Available online at www.sciencedirect.com
ScienceDirect
Electronic Notes in Theoretical Computer Science
Electronic Notes in Theoretical Computer Science 208 (2008) 95–111
www.elsevier.com/locate/entcs

# Performance of Reactive Interfaces in Stimulus Rich Environments, Applying Formal Methods and Cognitive Frameworks

Li Su[1,2], Howard Bowman[3]

*Centre of Cognitive Neuroscience and Cognitive Systems*
*University of Kent*
*Canterbury, United Kingdom*

Philip Barnard[4]

*MRC Cognition and Brain Sciences Unit*
*Cambridge, United Kingdom*

**Abstract**

Previous research has developed a formal methods-based (cognitive-level) model of the Interacting Cognitive Subsystems central engine, with which we have simulated attentional capture in the context of Barnard's key-distractor Attentional Blink task. This model captures core aspects of the allocation of human attention over time and as such should be applicable across a range of practical settings when human attentional limitations come into play. Thus, we have used this model to evaluate the performance trade-offs that would arise from varying key parameters in Stimulus Rich Reactive Interfaces. A strength of formal methods is that they are abstract and thus, the resulting specifications of the operator are general purpose, ensuring that our findings are broadly applicable.

*Keywords:* Formal methods, HCI, SRRIs, Attentional blink, ICS

## 1 Introduction

Applying formal modelling and cognitive frameworks to the design and analysis of interactive systems has gained much attention, particularly in the domain of mission and life critical applications. It has been argued by a number of researchers [12,8] that formal methods provide a powerful way to specify, evaluate, and verify such

[1] The work reported here was performed in the project Computational Modeling of Salience Sensitive Control in Humans and Artificial Systems, which was funded by the EPSRC under grant number GR/S15075/01.

[2] Email: ls68@kent.ac.uk

[3] Email: h.bowman@kent.ac.uk

[4] Email: philip.barnard@mrc-cbu.cam.ac.uk

1571-0661/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.entcs.2008.03.109

systems. As an important component in interactive systems, human operators and their cognitive capabilities also determine the success or the failure of such systems. Thus, attention needs to be paid to specifying and analysing the human cognitive system. In this paper, we argue that formal methods can specify and simulate cognitive frameworks as well as interaction between users and computer systems, allowing performance measurement, both qualitatively and quantitatively.

We are interested in a particular type of interactive system, *Stimulus Rich Reactive Interfaces* (SRRIs). In such systems, stimuli are presented rapidly to a human operator, as found in interfaces in cockpit environments or operations control rooms. A fundamental problem in perceiving stimuli in SRRIs is information overload. When capacity limited attentional mechanisms are overloaded, errors are likely to occur. Salient stimuli (targets [5]) are prioritised and processed more extensively than others (background items). In the context of SRRIs, stimuli may arrive so rapidly that they may interfere with each other. Thus, human operators could potentially miss some highly salient stimuli. It has been discovered that humans may miss a second target (T2) following a previous one (T1) after a few hundred milliseconds [23,4]. This temporal attentional limitation is supported by the results of a large number of psychological experiments, known as *Attentional Blink* (AB) tasks [23].

In order to explore the limits of temporal attention deployment in humans, we have developed a cognitive-level formal methods based model [26]. We have used a modelling technique called process algebra [9,18,21], which originated in theoretical computer science, being developed to specify and analyse distributed computer systems [9]. A process algebra specification contains a set of top-level subsystems (processes) that are connected by a set of communication channels. Subsystems interact by exchanging messages along channels. Furthermore, process algebra components can be arbitrarily nested within one another, allowing hierarchical decomposition in the manner advocated in [5]. The advantages of using formal methods in the context of cognitive modelling will be discussed shortly. The main purpose of our model was to simulate attentional capture in the context of a specific experimental paradigm [4]. Assuming that this model captures fundamental limitations on attentional deployment, it should be generalisable to practical settings that share its key properties. In this paper, we first outline the experimental paradigm and model. We then go on to examine in some detail how it could be extended to SRRIs.

Many real world applications require both high reliability in detecting targets and also that detection occurs in a timely fashion. This may lead to trade-offs in designing such interfaces. For example, a way to ensure that human operators have successfully perceived all targets is to ask them to acknowledge each target. Then, the system can re-present the missed targets. Alternatively, we can design the interface by considering the user's attention, so that it may avoid interference between targets. For example, we could separate targets by inserting blanks or background items. We call the first solution a reactive approach, and the second

[5] For simplicity of presentation, we use the term "target", although, more accurately, we should probably talk in terms of "potential targets", since, while the system knows what is likely to be salient, it is only the human operator that knows what is truly significant.

solution a constructive approach. However, there is a significant drawback with both approaches, that is, either acknowledging or inserting items may delay the presentation. In this paper, we will concentrate on illustrating how the modelling approach can provide insight into the design of constructive approaches, but will also explore some general implications for reactive solutions from the same theoretical position.

## 2 Formal Modelling of Cognitive Architecture

The general applicability of cognitive frameworks in the context of human-computer interaction has been demonstrated by the success of a number of broad cognitive theories, such as SOAR [22], ACT-R [1] and EPIC [20]. These are centralised production systems based models. Another cognitive theory is the *Interacting Cognitive Subsystems* (ICS) architecture [3]. This is a highly parallel architecture, in which control of system wide interactions is distributed rather than centralised. This theory assumes that our mental architecture is composed of nine subsystems:- 1) sensory subsystems: acoustic, visual, and body-state; 2) effector subsystems: articulatory and limb; 3) central subsystems: morphonolexical, object, propositional, and implicational.

Each subsystem has the same internal structure. Information arrives in a subsystem, is copied into its local memory, called an image record, and is transformed for use in other subsystems. The subsystems differ in their inputs and outputs, that is, they are each specialised for storing and processing a qualitatively different form of mental representation [3]. ICS stresses that mental activity concurrently occurs in multiple domains (realized by their corresponding subsystems) at the same time. Although each subsystem in ICS has a restricted functionality, much richer mental behaviour can arise out of the interaction among multiple subsystems. It will become clear later that our model can be fitted into the framework of ICS, but a complete description of ICS is outside the scope of this paper.

Many broad cognitive frameworks are symbolic in nature, since there are a number of attractions for using symbolic models instead of connectionist networks in formalizing large scale mental architectures [5]. Firstly, symbolic systems are good at manipulating, explaining and reasoning about complex data structures. Secondly, the knowledge or rules are visible. Thirdly, large systems can be composed from small structures. Finally, symbolic representations facilitate compositional and hierarchical representation of knowledge.

Broad theories, such as ICS, also present a challenge for computational modelling. For example, ICS and a large number of other psychological theories are presented using "box and arrow diagrams". Thus, the computational models must capture the interactions between mental subsystems and do it at an abstract level. Although connectionist networks are commonly regarded as concurrent and distributed, they are typically limited to only one level of concurrently evolving modules: the primitive elements of neural networks are neurons but not neural networks. To some degree, it is hard to construct and understand large architectures without

hierarchical structuring. In certain respects, modelling based on neural networks is low-level in character, i.e. it is hard to relate to primitive constructs and data structures found in high-level notations preferred by the symbolic modelling community.

Barnard and Bowman [5] have argued that the mathematical models developed by computer scientists, such as process algebra [9], offer the prospect to directly model "box and arrow diagrams". For example, in their process algebraic models, the "boxes" are modelled by processes and the "arrows" are modelled by communication channels. Issues in developing such broad models include the *irrelevant specification problem*[22], which is caused by a large number of assumptions being made during implementation. However, it is sometimes unclear what assumptions correspond to known cognitive behavior. Hence, we need to limit such overspecification.

We argue that the requirements for modelling such mental theories, in particular ICS, are similar to the requirements for modelling distributed computer systems. This is because control is distributed in ICS: subsystems are independent components, which interact through exchange of data representations over communication channels [3,8,5]. ICS asserts that cognition emerges as the product of the interaction between a set of autonomous subsystems. The advantages of process algebra are that modelling is at an abstract level, but the ability to execute is preserved. Hence, process algebra has already been applied to modelling ICS [8,5].

## 3 Temporal Bottleneck in Attention

In the context of this paper, we first explore the temporal attentional limitation of human users, and then go on to develop arguments about how this affects their ability to interact with computer systems. Raymond et al [23] designed an AB task, which involves letters being presented at the same spatial location using *Rapid Serial Visual Presentation* (RSVP) at around 10 items per second. One letter (T1) was presented in a distinct colour and was the target whose identity was to be reported. A second target (T2) followed after a number of intervening items. Typically, participants had to report whether the letter "X" was among the items that followed T1. This condition is called the blink condition, where subjects have to report the identity of both T1 and T2. The key finding was that selection of T2 was impaired with a characteristic serial position curve; see curve with unfilled squares in Figure 1, where lag indicates the number of distractors intervening between T1 and T2, i.e. lag-1: T2 immediately follows T1; lag-2: one intervening distractor, etc. T2s occurring immediately after T1 were accurately detected (a phenomenon typically described as lag-1 sparing). Detection then declined across serial-positions 2 (and also sometimes 3) and then recovered to baseline around lags 5 or 6 (corresponding to a target onset asynchrony in the order of 500 to 600 ms). However, subject's ability to report the identity of T2 when T1 being absent is unaffected by the serial position of T2. This condition is called the baseline condition; see the curve with filled diamonds in Figure 1.

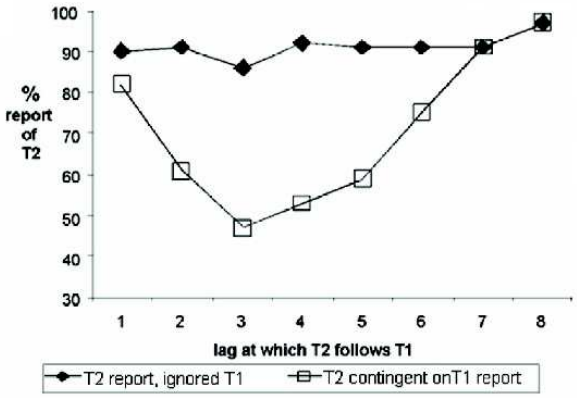


Fig. 1. The basic AB effect for letter stimuli [23]. Here, the blink condition (unfilled squares) is shown as T2 accuracy conditional on T1 report, reflecting the effect on T2 report of successfully attending to T1. Baseline (filled diamonds) represents a person's ability to report the presence of T2 when T1 was absent.

As research on the blink and RSVP in general has progressed, it has become evident that the allocation of attention is affected by the meaning of items [19] and their personal salience [24]. There is also evidence from electrophysiological recording that the meaning of a target is processed even when it is not reported [25]. In addition, there are now reports of specific effects of affective variables, e.g. [6]. In particular, [2] has shown that the blink is markedly attenuated when the second target is an aversive word.

These AB experiments have their counterpart in real life. For instance, when driving a car at 70 mph on the motorway, the driver monitors the traffic ahead and signposts. Events appear continuously and rapidly as those in RSVP streams, but they are somewhat inter-related. When a hazard occurs, e.g. another car suddenly changes lanes (cuts in front of you), it intrudes into this constant schema. These are high salient stimuli (targets in AB) for safe driving, but attending to such hazards (maybe reacting by breaking) could potentially knock out the driver's attention for about 500 ms. Under this situation, a driver is more likely to miss an important signpost or a junction. So, the robust AB serves as a useful paradigm to be generalised from lab to practical settings, and the semantic salience is particularly relevant because most real world tasks relate to meaning or executive function.

In order to examine semantic effects, Barnard et al [4] used a variant of the AB paradigm in which words were presented at fixation in RSVP format, at around 10 items per second. Targets were only distinguishable from background items in terms of their meaning. Participants were simply asked to report a word if it refered to a job or profession for which people get paid, such as waitress, and these targets were embedded amongst background words that all belonged to the same category, e.g. nature words. However, streams also contained a key-distractor item, which, although not in the target category, was semantically related to that category; see Figure 2. The serial-position that the target appeared after the key-distractor was varied. The effect of attentional capture by meaning is encapsulated in the serial position curve (unfilled diamonds) in Figure 3. That is, the key-distractor drew attention away from the target with a clear temporal profile. In the next section, a

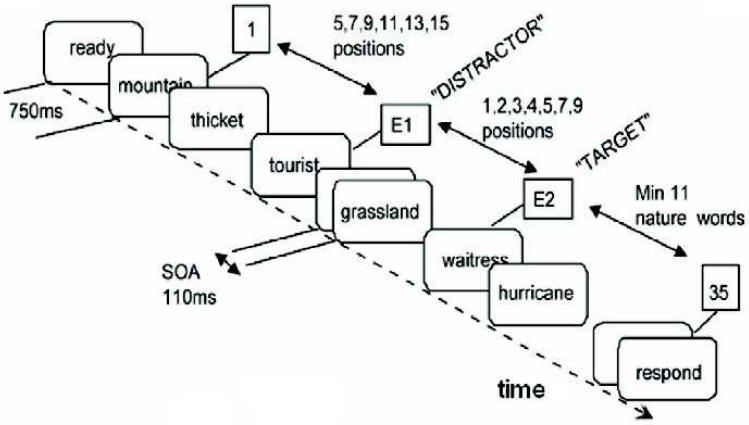


Fig. 2. Task schema for the key-distractor blink; adapted from [4].

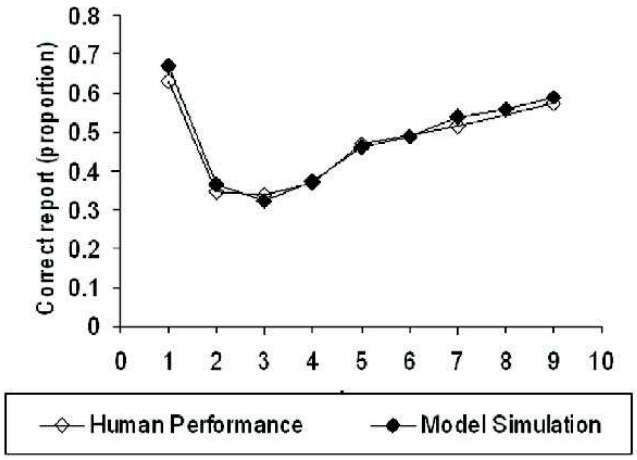


Fig. 3. The attentional capture by meaning effect in humans [4] and model simulations [26], target report accuracy by lag of target relative to the position of the key-distractor.

cognitive model of human temporal attention is described. The simulation result of this model is the serial position curve (filled diamonds) in Figure 3, which can be compared to the performance of real participans shown in the same graph.

# 4 Cognitive model of Human Operators

We have provided a concrete account of attentional capture by meaning and the temporal dynamics of that process [26]. Key principles that underlie this account are sequential processing, 2-stages, and serial allocation of attention. We discuss these principles in turn.

## *4.1 Sequential Processing*

With any RSVP task, items arrive in sequence and need to be correspondingly processed. Thus, we require a basic method for representing this sequential arrival

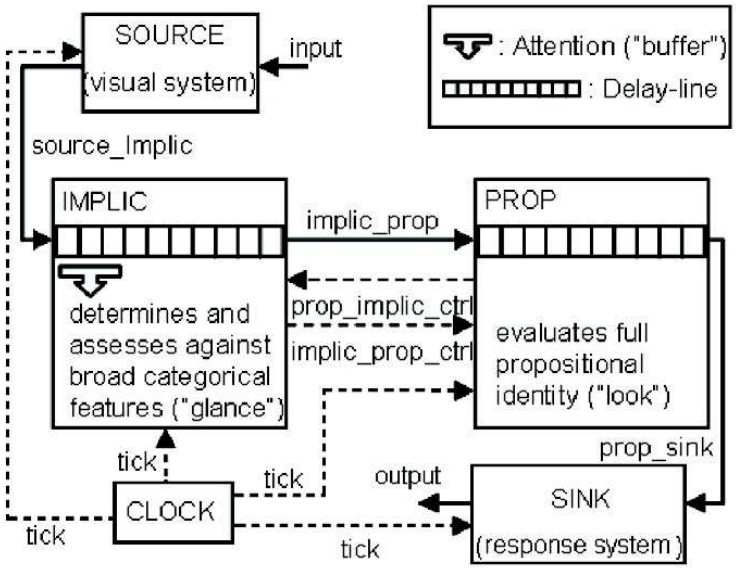


Fig. 4. Top-level structure of the two subsystems model with implicational subsystem attended.

and processing of items. At one level, we can view our approach as implementing a pipeline. New items enter the front of the pipeline from the visual system; they are then fed through until they reach the back of the pipeline, where they enter the response system [6]. Every cycle, a new item enters the pipeline and all items currently in transit are pushed along one place. The key data structure that implements this pipeline metaphor is a delay-line as shown in Figure 4. It could also be viewed as a symbolic analogue of a sequence of layers in a neural network. It is a very natural mechanism to use in order to capture the temporal properties of a blink experiment, which is inherently a time constrained order task.

### *4.2 2-Stages*

Like [10,11], [4] argued for a two-stage model, but this time recast to focus exclusively on semantic analysis and executive processing. In particular, [5,26] modelled the key-distractor blink task using a two-stage model. In the first stage, a generic level of semantic representation is monitored, and initially used to determine if an incoming item is salient in the context of the specified task. If it is found to be so, then, in the second stage, the specific referential meaning of the word is subjected to detailed semantic scrutiny; thus, a word's meaning is actively evaluated in relation to the required referential properties of the target category. If this reveals a match, the target is encoded for later report. The first of these stages is somewhat akin to first taking a "glance" at generic meaning, with the second akin to taking a closer "look" at the relationship to the meaning of the target category. These two stages are implemented in two distinct subsystems, as shown in Figure 4: the implicational subsystem or Implic and the propositional subsystem or Prop [3].

These two subsystems process qualitatively distinct types of meaning. One,

[6] Note, visual system and response system do not directly map to ICS subsystems. They are abstract representations of a number of subsystems necessary to perceive stimuli and to make responses.

implicational meaning, is holistic, abstract and schematic, and is where affect is represented and experienced [3]. The other is "rational", being based upon propositional representation, capturing referentially specific semantic properties and relationships. Semantic errors make clear that sometimes we only have referentially non-specific semantic *gist* information available, e.g. the Noah illusion illustrates implicational meaning [14]. That is, in a Noah specific sentence, such as "How many animals of each kind did Moses take on the Ark?" most people respond "two" even though they know that it was Noah, not Moses, who took the animals on the Ark. Substitution of Moses for Noah often fails to be noticed, while substitution with Nixon, or even Adam, is noticed. This is presumably because both Moses and Noah fit the generic (implicational) schema "aged male biblical figure"[4], but Nixon and Adam do not. To tie this into the previous principle, Implic and Prop perform their corresponding salience assessments as items pass through the pipeline.

### *4.3 Serial Allocation of Attention*

Our third principle is a mechanism of attentional engagement. It is only when attention is engaged at a subsystem that it can assess the salience of items passing through it. Furthermore, attention can only be engaged at one subsystem at a time. Consequently, semantic processes cannot glance at an incoming item, while looking at and scrutinising another. It is worth noting that this is a key aspect of distributed control, i.e. limitations arise because the system can effectively look at only one source at a time. This constraint will play an important role in generating a blink in our models. When attention is engaged at a subsystem, we say that it is buffered [3]. (In the context of this paper, the term buffer refers to a moving focus of attention.) Thus, salience assignment can only be performed if the subsystem is buffered and only one subsystem can be buffered at a time, as shown in Figure 4. The buffer mechanism ensures that the central attentional resources are allocated serially, while items pass concurrently, i.e. all items throughout the delay-line are moved on one place on each time step.

As we have previously mentioned, the model presented here can be placed within the context of ICS, i.e. distributed control is inherent in ICS, and both the delay-line and buffering concepts that we use have their roots in ICS. However, most significantly, the implicational - propositional distinction reflects ICS' dual-subsystem central engine [27].

### *4.4 How the Model Blinks*

The buffer movement dynamic provides the underlying mechanism that causes the blink. Initially, Implic is buffered as shown in Figure 4. When, in response to the key-distractor being found implicationally salient, the buffer moves from Implic to Prop, salience assessment cannot be performed on a set of words (i.e. a portion of the RSVP stream) entering Implic following the key-distractor. So, when these implicationally uninterpreted words are passed to Prop, propositional meaning (which builds on implicational meaning) cannot be assessed. Target words falling within

this window will not be detected as implicationally salient and thus will not be reported.

There is normally lag-1 sparing in key-distractor AB experiments, i.e. a target word immediately following the key-distractor is likely to be reported. This arises in our model because buffer movement takes time, hence, the word immediately following the key-distractor may be implicationally interpreted before the buffer moves to Prop.

When Prop is buffered and detects an implicationally uninterpreted word, the buffer is passed back to Implic, which can assign salience again. After this, target words entering the system will be detected as implicationally and propositionally salient and thus will be reported. Hence, the blink recovers.

### *4.5 LOTOS Specification*

The top-level model shown in Figure 4 is an example of a "box and arrow diagram" as previously discussed. Thus, it is informal and not executable. In order to further explain the AB phenomenon and constraints on the interaction of these subsystems, such informal theories have been realised as computational models. As previously discussed, we used process algebra to model cognitive frameworks. The particular specification language used here is called LOTOS [7]. The top-level specification is shown as follows:

```
specification keyDisAB[para1,para2,input,output]:noexit
library keydisab endlib
behaviour
para1?implic_buf_mov:Nat; para2?prop_buf_mov:Nat;
hide source_implic,prop_sink,implic_prop,implic_prop_ctrl,
prop_implic_ctrl,tick in
(((IMPLIC[source_implic,implic_prop,implic_prop_ctrl,
    prop_implic_ctrl,tick](implic_buf_mov,true)
      |[implic_prop,implic_prop_ctrl,prop_implic_ctrl,tick]|
PROP[implic_prop,implic_prop_ctrl,prop_implic_ctrl,prop_sink,tick]
    (prop_buf_mov,false))
      |[source_implic,prop_sink,tick]|
(SOURCE[input,source_implic,tick]
      |[tick]|
SINK[prop_sink,output,tick](mk(<>,U,U))))
      |[tick]|
CLOCK[tick](0 of Nat))
where (* Specification of subsystems *)
library  implic,prop,source,sink,clock endlib
endspec
```

In the above specification, names in uppercase are processes, e.g. `IMPLIC`, names in lowercase are communication channels, e.g. `implic_prop`. `IMPLIC`, `PROP`, `SOURCE`, and `SINK` represent implicational subsystem, propositional subsystem, visual sys-

tem, and response system respectively. Each process has a set of communication channels attached to it, for example, `IMPLIC` has five channels. `A|[`$a_1, a_2, ..., a_n$`]|B` denotes a parallel composition of process `A` and `B`, which synchronize on channel $a_1, a_2, \ldots, a_n$.

Input channels `para1` and `para2` set the simulation up with appropriate parameters, which were determined from our previous modelling of the AB task [26]. Input channel `input` connects stimuli from the SRRIs to the user model. In each run of the simulation, an entire presentation is randomly generated and fed into the simulation. A run stops by creating a *deadlock* in `SOURCE` when all items have been presented. The output channel `output` initiates a subsequent response[7]. All subsystems are synchronized by a global clock `CLOCK`, which ticks every 20ms of simulated time.

It can be seen that the LOTOS specification has a similar overall structure and interaction as our "box and arrow diagram" shown in Figure 4. The internal structure of each process was also modelled as reusable library files, however, detailed descriptions of psychological processes is beyond the scope of this paper.

The LOTOS model was simulated using the CADP toolbox [15]. The specification of the data types written in ACT ONE [13] were transformed to C code using the EXEC/CAESAR environment [16]. Psychological simulations often obtain good results from the average of thousands of runs. So, C code generated using EXEC/CAESAR can be performed efficiently. The results of the simulation were compared to human performance in order to verify our theories of temporal attention [26]; see Figure 3.

## 5 Models of SRRIs

As previously argued that the AB paradigm is generlisable to SRRIs, we believe that we can make the following assumptions in simulating SRRIs:- 1) all items are presented for 120ms, 2) the interface presents a maximum of 200 items (this constrains the user to respond to stimuli in a timely fashion), 3) items may be targets, background items or blanks, 4) items appear at the same spatial location, 5) stimuli may have observable bursts on all time scales (sometimes called self-similar traffic[8]).

There are a number of methods to generate self-similar traffic, such as the multiple ON/OFF source aggregation process [17]. However, many of these are complex and hard to implement. We choose a particular model, called the $b$-model [28]. It is a simple model with very few parameters, but it can generate self-similar and bursty traffic for any given time scale. Its parameters include the number of targets, the number of items (both targets and background items), and the burstiness of the traffic, which is characterised using a single parameter $b$ that ranges between 0 and

[7] In reality, the interface has no access to this output signal. Indeed, the output is used by the simulation environment to generate a report. Channels `input` and `output` are connected to the simulation environment implemented in C code.

[8] There are many models of traffic proposed by computer scientists, such as Poisson arrivals and self-similar models. The traditional Poisson arrivals is not used here, since it is only bursty within a short time scale, but it will smooth out if it is applied to large time scales.

0.5. The smaller $b$ is, the more bursty the stimuli. A $b$ of 0.5 results in items being evenly distributed in the stream. A complete description of the model is presented in [28]. Intuitively speaking, the $b$-model randomly divides events into two halves and distributes them into two equal periods of time. This process repeats on each of these two parts of the presentation until all events have been distributed.

We would like to use the $b$-models to simulate how stimuli appear in SRRIs. However, such stimuli are different from those arising from RSVP streams, since multiple events/stimuli can arrive simultaneously. Thus, we assume that SRRIs have a buffer [9], which stores all events and presents them to the user serially. This assumption ensures stimuli appear in RSVP fashion, restricting our research to the area of temporal attention. The total number of items (background and targets) to be presented is $2^n$, where $n$ is a parameter in the $b$-model, called the aggregation level. Intuitively speaking, the aggregation level determines the length of the traffic, i.e. aggregation level 2 generates 4 items; aggregation level 6 generates 64 items, which contains both targets and background items. An example of such stimuli is shown in Figure 5(a). Note, the interface outputs blanks when all items have been presented.

This interface was composed with the user model described in the previous section. Then, we vary a number of key parameters of the interface, and measure the probability of reporting target identity. Our model of the AB task suggests that the ability of reporting the subsequent item is significantly impaired between 100ms and 600ms after the onset of the previous salient item. Hence, an approach to improve performance is to separate salient items, so that their interference is minimised. Such a constructive approach ensures that interfaces leave a window between any two targets by inserting blanks. The window size is measured by the number of blank items inserted. Figure 5(b) illustrates an example of AB-aware system with blink window size 5. That is, between every two adjacent targets, there are minimum 5 blanks. A typical window size is 5, which corresponds to 600ms. We call such interfaces AB-aware systems, and the original one AB-unaware; see Figure 5.

## 6 Performance Evaluation

We now present the simulation results, focusing on the performance of the interface measured as the probability of seeing a target. In the first experiment, both the burstiness and the number of targets were varied. In the second experiment, both the burstiness and the aggregation level (total number of targets and background items) were changed. Finally, we explore how the window size affects the effeciency of the interface.

[9] Using the standard computer science meaning here.

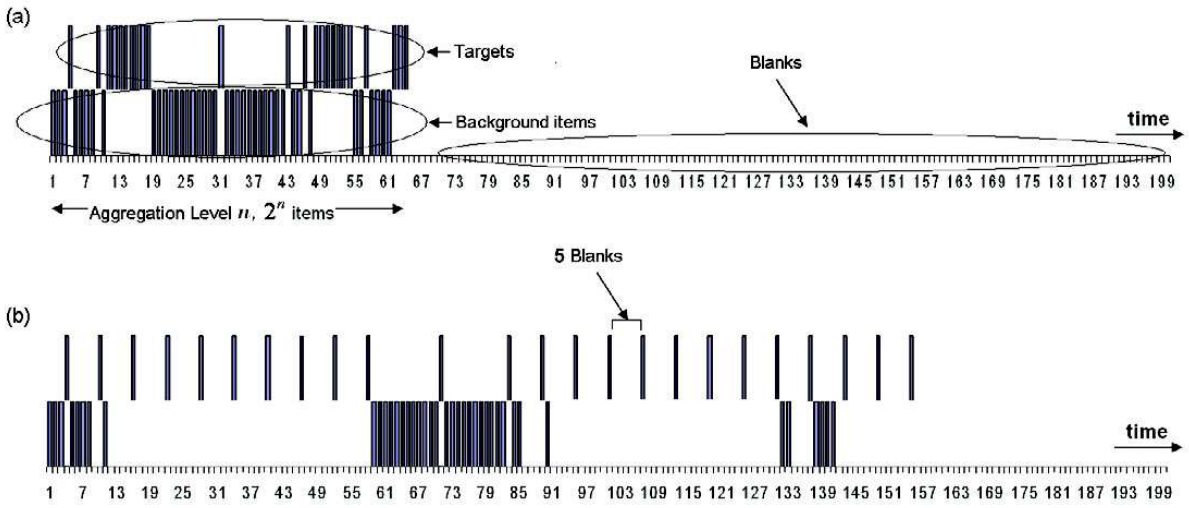


Fig. 5. Examples of stimuli in SRRIs, with time expressed in terms of the number of items. Vertical bars on the top half of each graph represent targets. Bars on the bottom half represent background items. Empty spaces represent blanks in SRRIs. (a) an AB-unaware system, (b) an AB-aware system.

### *6.1 Experiment 1*

In this experiment, we fix the aggregation level to 6, window size to 5, and vary the burstiness and the number of targets. Figure 6(a) shows that the performance of both systems drops as the number of targets increases. This is because salient items tend to be close together when the number of targets is high. In general, the performance also decreases when the traffic becomes burstier, since targets are closer to each other and interfere with each other more often. In order to compare these two systems in detail, we subtracted the performance of the AB-unaware system from the AB-aware one. As shown in Figure 6(a), the AB-aware system is less affected by the burstiness because the AB-aware system can reduce bursts and smooth the traffic. It effectively acts as a buffer [10], maintaining space between salient items. When the number of targets is low (below 20), targets are more separated, thus the difference between the two systems is less significant. However, we can see that the AB-aware system is significantly superior when the number of targets is medium (between 20 and 40). The benefits of AB-aware systems start to drop when the number of targets is 40, since the AB-aware system is unable to present all targets within the 200 item's time limit, and keep a sufficient gap (5 blanks in the current setting) between targets. Moreover, the AB-aware system starts to perform worse than the AB-unaware system when the number of targets is high (above 70). This reflects the effects of aggregation level on the performance. So, we investigated this parameter in the next experiment.

### *6.2 Experiment 2*

In this experiment, we fixed the number of targets to 20, window size to 5, and varied the burstiness and the aggregation level. Figure 6(b) shows that the performance of both systems improves as the aggregation level increases. This is because the targets

[10] Using the standard computer science meaning here.

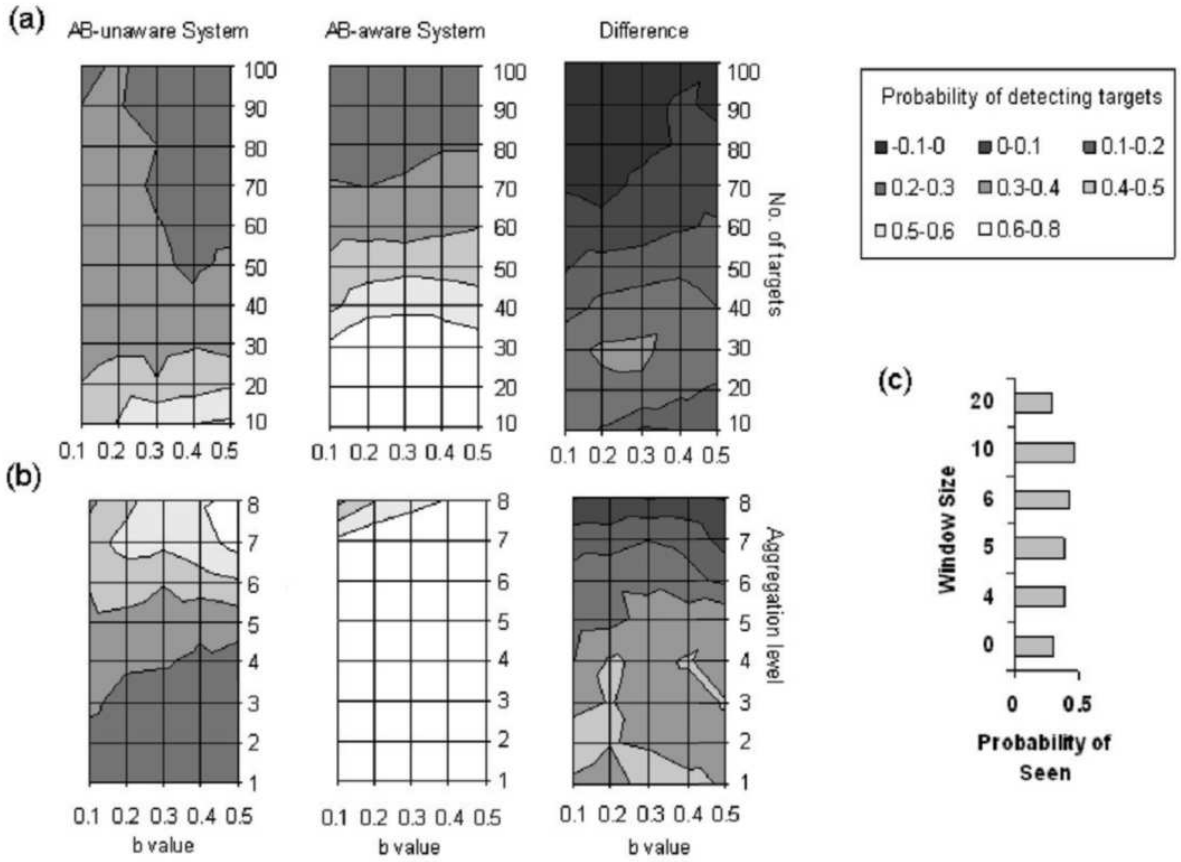


Fig. 6. (a,b) Performance (measured as probability of detecting the targets) of AB-unaware and AB-aware systems by varying the number of targets, the aggregation level, and the burstiness (i.e. the $b$ value, with burstiness increasing as $b$ decreases). (c) Performance of SRRIs with different window sizes of the stimuli. AB-unaware system is a special case of AB-aware system with a window size of 0.

are more separated when the aggregation level is high. Performance drops when the stimuli are more bursty. Similar to the previous section, difference in performance is shown in Figure 6(b). It can be seen that the AB-aware system consistently performs better than the AB-unaware system. Note that the performance did not improve continuously as the aggregation level increases, that is, it starts to drop at aggregation level 8. This is because our simulation only presents a maximum of 200 items, thus the rest will be discarded. Unlike the previous experiment, the AB-aware system is also sensitive to the burstiness of the stimuli as we vary the aggregation level. This is because increasing the length of the traffic reduces the blank period at the end of the traffic (see Figure 5 for how the AB-aware system uses the blank period), leaving the AB-aware system less space to arrange targets.

### *6.3 Experiment 3*

In this experiment, we repeated the previous two experiments, but this time using different blink window sizes of 4, 5, 6, 10 and 20. Figure 6(c) shows the performance of the AB-unaware system (which can be seen as a special case of the AB-aware system with a window size of 0) and AB-aware systems using different window sizes. A trade-off can be found. That is, if we want to ensure that the human user perceives targets as accurate as possible by avoiding interference between targets,

we should increase the window size. However, large window sizes potentially result in fewer targets being presented within a bounded time. Hence, the balance between accuracy and urgency must be considered.

## 7 Conclusions

We varied a number of parameters in SRRIs, and evaluated the performance trade-offs that arise in AB-aware systems. They could reduce the effect of burstiness by smoothing the traffic, and improving the probability of reporting targets by sacrificing urgency. The disadvantage is that they could potentially delay the presentation of targets thereby impairing performance (in particular when the aggregation level is high), and breakdown the absolute timing of events, making some presentations unintelligible. We also noticed that the blink window size affects the performance of the AB-aware system, and a medium size is suitable for the type of stimuli used here.

Figure 6 shows that the biggest probability of detecting a target is around 0.6 in all experiments, which is close to the baseline performance of AB experiments. The smallest probability of detecting a target is around 0.2, which is similar to the worst performance during the blink. This reflects the fact that AB-aware systems can direct a user's attention, therefore improving the performance, but they are unable to exceed the maximum attentional capacity of the user. In other words, a fundamental limitation of such an approach is that an AB-aware system cannot completely remove the noise of the biological and mental processes as reflected by the baseline performance in the AB. Note, 60 percent chance of detecting a target is mainly due to high presentation rate and masking after the target, so it may not be really representative of real applications. However, with only one target and no masking after the target, it would rise to almost 90 percent.

## 8 Discussion and Future Work

As our first step to applying formal methods in interactive systems that consider the attentional state of users, we have designed a model of SRRIs based on studies of the AB phenomenon. AB-aware systems proposed here could avoid presenting targets when human operators are not ready to perceive them. However, a significant cost is that the presentation rate may be reduced dramatically. Another observation from these experiments is that such a constructive approach cannot improve the performance beyond the baseline condition, i.e. when there is only one target in an RSVP stream.

Hence, future research may consider reactive approaches to improving SRRIs, for example, using the electroencephalogram (EEG) as an acknowledgement from the human operator. Such an approach may provide faster feedbacks than explicitly asking user to acknowledge. Wyble et al [29] have performed a feasibility investigation on SRRIs using EEG. They have discovered that two potential EEG measures (reduced EEG power in the alpha band at posterior brain areas and a

P3-like deflection over parietal areas) may be correlated with whether a stimulus has been perceived. Critically, they have investigated the possibility of extracting these signals in real-time using compact devices implemented and integrated in a head-mounted display. It is possible that a reactive approach could improve performance above baseline. This is because such systems can identify missed targets and re-present them, while in constructive approaches, the user has only one chance to perceive each stimulus. It was also noticed that such EEG feedback may not always accurately predict whether the user has perceived targets. Hence, we could extend our formal model to produce feedback in a similar fashion as EEG signals and use them to investigate the feasibility of such interfaces. We could, for example, evaluate how accurate the acknowledgement should be, in order to make such an approach worthwhile.

Finally, we should also notice that most AB experiments have very small numbers of targets. However, in the case of SRRIs, human users have to cope with much larger numbers of targets. Thus, they will almost certainly exceed working memory capacity. We have ignored this issue in our studies, but it could have a dramatic influence on performance. The effect of such cognitive load awaits future investigation. Also, effects of context or of schematic fit may be important. Note, for example that AB experiments use discrete items, whereas in many real SRRIs visually presented items are going to be part of a more or less predictable knowledge structure. Indeed, this may be a really interesting point to pursue in the future. Do humans pick up schema discrepant material more readily than schema conforming stimuli?

In summary, this model is a detailed psychological description of human temporal attention mechanism. It has been demonstrated elsewhere [26] that this model accurately reproduces a large set of behavioural data in Barnard's key-distractor AB task. Hence, we believe that this model should, to some degree, help us to understand and predict the effect of temporal bottleneck of attention on SRRIs. However, validation of our observations / predictions from such abstract model requires highly intensive experiments of real participants. It should also be noticed that we have made a number of assumptions about the SRRIs in order to simplify them to RSVP streams. Although we tried to minimise our assumptions, many aspects of system design have been ignored in the current study, for example, spatial attention, contextual influence, multimodal input, and so on. Hence, we hope future studies will address these interesting issues.

# References

[1] Anderson, J. R., "Rules of the Mind", Hillsdale, NJ, Erlbaum, 1993

[2] Anderson, A. K., *Affective influences on the attentional dynamics supporting awareness.* J. Exp Psychol Gen, **134**(2), (2005), 258-281.

[3] Barnard, P.J., *Interacting Cognitive Subsystems: modelling working memory phenomena within a multi-processor architecture.* Models of Working Memory: Mechanisms of active maintenance and executive control (1999), 298-339.

[4] Barnard, P.J., S. Scott, J. Taylor, J. May, and W. Knightley, *Paying Attention to Meaning*, Psychological Science, **15**(3).(2004), 179-186.

[5] Barnard P.J. and H. Bowman, *Rendering Information Processing Models of Cognition and Affect Computationally Explicit: Distributed Executive Control and the Deployment of Attention*. Cognitive Science Quarterly, publisher: Hermes Science Publications, **3**(3), (2004), 297-328.

[6] Barnard, P. J., C. Ramponi, G. Battye, and B. Mackintosh, *Anxiety and the Deployment of Visual Attention over Time*. Visual Cognition, **12**(1), (2005), 181-211.

[7] Bolognesi, T., and E. Brinksma, *Introduction to the ISO Specification Language LOTOS*. Computer Network and ISDN Systems, **14910** (1988), 25-29.

[8] Bowman H. and G. Faconti, *Analysing cognitive behaviour using LOTOS and Mexitl*. Formal Aspects of Computing, **11** (1999), 132-159.

[9] Bowman, H., and R. S. Gomez, "Concurrency theory, calculi and automata for modelling untimed and timed concurrent systems", Springer, 2006.

[10] Bowman H. and B. Wyble, *The simultaneous type, serial token model of temporal attention and working memory*. Psychological Review, **114**(1), (2007), 38-70.

[11] Chun, M. M., and M. C. Potter, *A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation*. Journal of Experimental Psychology: Human Perception and Performance, **21**(1), (1995), 109-127.

[12] Duke, D. J., P. J. Barnard, D. A. Duce, and J. May, *Syndetic Modelling*. Human-Computer Interaction, **13**(4), (1998), 337-393.

[13] Ehrig, H., W. Fey, and H. Hansen, "ACT ONE - An Algebraic Specification Language with two Levels of Semantics", ADT, 1983.

[14] Erickson, T. D., and M. E. Mattson, *From words to meaning: a semantic illusion*. Journal of Verbal Learning and Verbal Behavior, **20** (1981), 540-551.

[15] Garavel, H., F. Lang, and F. Mateescu, *An overview of CADP 2001*, EASST Newsletter, **4** (2002):13-24.

[16] Garavel, H., C. Viho,, and M. Zendri, *System Design of a CC-NUMA Multiprocessor Architecture using Formal Specification, Model-Checking, Co-Simulation, and Test Generation*. Springer International Journal on Software Tools for Technology Transfer (STTT), **3**(3), (2001), 314-331.

[17] Gomez, M.E., and V. Santonja, *Self-similiary in I/O workload: Analysis and modeling*. In Workshop on Workload Characterization, (1998).

[18] Hoare, C. A. R., "Communicating Sequential Processes". London: Prentice-Hall, 1985.

[19] Maki, W. S., K. Frigen, and K. Paulsen, *Associative Priming by Targets and Distractors During Rapid Serial Presentation*. Journal of Experimental Psychology: Human Perception and Performance, **23** (1997), 1014-1034.

[20] Meyer, D. E. and D. E. Kieras, *A computational theory of executive cognitive processes and multiple task performance: Part 1. Basic mechanisms*, Psychological Review **104** (1997), 3-65.

[21] Milner, R., "Communication and Concurrency". Hemel Hempstead, UK: Prentice-Hall, 1989.

[22] Newell, A., "Unified Theories of Cognition". Cambridge, Massachusetts: Harvard University Press, 1990.

[23] Raymond, J., K. Shapiro, and K. M. Arnell, *Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink*, JEP:HPP, **18**(3), (1992), 849-860.

[24] Shapiro, K. L., J. I. Caldwell, and R. E. Sorensen, *Personal names and the attentional blink: The cocktail party revisited*. Journal of Experimental Psychology: Human Perception and Performance, **23** (1997), 504-514.

[25] Shapiro, K. L., and S. J. Luck, *The attentional blink: A front-end mechanism for fleeting memories*. In Fleeting memories, Cognition of Brief Visual Stimuli. Boston, Massachusetts: A Bradford Book, MIT Press, (1999), 95-118.

[26] Su, L., H. Bowman, and P. J. Barnard, *Attentional Capture by Meaning, A Multi-level Modelling Study*, In D. S. McNamara and J. G. Trafton (Eds.), Proceedings of the 29th Annual Cognitive Science Society. Austin, TX: Cognitive Science Society, (2007), 1521-1526.

[27] Teasdale, J.D., and P.J. Barnard, "Affect, Cognition and Change: re-modelling depressive thought", Hove, Lawrence Erlbaum Associates, 1993.

[28] Wang, M., T. Madhyastha, N.H. Chan, S. Papadimitriou, and C. Faloutsos, *Data Mining Meets Performance Evaluation: Fast Algorithms for Modeling Bursty Traffic*. 18th International Conference on Data Engineering, (2002).

[29] Wyble, B., P. Craston, and H. Bowman. Electrophysiological feedback in adaptive human computer interfaces. Technical Report 8-06, Computing Lab, University of Kent at Canterbury, (2006).