The Simultaneous Type, Serial Token Model of Temporal Attention and Working Memory

Howard Bowman and Brad Wyble

Center for Cognitive Neuroscience and Cognitive Systems, University of Kent, Canterbury, Kent, CT2 7NF, United Kingdom, Phone: +44-1227-823815, Fax: +44-1227-762811 (H.Bowman,B.Wyble)@kent.ac.uk

Abstract

A detailed description of the Simultaneous Type Serial Token (ST^2) model is presented. ST^2 is a model of temporal attention and working memory, which encapsulates five principles: 1) Chun and Potter's (1995) 2-stage model; 2) a stage one Salience Filter; 3) Kanwisher's Types-tokens distinction; 4) a Transient Attentional Enhancement; and 5) a mechanism for associating types with tokens called the Binding Pool. We instantiate this theoretical position in a connectionist implementation, called Neural-ST^2, which we illustrate by modeling temporal attention results, focused on the Attentional Blink (AB). We demonstrate that the ST^2 model explains a spectrum of AB findings. Furthermore, we highlight a number of new temporal attention predictions arising from the ST^2 theory, which we test in a series of behavioral experiments. Finally, we review major AB models and theories, and compare them to ST^2.

Keywords

Temporal Attention, Working Memory, Attentional Blink, Neural Networks

Introduction

1 Temporal Attention

1.1 Background

The study of *attention* (Driver, 2001; Logan, 2004) is focused on understanding how the brain identifies salient items from amongst a set of competing stimuli. The majority of visual attention research has considered how selection occurs when the competing stimuli satisfy two criteria: 1) they are spatially offset from one another and 2) they are presented simultaneously (Duncan, 1981; Eriksen & Eriksen, 1974; Triesman, 1998). However, attentional processing still obtains even if criteria 1) and 2) do not apply. Indeed, the problem of how salient items are selected when the competing stimuli are offset in time, rather than space, must also be solved by a cognitive system. However, although there is increasing investigation of *temporal attention*, it remains, especially from theoretical and computational standpoints, considerably less explored than spatial attention.

When temporal attention is explored, it is natural to present stimuli briefly (typically for around 100ms) with each being replaced by the next at the same spatial location, generating what is called Rapid Serial Visual Presentation (RSVP) (Chun & Potter, 1995; Kanwisher, 1987; Raymond, Shapiro, & Arnell, 1992; Shih & Sperling, 2002). At this rate of presentation, performance is below ceiling, permitting investigation of factors affecting perception of targets. Furthermore, since all items are presented at the same spatial location, each item (apart from the first) masks the previous item (Giesbrecht, Bischof, & Kingstone, 2003; Giesbrecht & Di Lollo, 1998). Thus, a key characteristic of explorations of temporal attention is that representations of stimuli are fleeting (i.e. only briefly available to the retina).

In addition, while the question of how the brain detects a single salient item in an RSVP stream is interesting, to reveal the characteristics of temporal attention, streams containing multiple salient items need to be explored. This allows a number of key questions to be investigated. For example, how long is attention allocated to an initial target before it is free to be allocated to a second; can a salient item interrupt processing of an earlier item and cause attention to be redirected; and what constitutes salient in this context: relevance to the task at hand, emotional significance, similarity to a target category, etc?

The brain not only enables selective attention to salient items, but also enables higher-level cognition (e.g. executive processes, reasoning, language comprehension) to act upon attended items. The basis of such processing is the capacity to encode attended items into Working Memory (WM). Indeed, when considering temporal attention, the issue of WM encoding is particularly pivotal since, not only does this process involve storage of items, it also supports encoding of information about the temporal context in which an item occurred. With RSVP streams, assigning such *episodic contexts* is particularly difficult, since it requires the stream to be "individuated" through time; that is, a discrete episode needs to be associated with the occurrence of each item that is encoded into working memory.

In support of this position, it is clear that humans can report episodic information about when and how events occurred. In particular, they can report the order in which events arose. Furthermore, temporal order is a key information-bearing dimension, the extraction and recording of which is critical to high-level cognition. Linguistic abilities make this clear, e.g., word order in the sequence "John hit Jane" is central to attribution of meaning. Indeed, one of our main theses in this article is that, the human cognitive system is good at "filtering" non-salient items during visual processing and that thus the main constraint on the capacity to direct attention through time is the "cost" incurred in associating episodic contexts while encoding items into working memory.

The basic problem this paper focuses on is, given a sequence of fleeting representations of stimuli, how is it that the brain identifies the stimuli that are salient, sustains them and encodes them into Working Memory (WM). Central to this process is the allocation of discrete episodic contexts during encoding. Specifically, we will devise a comprehensive theoretical proposal for this process, which is encapsulated as the Simultaneous Type Serial Token (STST or ST^2) Model. This theory is realized in a connectionist model, Neural-ST^2, that is constrained both by behavioral data on RSVP tasks (particularly the Attentional Blink (AB) (Chun & Potter, 1995; Raymond et al., 1992)) and by neurophysiology. In order to frame our theory, we identify a set of functional requirements.

1.2 Functional Requirements for Model

When we consider behavioral data, it will become clear that, in certain very demanding visual environments, some of the following functional requirements will not obtain.[i] Thus, our claim is that the following capacities characterize the "normal functioning" of temporal attention.

1. *Salience Detection.* The system should have the capacity to rapidly detect that a stimulus is salient when a fleeting representation of it arises amongst a set of (temporally) competing stimuli.
2. *Sustain Representations.* The system should have the capacity to sustain fleeting, but salient, representations of visual stimuli, in order that they can be acted upon (and encoded into WM).
3. *Ascribe Episodic Context.* The system should have the capacity to ascribe (discrete) episodic contexts to items as they are encoded into Working Memory (WM), thereby enabling order information to be retrieved.
4. *WM Maintenance.* The system should have the capacity to maintain multiple items in WM.

5. *Repetitions.* The system should have the capacity to represent multiple instances of the same item. Although this may seem like a trivial constraint, in fact, it is not always that easily obtained, see section 6.1. The reason for this difficulty is that perceiving and encoding repetitions requires the neural resources used for detecting the first instance of an item, to be "freed-up" before the second instance arises.

The paper is organized as follows. In section 2, we present an informal description of the Simultaneous Type Serial Token (ST^2) model, motivating the approach from the perspective of our functional requirements. This is followed, in section 3, by a discussion of the key empirical findings we will model; these are based upon the Attentional Blink (AB) phenomenon. Then, in section 4, we present a formal, neurally explicit, description of the model, called Neural-ST^2 and show how it reproduces the empirical findings. Following this, section 5 identifies new predictions arising from the model and verifies them. Then sections 6 and 7 relate ST^2 to competitor models of the blink. Finally, section 8 discusses how the approach relates to theories of temporal attention in general and presents concluding remarks.

The Simultaneous Type Serial Token Theory

2 Informal Description of Model

The following principles underlie ST^2. We consider these in turn in sections 2.1-2.5.

1. *Two Stages.* We postulate a cascaded parallel first stage of visual processing, followed by a serial second stage tied to WM encoding.
2. *Salience Filter.* We assume a mechanism for attributing salience, which filters the set of competing items.

3. *Types-tokens.* We distinguish between the featural representation of an item (its type) and the episodic context in which it occurs (a token). In ST^2, associating a token with a type is the process of WM encoding.
4. *Transient Attentional Enhancement (TAE).* We postulate a mechanism that, in response to the detection of a salient stimulus, provides a temporally brief (but spatially specific) enhancement of type representations.
5. *Binding Pool.* We introduce a neurally plausible means to associate types with tokens.

2.1 Two Stage Theory

Two stage theories arise throughout the recent study of attention (Broadbent, 1958; Chun & Potter, 1995; Shih & Sperling, 2002; Treisman & Gelade, 1980). These theories are distinguished by the processes they ascribe to each stage and the properties of the interface between the two stages (Driver, 2001). The particular two-stage theory that ST^2 employs is best explained by considering the stages in turn.

Stage One

In the ST^2 model, stage 1 implements standard visual processing: from visual feature extraction to semantic categorization. Importantly, (modulo constraints arising from stimulus presentation[ii]) stage one is *parallel*. Thus, multiple items can be concurrently processed, with little interference between them. This is illustrated in Figure 1, where progression through the stage is depicted as a vertical pass through a series of processing levels (horizontal bands).

Insert Figure 1 Here

In addition, processing is *cascaded* across the levels of stage 1. This is illustrated in Figure 1 by depicting item representations as rectangles. Thus, item representations are "smeared" across a number of levels at any instant, with decreasing strength to the extremes of the rectangle. This is consistent with neural activation dynamics, which build up gradually and

persist (in a decaying form) beyond the activation peak. A consequence of this cascaded processing is that multiple items can be coactive at the same level, even if they are presented to the system one after the other.

Effectively, each level of stage 1 acts as a very short-term memory, the duration of which is governed by the speed of activation dynamics. Indeed, early levels of stage 1 could be related to Iconic Memory (Coltheart, 1983; Shih & Sperling, 2002). Thus, representations persist for hundreds of milleseconds unless overwritten, as can arise from masking (Keysers & Perrett, 2002). However, these memories are indeed very short term. Thus, in the absence of top-down enhancement, stage 1 representations are subject to rapid decay, as well as being vulnerable to overwriting.

Stage Two

For an item to obtain a more durable representation, it has to make it through to stage 2. That is, stage 2 is the "entrance" to WM, with a probable role for conscious perception in this process. In addition, in contrast to stage 1, the second stage imposes sequentiality constraints. These arise since the system attempts to associate items with discrete episodic contexts. We consider the mechanism by which such associations are built when we discuss the types - tokens distinction.

2.2 Salience Filter

We use the term salience filter to include all mechanisms that provide top-down selective enhancement, whether driven by long-term cognitive goals or emotional significance. Such mechanisms can operate at many points in the stage 1 processing pathway. In particular, task set modulates the response of neurons throughout the visual processing hierarchy (Desimone & Duncan, 1995).

We depict this mechanism (see Figure 1) as a processing layer that selectively emphasizes certain channels and de-emphasizes others. This approach is consistent with the task demand systems used in connectionist models (Cohen & Huston, 1994; McClelland, 1993). However, Figure 1 is not intended to indicate that the focus of the salience filter is permanently fixed. Rather, the filter is being constantly adjusted, both horizontally (in terms of stage 1 channels emphasized and de-emphasized) and vertically (in terms of level of operation). Indeed, it is likely to operate in graded fashion at many levels at the same time.

It is through the salience filter that cognitive control (Botvinick, Braver, Barch, Carter, & Cohen, 2001) enforces the task set (Monsell, 2003). Thus, the salience filter enhances task relevant items, enabling them to progress into stage 2. In contrast, the filter typically ensures that task irrelevant items do not reach stage 2.

2.3 Types-tokens

The types-tokens distinction has been useful in explaining a number of temporal attention phenomena (Chun, 1997; Kanwisher, 1987; Kanwisher, 1991) and it plays a central role in the ST^2 model. Before we come to this though, we define terms. *Types* encompass all featural properties of an item. For example, the type representation of the letter K would contain 1) semantic features, e.g. that it is in the category of letters and it follows J in the alphabet, and 2) visual features, e.g. its shape, the angled line segments that comprise it and the color in which it appears.

In contrast, a *token* is a compact encoding of instance specific (or episodic) information. Thus, a token indicates that a particular type has occurred and also, *where* relative in time to other items it occurred. Note that, because of our focus, "where" is interpreted in temporal (rather than spatial) terms. Our hypothesis is that, while type representations are activated when detecting and identifying an item, it is unlikely that it is the type that is held active when the item is maintained in WM (see section 6.1 for a

discussion of how this impinges upon active memory models). Rather, we argue that there is a transition from full type activation to compact token encoding, as an item passes to a more durable representation. Thus, in the ST^2 model, WM encoding is exactly the process of associating (or binding) a token to a currently active type; we also use the term *tokenization* to describe this process. In this sense, once bound, tokens act as "pointers" from which the corresponding type can be regenerated when required.

Furthermore, once a token is bound to a type, in the absence of continued bottom-up support, activation of that type can "safely" be dissipated. In terms of Figure 1, tokenization frees up the stage 1 channel of the type that has been tokenized. The fact that a second instance of a previously encoded item can be perceived at all, suggests such a type dissipation (although, repetition blindness (Kanwisher, 1987) suggests that this liberation of resources can incur costs). That is, if WM maintenance were to commit the neural resources of a type, then those resources would not be available to detect a repetition of that type. Indeed, such an approach would also have difficulty retaining distinct items with common features. The alternative to this is that the brain holds duplicates of each type, with the number of duplicates required being bounded, at the least, by the capacity of WM.[iii] However, this would 1) require a complex type handover mechanism and 2) be exceptionally expensive in respect of neural resources.

Figure 1 depicts a snapshot of the tokenization process. Tokens are made available in sequence, and the model is shown in a state in which the first token is available, and all other tokens are unavailable. Item i1 has reached the end of stage 1 and thus, is fully (type) processed. In particular, i1 has largely passed the salience filter, and thus, has been identified as suitable for WM encoding.

Tokenization is initiated when an item passes the salience filter. Thus, Figure 1 shows a state in which i1 is being bound to the available token. Once this tokenization is complete,

the system enters a state such as that shown in Figure 2. Once token 1 has been bound, a handover takes place and tk2 becomes available, i.e. tokens are serial.

However, a neurobiological realization of tokenization would necessarily be complex; as a result, tokenization is not an instantaneous process. This has important implications, since the binding of tokens to types is promiscuous in nature. While stage 1 can represent multiple items concurrently, the tokenization process binds the currently available token to *all* types active post the salience filter. For this reason, it is important that during token binding, subsequent items be prevented from becoming strongly active. Thus, to protect the integrity of ongoing tokenization, attentional resources are withheld to prevent interference from other stage 1 items. In ST^2 terms, the attentional resource is the *Transient Attentional Enhancement (TAE)*, see section 2.4.

Insert Figure 2 Here

In RSVP, the task is to detect and / or identify the salient items in the sequence. In order to do this, participants must segment the presentation stream into discrete components, with key components being the salient items. The token system allocates discrete episodic contexts consistent with such segmentation. If a sequence is presented at "normal" speeds, i.e. a good deal slower than in RSVP, each item encoded into WM is bound to a separate token. As a result, serial order can be retrieved, i.e. the item bound to token 1 is first, the item bound to token 2 is second, etc.

In addition, the ST^2 approach can encode a rich variety of type-token associations. Firstly, the system enables multiple instances of a type to be recorded. That is, it can reach a state in which multiple distinct tokens are associated with the same type. It is also possible for two tokens associated with the same type to be adjacent in time. This will arise if an item is repeated with a sufficient gap in time between the two instances to ensure that the first tokenization has completed before the second starts.

Secondly, in extreme situations, the system can reach states in which two types are bound to the same token. This represents a state in which the system has failed to encode a unique episodic context to the occurrence of each of the types and will play an important role in our modeling of the attentional blink; see the Lag 1 Sparing subsection of section 3.3. Finally, inherent in the types-tokens approach is a justification for why the two-stage theory has the characteristics it does. That is,

1. stage 1 is parallel, since multiple distinct types can be simultaneously represented;
2. stage 1 is subject to rapid forgetting, since active types decay;
3. stage 2 ensures a more durable representation, since token binding compactly encodes a recognition event, which can be maintained without committing type representation resources;
4. stage 2 is serial, since it seeks to provide an unambiguous binding of types to discrete episodic contexts in the setting of a promiscuous tokenization process.

2.4 Transient Attentional Enhancement

An item passing the salience filter in a strongly active form, as item 1 has done in Figure 1, initiates an attentional enhancement, which elevates activation across the later levels of stage 1. In this sense, the mechanism is exogenous in character, being instigated by the occurrence of an environmental stimulus. This mechanism highlights a very brief window of time and space that is particularly salient. The mechanism has a number of characteristics.

1. The TAE enhances the activation level of salient items, aiding their encoding into WM. This is particularly important when considering demanding stimulus environments (such as RSVP), since, unaided, fleeting representations have insufficient bottom-up activation to facilitate encoding. That is, attentional enhancement plays a key role in facilitating type - token binding.

2. The TAE has two important characteristics, a) the attentional enhancement is brief (a pulse of around 100ms), and b) it only fires once per tokenization. Thus, detection of a salient item fires the TAE, but then it is maintained offline until this tokenization is complete. Unavailability of the TAE protects ongoing tokenization from the intrusion of newly arriving items, in an attempt to ensure discrete allocation of episodic contexts (i.e. tokens).
3. We hypothesize that the TAE is spatially specific, but is not featurally selective in its enhancement. That is, it is initiated by detection of features characterizing the task set but the effect of the TAE is not restricted to those features.

A Transient Attentional Enhancement is suggested by RSVP data, particularly, one of the trademarks of the Attentional Blink (AB), lag 1 sparing (see the Lag 1 Sparing subsection of section 3.3). However, the TAE is also supported by other psychophysics research. For example, Nakayama and Mackeben (1989) described two forms of covert visuospatial attention. One was a sustained component that was slow to deploy. The other was a transient component with behavioral effects beginning 50ms after a task relevant cue, but then fading within 150 ms. In their paradigm, these two forms of attention sum.

In addition, Muller and Rabbitt (1989) found elevated performance when cues preceded a fleeting target by 100 or 175 ms. They made the distinction between peripheral cues that appeared in the same location as the target they cue, and central cues (e.g. arrows) which appeared at fixation and directed attention to the target. They found the same pattern as Nakayama and Mackeben (1989); that is, peripheral cues first evoked a transient pattern of improved accuracy over SOA's 100 and 175, which then fell to a baseline defined by the sustained component of attention. Recent work exploring transient attention has identified similar effects (Kristjansson, Mackeben, & Nakayama, 2001; Kristjansson & Nakayama, 2003).

Deploying attention to any stimulus that is fleeting requires a mechanism with a rapid onset, as we describe here. This situation would, for example, arise when attempting to identify objects that are intermittently occluded. Furthermore, its rapid onset is well suited to deploying a brief episode of attention during a single saccade.

2.5 The Binding Pool

Up to this point, we have drawn arrows linking type channels and tokens as required, however this simplicity of depiction belies the technical difficulties involved. Indeed, from a neural perspective, token binding is the most challenging aspect of our theory. This section introduces a suitable mechanism: the binding pool. The following are requirements for this mechanism.

1. It needs to be able to represent arbitrary type - token bindings. In particular, it should be possible to bind multiple tokens to the same type; see section 2.3.
2. Token bindings need to be set-up in real-time and on demand, i.e. when a salient item is detected.
3. We require that the mechanism encodes bindings efficiently and in a neurally feasible manner (see section 4.6). It is easy to describe strategies that are inefficient. For example, the number of neurons needed to fully characterize a type is likely to be very large. Therefore, an approach requiring the full neural profile of each type to be duplicated (in order, say, to encode repetitions) would be extremely wasteful and is unlikely to have arisen from evolution.
4. Following on from the previous point, it is also necessary to identify how the approach would scale to a brain-level realization

Hebbian learning is selective for synapses that are conjunctively (pre and post synaptically) active (Hebb, 1949; O'Reilly & Munakata, 2000) and thus, seems to be ideal for binding types to tokens. However, the time scale required for these changes, which occurs in vitro

over the course of minutes, suggests that it is unsuitable. This is especially true in the temporal attention context where multiple WM encodings can occur in a few hundred milliseconds.

Another candidate mechanism is short-term synaptic modification, whereby the responsiveness of a synapse is a complex function of its recent activity (Liaw & Berger, 1996; Pantic, Torres, Kappen, & Gielen, 2002). This mechanism is certainly rapid enough, however it occurs without post-synaptic specificity. Thus, co-activation of a type and a token would cause non-selective potentiation from a token to all types or vice versa. Consequently, information about which type co-occurred with which token would not exist.

We are left with an account based upon sustained activity, which is probably the standard explanation for WM (Cowan, 2001). In particular, a number of buffer models exist that posit sustained activity at various levels of representation (Amit, Bernacchia, & Yakovlev, 2003). Typically though, these models do not fulfill the requirements identified in section 1.2 (see section 6.1 for a detailed justification). For example, objects would be stored by sustained activity in the same (type) nodes used to recognize those items, making it impossible to detect repetitions, amongst other limitations. Note, this same problem would arise with an approach based upon temporal synchrony (Hummel & Biederman, 1992; von der Malsburg, 1981). Further, pure active memory models typically leave unexplained the means by which stored memories are accessed. This is in contrast to types-tokens accounts, where active tokens are a localized focus through which encoded types can be regenerated.

Insert Figure 3 Here

In response, we propose a significant modification of the active memory approach. In ST^2, a dedicated pool of binding nodes (see Figure 3) forms a self-sustaining pattern of activity during encoding to store the coincidence of types and tokens. Binding pool units are selective for specific types and tokens, with each type and token being coded for by overlapping

subsets of the pool. Figure 4 shows a binding pool state arising when binding type i1 to token tk1. As is typical of the approach, some binding pool units code neither i1 nor tk1, while some code just tk1, some just i1 and others both (i.e. the conjunction of) tk1 and i1. It is these conjunctive units that bind i1 to tk1 and which become (durable) active representations. Thus, the sort of binding link we have been depicting as an arrow between a token and a type, is effectively piggy backed on top of the active binding pool representation.

Insert Figure 4 Here

The binding pool satisfies the requirement that WM does not commit type representation space, thus allowing a type to be processed even if it is already present in the memory set. In section 4.6 we discuss an implementation, which clarifies how binding pool units are realized as neural circuits and the connectivity into and out of the pool. In this implementation, the binding pool will contain localist representations of every possible combination of target type and token, similar to Deco, Rolls and Horowitz's (2004) approach to binding items to spatial locations.

In scaling such localist approaches to brain-level implementations, there is an inherent problem of combinatorial explosion in the size of the pool of neurons required to represent every possible combination (in our case, (number of tokens) × (number of types) nodes). The use of a distributed representation within the binding pool solves this problem. A small pool of nodes can store a compact representation of all possible combinations of types and token (Wyble & Bowman, 2006). However, with a distributed approach, there is overlap between the representations of different token - type bindings, with the weights into and out of the binding pool governing the level of overlap (indeed, it is this overlap that leads to a more efficient encoding). As a result, the distributed approach produces a falloff in retrieval fidelity as memory load increases, replicating Wilken and Ma's (2004) experimental results. Thus,

although for simplicity of modeling and presentation, we use a localist binding pool in this paper, we believe that distributed binding pools would scale.

2.6 Discussion

The ST^2 model, satisfies the five functional requirements of temporal attention highlighted in section 1.2.

1. *Salience Detection.* ST^2's salience filter detects salient stimuli when briefly presented in visual environments containing (temporally) competing stimuli.
2. *Sustain Representations.* The TAE has the capacity to sustain fleeting, but salient, representations of visual stimuli, enabling WM encoding.
3. *Ascribe Episodic Context.* ST^2's types-tokens system enables discrete episodic contexts to be ascribed to items. Indeed, the allocation of such a token *is* the process of encoding into WM. In addition, tokens are allocated in sequence, enabling the order of perceived occurrence of items to be retrieved.
4. *WM Maintenance.* Since multiple tokens are available, many items can be maintained together in WM.
5. *Repetitions.* Multiple tokens can be bound to the same type, enabling repetitions to be maintained in WM. Furthermore, since type resources are freed-up once an item has been tokenized, those resources are available to perceive and encode further instances of the same item.

Finally, in respect of the standard debate between early and late selection in attentional processing (Driver, 2001; Logan, 2004), we are largely noncommittal, since configurations of the salience filter at early levels of stage 1 would generate early selection, while configuration at late levels would generate late selection. This said, the salience filter configurations considered in this article would enforce late selection, as suggested by the characteristics of the Attentional Blink (see section 3.2). It is important though to realize that,

in contrast with many other prominent two-stage theories (e.g. Broadbent, 1958), it is not the filter that is the interface between stage 1 and stage 2 or (in an absolute sense) imposes seriality. In ST^2, the salience filter certainly is a bottleneck, since it restricts what passes through it, but it does not, per se, enforce seriality.

The Empirical Landscape

3 Temporal Attention Studies

We will assess the ST^2 model against a set of archetypal temporal attention data, focused on the Attentional Blink (AB).

3.1 The Attentional Blink

Perhaps the most prominent RSVP paradigm is the Attentional Blink (AB) (Raymond et al., 1992). In this task, two targets are placed in a sequence of items presented at a rate of around 10 items per second. We focus on the *letters-in-digits* task (Chun & Potter, 1995) in which participants must report the identity of two letter targets (T1 and T2) presented in a stream of digit distracters. Our reason for selecting this task is that certain aspects of other tasks would not be straightforward to model. In particular, the letters-in-digits task does not require binding across features, e.g. between color and identity, as found in color marked tasks (such as, Maki, Couture, Frigen, & Lien, 1997). Furthermore, the task can be argued to yield a *pure* test of the blink, since there is no task switch between the T1 and T2 tasks (Chun & Potter, 2000).

3.2 Findings to be Reproduced

The specific findings that we will reproduce are as follows,

1. *The Basic Blink*. A typical AB serial-position curve, which arises from the letters-in-digits task, is shown in Figure 5. Points to note are,
    a. the blink is a 200ms – 500ms (approx) interval post T1 onset in which performance on T2, conditional on correct report of T1 (i.e. T2 | T1), is significantly below baseline;
    b. generally, the blink has a sharper onset than offset; and
    c. if T2 immediately follows T1 it is reported at baseline levels, which is described as *lag 1 sparing*.

    Of particular note is lag 1 sparing, which seems counter-intuitive, since one may believe that limited attentional resources are maximally engaged in processing T1 when T2 "arrives" at lag 1.

Insert Figure 5 Here

2. *T1+1 and T1+2 blank*. The blink was attenuated when a blank was placed in the T1+1 position, but not when it was placed at T1+2 (Chun & Potter, 1995; Raymond et al., 1992), see Figure 5.
3. *T2 unmasked.* The blink was significantly attenuated if T2 was the last item in the stream (Giesbrecht & Di Lollo, 1998), see Figure 5. In addition, we suggest that, placing blanks after T2 will attenuate the blink.
4. *T1 Performance.* T1 performance was reduced at lag 1; see Figure 6.

Insert Figure 6 Here

5. *Swaps.* It is possible that both T1 and T2 are identified, but they are perceived in the wrong order. In Chun and Potter (1995), such T1 - T2 swaps were maximal at lag 1 and effectively disappeared by lag 3; see Figure 6. In fact, at lag1, temporal order judgment was only a little above chance.

6. *Stage of the AB bottleneck.* A number of studies indicate that the blink results from a late stage bottleneck.
    - *Priming*. Shapiro, Driver, Ward and Sorensen (1997) considered AB paradigms with three targets. In one task, T2 was an upper case letter and T3 was lower case. The basic finding was that when a T2 was missed, correct report of T3 was higher if the two targets had the same identity. In the second experiment, the stream items were words and it was discovered that missed T2s semantically primed T3.
    - *T2 breakthrough Effects*. High salience T2s "breakthrough" the blink. For example, substantial blink attenuation was shown when personal names were used as the T2 (Shapiro, Caldwell, & Sorensen, 1997). In addition, emotionally salient T2s attenuated the blink (Anderson & Phelps, 2001).
    - *ERPs*. Early visual components of the ERP waveform (the N1 and P1) elicited by T2s have been shown to be comparable whether T2 was missed or seen. In addition, semantic effects (N400) of a T2 word were also present when that item was missed. In contrast, a component associated with WM update (the P3) was reduced when a T2 was presented during the blink compared with outside that period (Luck, Vogel, & Shapiro, 1996; Vogel, Luck, & Shapiro, 1998).

    These investigations suggest that even when a T2 is missed during the blink, it is nonetheless extensively processed, in respect of both visual and semantic features, allowing priming and breakthrough. Furthermore, the ERP P3 data suggest that the blink bottleneck is located at the point of encoding into WM.
7. *Priming from T1+1 distracters*. Chua, Goh and Hon (2001) found that performance on a color marked T2 letter was improved if the same letter was presented as a distracter in the T1+1 position, suggesting that distracters following the T1 prime later T2s.

3.3 The ST² Model and the Attentional Blink

Before we move to introducing our neural implementation, we consider how the AB fits with our informal formulation of the ST^2 model.

Fleeting Representations and Bottom-up Trace Strength

In the context of RSVP, we will argue that the fleetingness of visual stimulus representations arises from masking. That is, in a basic blink stream, each item (apart from the first) acts as a mask for the item that precedes it[iv]. As a result, masking attenuates sensory traces of items. Accordingly, we will distinguish between *weak* and *strong sensory traces*, which arise when an item is masked and unmasked, respectively. The key situation in which items are unmasked is when they are followed by a blank.

In a manner that we will clarify shortly, strong traces can more easily be encoded into WM than weak traces. It is for this reason that the blink is attenuated in the T1+1 blank and T2 as end of stream conditions; see Figure 5. This interpretation is consistent with a number of studies that have explored the effects of low-level masking in the AB (e.g. Seiffert & Di Lollo, 1997).

Why the ST^2 Model Blinks

In terms of the ST^2 model, the blink occurs because, while stage 2 is occupied encoding T1, T2 is decaying at stage 1 and may have decayed completely by the time stage 2 is again free. Specifically, the blink arises because the tokenization process is overloaded, in the sense that targets arrive so rapidly that T1 is still being tokenized when T2 arrives. As a result, the system sacrifices tokenization of T2 (by taking the TAE offline) in order to prevent it interfering with the tokenization of T1. That is, the blink is a mechanism to avoid spurious and confused binding of T2 to a token that is in the process of being allocated to T1.

Lag 1 Sparing

It may seem that lag 1 sparing is a problem for this explanation of the blink, since at lag 1, T2 | T1 performance is high even though the T1 and T2 tokenization processes maximally overlap in time. However, lag 1 sparing does not come free of cost. Firstly, as discussed above, T1 accuracy is degraded at lag 1, suggesting that the competing T1 and T2 tokenizations can resolve in favor of T2 and at the expense of T1. Secondly, as again previously highlighted, swaps are very high at lag 1.

In ST^2 terms, in the lag 1 case, when the TAE fires in response to T1 detection, T2 is sufficiently advanced in stage 2 that it benefits from that attentional enhancement (indeed, as suggested by Potter, Staub and O'Conner (2002) it is likely that the T2 benefits more than the T1). As a result, both T1 and T2 become strongly active together in the post salience filter levels of stage 1, which yields the degenerate situation in which two types are bound to the same token. Loss of temporal order information is characteristic of such a failure of the tokenization process.

More generally then we would argue that due to the rapid arrival of fleeting representations, the system is unable to extract all the information available in the stimulus stream. Thus, it is forced to trade extraction of one information-bearing dimension off against extraction of another. Indeed, the transition from lag 1 to lag 2 represents an adjustment of this trade-off, from extraction of information about both targets at the expense of confused binding (at lag 1), to extraction of unequivocal T1 information at the expense of T2 information (at lags 2, 3 and 4).

Neural Realization

4 The Neural ST^2 Theory

This section describes a neural realization of the ST^2 model (called Neural-ST^2). We begin, in section 4.1, by introducing a set of neural principles. Then, in sections 4.2 - 4.7 we describe the neural realization. In section 4.8 we show the data the model reproduces and, finally, in section 4.9 we relate the model to neurophysiology.

4.1 Neural Representation Scheme

We first describe activation equations and then highlight key neural mechanisms. However, our description of the model can be comprehended without understanding these mathematical details.

Simple Elements

Activation functions of neural units are simple combinations of bias, excitation, inhibition and leak currents. Connections between elements are excitatory or inhibitory and are not modifiable. The membrane potential update function is:

Equation 1: Membrane Potential

$$MP_{(i,j,t)} = MP_{(i,j,t-1)} + DT\_VM_{(j)} \times [(Bias_{(j)} + Excite_{(i,j,t-1)}) \times (EE_{(j)} - MP_{(i,j,t-1)}) + Inhib_{(i,j,t-1)} \times (EI_{(j)} - MP_{(i,j,t-1)}) + Leak_{(j)} \times (EL_{(j)} - MP_{(i,j,t-1)})]$$

where $MP_{(i,j,t)}$ is the membrane potential for node *i* of layer *j* at time *t*.

The membrane potential is a measure of the extent to which a node is excited. Thus, when *MP* is high, the node is very excited and when it is low it is unexcited. *DT_VM* is a constant determining the rate of change of a node. *Bias* is a constant input current, which regulates the basic excitability of a node, thus, the higher the bias, the more easy it is to push the

membrane potential over its output threshold (see Equation 2). *Excite* and *Inhib* are weighted summations of all excitatory and inhibitory inputs respectively. *Leak* is a leak current, which determines how quickly the node decays back to its resting level.

EE, *EI* and *EL* are reversal potentials for excitatory, inhibitory and leak currents. Excitation and bias push the node up to the value of *EE*. In contrast, inhibition and leak push the node down to *EI* and *EL* respectively. Thus, *EE* bounds the membrane potential from above and *EI* (which is set to be smaller than *EL*) bounds the potential from below. In addition, *EL* sets the resting level, since the node will stabilize at this value if it has no excitatory or inhibitory input for a period of time. The output mapping is a sigmoid function, which bounds a node's output to the range [0,1).

Equation 2: Output Function

$$Out_{(i,j,t)} = \frac{[MP_{(i,j,t)} - \theta_{(j)}]_+ \times \gamma_{(j)}}{[MP_{(i,j,t)} - \theta_{(j)}]_+ \times \gamma_{(j)} + 1} \quad where \quad [x]_+ = \begin{cases} x & if \quad x > 0 \\ 0 & otherwise \end{cases}$$

Thus, if the membrane potential is below the threshold θ then the node outputs zero. However, the output rises rapidly as the membrane potential becomes bigger than θ and then saturates at large values. How rapidly output approaches the saturation limit of 1 is governed by the scaling parameter, γ. These equations are similar to those found in O'Reilly and Munakata's (2000) biologically plausible neural networks and can be related to Hodgkin-Huxley equations (Koch, 1999).

Key Neural Mechanisms

We now consider three mechanisms of which we will make liberal use.

- *Self-excitation*. In the absence of excitatory input, leakage ensures that a node's activation decays exponentially back to resting levels. However, we will need to model a more temporally sustained maintenance of activation, e.g. activation based memory (Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Davelaar & Usher, 2004). In

localist models, this effect is obtained by including an excitatory link from the node back to itself, see Figure 7(a). Such a self-excitatory node can be viewed as an abstract representation of an assembly of biological neurons, which sustains activation across the assembly through recurrent interactions between constituent neurons.

Insert Figure 7 Here

- *Off Nodes.* It can also become necessary to break a self-sustaining cycle. A simple way to do this is to use an inhibitory inter-neuron (i.e. an off node) (see Figure 7(b)), which builds-up in response to activation of its associated self-sustaining node and, once the inhibitory inter-neuron crosses its output threshold, it suppresses the self-sustaining node. In fact, neurobiologically speaking, off nodes could be realized in a number of ways. For example, off nodes could be instantiated as an intracellular process, e.g. a calcium activated after-hyerpolarization (AHP) potassium conductance current is found in most neurons. This mechanism reduces excitability following periods of strong activation (Bal & McCormick, 1993; Hotson & Prince, 1980). Alternatively, off nodes could be viewed as a localist implementation of a layer of inhibitory inter-neurons (O'Reilly & Munakata, 2000).

Insert Figure 8 Here

- *Gate-Trace Circuit.* This neural structure revises the off node circuit shown in Figure 7(b) by moving the self-loop to the off node, as shown in Figure 7(c). We call this a gate-trace pair. As a result, the (off) trace node can only be activated through the (on) gate node and, once activated, the trace node stays active. (The fact that the trace node stays active is a major difference to the previously discussed (off node) circuit, where, once the off node fires no activity is maintained in the circuit.) Typical behavior of such a circuit is depicted in Figure 8. Panel A shows the circuit in its initial state. Panel B shows the state reached when the circuit has been externally excited through

the gate node. As the gate node becomes activated it excites its trace node, which, when strongly active, suppresses the gate. Eventually the circuit settles into a new stable state with a strongly active trace node and a suppressed gate node (as shown in panel C). Thus, the trace node is an active memory, which is accessed through the corresponding gate node. In addition, when in its sustaining state (see panel C), the gate node can be reactivated by suppressing the trace node, thus, freeing the gate to be biased by extrinsic influences. This method can, for example, be used to retrieve from the active memory. A layer of such gate-trace pairs amounts to a *gated active memory*. A gate-trace pair provides sustained activation without "tying" up the whole layer. For example, competition (via lateral inhibition) amongst gate nodes can continue in the presence of sustained trace activation. A simple layer of competing self-sustaining nodes does not behave in this way: once a node in the layer is in a self-sustaining state there is no future opportunity for competition.

4.2 Overview of Neural-ST^2

The Neural Simultaneous Type Serial Token model is depicted in Figure 9. ST^2's first stage is realized by the bottom four layers (Input, Masking, Item and Task Filtered), which model the extraction of types. The connectivity (which is one-to-one between layers, and has no lateral inhibition at the TFL and Item layer) ensures that stage 1 can represent multiple items simultaneously. The second stage is realized by the WM layer (comprising tokens) and their interaction (through the Binding Pool) with the Task Filtered layer.

Insert Figure 9 Here

Furthermore, the second stage is serial in the sense that only one token is available at any time. Finally, ST^2's TAE, which we realize with a mechanism called *the blaster*, is implemented as a neural circuit, with the projections into the circuit controlling the TAE's availability. We now move to a more detailed description of these mechanisms. First, section

4.3 reviews the first stage. Then the second stage is described in section 4.4. Following this (in section 4.5) we consider Neural-ST^2's TAE and then in section 4.6 we discuss the binding pool. Finally, section 4.7 presents a brief summary of the neural realization.

4.3 Neural-ST^2's First Stage

Input

The RSVP stream is modeled by presenting a sequence of patterns to the input layer. Each pattern models an item in the sequence and is presented for 20 time steps. Trials are 550 time-steps, each step simulating 5 ms of time. Streams contain 2 targets and 16 distracters, with T1 always the 7th item and the two targets separated by 0-7 distracters. T1 and T2 strengths are varied systematically over the range 0.43 to 0.61 (in steps of 0.012) for each lag. This variance simulates featural aspects of letters and digits, which make some of them easier to perceive. For the sake of simplicity, the model does not explicitly represent the binding of visual features into complete items. Rather, at all layers, a single node represents an item presented to the model. The model is thus localist in nature. In addition, there is no explicit representation of spatial location in the model, although this could easily be added. We have abstracted from these details, since our main focus is higher up the visual processing pathway, i.e. the point of WM encoding.

Insert Figure 10 Here

Early Visual Processing: Masking Layer

Input activation propagates over unidirectional connections through layers that abstractly represent steps of visual processing. In RSVP streams, since stimuli are presented at the same spatial location, successive items are represented using the same neural tissue. As a result, items overwrite one another, generating masking. Such overwriting is abstractly modeled at

the masking layer, which includes inhibitory interactions between items; see Figure 9. The masking effects generated by such inhibition are a key determiner of bottom-up trace strength; see the Fleeting Representations subsection of section 3.3. Thus, activation traces are greatly weaker for masked as opposed to unmasked stimuli; see Figure 10. Our approach is consistent with single cell recordings (e.g. Keysers & Perrett, 2002; Kovacs, Vogels, & Orban, 1995; Rolls, Tovee, & Panzeri, 1999), that show strong amplitude differences in the firing rate of temporal cortex neurons to different types of masking in monkeys.

Masking is asymmetric; that is, backward masking is stronger than forward masking (Seiffert & Di Lollo, 1997). This asymmetry arises since feed-forward inhibition into the masking layer is stronger than lateral inhibition within the layer. Lateral inhibition creates symmetric suppression between masking layer nodes, i.e. the X^{th} and $X\text{-}1^{st}$ list items suppress each other to a similar degree. In contrast, since feed-forward inhibition only generates suppression from an active input node (as set by the current input pattern), it only enables the X^{th} item to inhibit the $X\text{-}1^{st}$ item.

Type Processing: the Item and Task Filtered Layers

The final two layers in stage 1 (see Figure 11) abstractly represent type oriented processing, including extraction of conceptual features (Potter, 1993). The first of these, the Item layer, supports a temporally sustained (equivalent to several hundred milliseconds), but decaying, representation of items (see Figure 10). Task set has no affect upon this layer (i.e. distracters and targets are treated identically) and similar items have weak excitatory connections (see Figure 11), along which an active item can prime similar items. Thus, even weak T2s, which are more likely to be missed, can prime future targets, as required by the late blink bottleneck hypothesis. Despite the fact that the Item layer is localist, the brain may realize the same effect through an efficient distributed representation scheme, with priming arising as a by-product of the resulting overlapping item representations.

Insert Figure 11 Here

In contrast, at the Task Filtered Layer (TFL) task demand selectively excites target nodes while suppressing distracter nodes (see Figure 11). Thus, in respect of the ST^2 theory, the TFL enforces the salience filter, with, in this context, task instructions fully prescribing salience. The Task Filtered Layer is the gateway to tokens. Excitatory connections from TFL items project to the binding pool, which, in turn, projects excitatory links to the gate nodes of tokens (see Figure 12). TFL and Item nodes are weakly self-excitatory, yielding temporally sustained traces. However, these layers are also connected in a 1 to 1 fashion with off layers, which curtail activation traces; see the Token System subsection of section 4.4.

4.4 Neural-ST^2's Second Stage

In accordance with the ST^2 theory, the Neural-ST^2 first stage yields a decaying trace of the visual and semantic features of items (see Figure 10), which is parallel, since different items can be simultaneously active at the Item and Task Filtered layers. In contrast, stage two realizes encoding into WM, with sequentiality emerging from mechanisms that attempt to ensure that items are discretely bound into WM. We begin by discussing the token system. Then we discuss the retrieval of items from tokens.

The Token System

Tokenization. During presentation of an RSVP stream, task-relevant stage 1 items can activate a token, this activation being relayed via the binding pool. This initiates a tokenization process, whereby binding pool units are allocated. The process of activating the token and allocating binding pool nodes is not instantaneous; it takes hundreds of milliseconds. However, the stronger the type trace, the faster it occurs.

Insert Figure 12 Here

The token mechanism is implemented as a layer of trace-gate pairs, each of which corresponds to a token; see Figure 12. That is, the rectangular tokens depicted in our informal exposition of the ST^2 theory, e.g. as shown at the top of Figure 2, are implemented as trace-gate circuits. Tokens have four states: *available*, *unavailable*, *binding* and *bound*, see Figure 13. An available token has a minimally active gate node and an inactive trace node. While one token is available, others are unavailable due to lateral inhibition between the gates. A token in the process of binding has a highly active gate node, being driven (via the binding pool) from the first stage, and a trace node, which is steadily accruing activation. A bound token has a suppressed gate node receiving continuous inhibition from its active trace node.

Insert Figure 13 Here

While a token is being bound, binding pool units sitting between the token's gate node and active TFL nodes are incrementally allocated. The rate at which these nodes are allocated is proportional to the strength of the individual traces in the TFL. The self-loop associated with each WM trace node ensures that, once the corresponding WM gate is bound (as shown in panel C of Figure 13 for Token 1), the trace node will remain active indefinitely. Such sustained trace node activation is used later to determine which tokens were bound. Allocated binding pool nodes between the WM gate nodes and the TFL determine what item or items were encoded by those tokens.

Properties of the Token System. WM gate nodes compete to become available at the beginning of a trial and also when a token has completed being bound. The winner is determined by an ordered pattern of biases applied to the WM gates, which determines which token is first, second and so on, see Figure 13. In this way, Neural-ST^2 ascribes discrete episodic contexts and encodes order when items are presented slowly enough.

It is possible that binding pool nodes allocated from different WM gates may project to the same TFL node. This reflects the fifth of the functional requirements we highlighted in

section 1.2, since it allows representation of multiple instances of the same item, i.e. repetitions. It is also possible for binding pool units to be allocated between one token and many TFL nodes. That is, many items are encoded into WM, but order information is lost, which will yield order inversions at recall, i.e. swaps.

Task Filtered Layer Off Nodes. While it is a common phenomenon to fail to perceive a rapidly presented object, participants do not report extraneous instances of a single item presented within an RSVP stream. To ensure that Neural-ST^2 also possesses this property, TFL nodes have their own off nodes, see Figure 11. These off nodes ensure that, after a sustained period of activation, an item is inhibited to prevent the creation of spurious tokens. Without this mechanism, an item (at the TFL) could be active from one token to the next; being bound to multiple tokens.

Recall

At the conclusion of an RSVP stream, the model simulates recall. There are two phases to this process: (i) *identity retrieval* and (ii) *order retrieval*.

Identity Retrieval. The retrieval system first locates bound tokens (as indicated by active WM trace nodes) and then identifies binding units allocated to each token; if binding units are allocated, the type (or types) "pointed to" by those units is retrieved. Thus, for successful identity recall of a target, a portion of the binding pool must be allocated from any token to that target, and that token must have an active trace node. Multiple identities are retrieved if only one token has been bound, which, through allocated binding units, points to multiple types.

Order Retrieval. If both targets were successfully encoded to either one or two tokens, the recall phase determines in which order they were perceived. This proceeds according to the following algorithm,

1. if either token (tk1 or tk2) is unambiguously bound (i.e. to a single target) then order is uniquely defined,[v]
2. otherwise there is a 50 : 50 chance of either report order.

Note that alternative (1) handles the case in which both tokens are unambiguously bound, since the condition will trivially hold. In addition, alternative (2) will be applied if all token bindings are ambiguous, whether there are one or two such bindings (both situations of which can occur at extreme presentation rates).

4.5 Neural-ST^2's Transient Attentional Enhancement: the Blaster

As discussed in sections 3.2 and the Lag 1 Sparing subsection of section 3.3, there are a number of pieces of AB data that support our TAE. Firstly, lag-1 sparing is suggestive of such a mechanism; see the Lag 1 Sparing subsection of section 3.3. Secondly, Chua et al (2001) demonstrate that a distracter is a more effective prime of a T2 if it is preceded by a T1. Thirdly, Potter et al (2002) show that at very short SOAs (particularly less than 100 ms) T2 performance is in fact better than T1 performance. The implication is that perception of the T1 generates a brief window (around 150ms) of enhanced processing that falls most strongly upon the (possibly distracter) item in the T1+1 slot.

The blaster is depicted in Figure 9 (full implementation details can be found in section A.2 of the appendix). Above threshold activity in any node of the TFL excites the blaster (through the projection marked (a) in Figure 9). However, task demand foregrounding ensures that, in practice, only perception of a target can trigger the blaster, which in turn sends a powerful excitatory projection to all nodes in the Task Filtered and Item layers (through the projections marked (b) in Figure 9). This causes a general excitation of all Item and Task Filtered nodes.

The blaster fires a burst of excitation and then rapidly shuts off. Furthermore, the shut-off is sufficiently strong to induce a refractory period, where the blaster is unavailable.

In addition, inhibition (through the projection marked (c) in Figure 9) while binding units are being allocated maintains the blaster in this off state even when the refractory period is complete.[vi] Therefore, the general behavior of the blaster is a spike of excitation followed by a period of inactivity until the completion of the current token binding. These dynamics protect the integrity of ongoing tokenization by limiting attentional resources that could cause binding intrusions.[vii]

4.6 Neural-ST^2's Binding Pool

Each token provides a powerful inhibitory projection to three quarters of the binding pool, while type nodes in the TFL have excitatory projections to some of the units in the pool (see Figure 14). These projections ensure that, for each combination of type and token, there will be one element of the binding pool that is (1) not inhibited by the token and (2) excited by the type. Thus, the pool can uniquely represent every combination of token and type, as well as combinations of multiple types bound to a single token or multiple tokens bound from a single type.

The binding pool itself is comprised of gate-trace pairs that function much like the nodes comprising each token. When confusion can arise, we use the terms *WM* gate - trace (or token gate - trace) to denote constituents of tokens and *binding* gate - trace to denote constituents of binding units. Within such a binding unit, the gate node is activated by input from the TFL, and in turn excites the corresponding trace node, which slowly accrues activation. When the trace node reaches its threshold it inhibits its corresponding gate node. The self-sustaining activity in the trace node serves to maintain the corresponding type/token binding.

Insert Figure 14 Here

Each binding gate node also excites its corresponding token gate, which serves to activate the token. Continued activation of the type and token nodes is not necessary to maintain the

binding once the binding trace node is active. However, if the type input from the TFL is too weak, the binding trace node will not reach its threshold and that item will not be encoded into WM.

4.7 Summary of the Neural-ST^2 Model

The following reviews the functional significance of each element of the model:

- First Stage:

1. *Input & Masking layers*. Through inhibitory interactions, masking arises at the masking layer.
2. *Item layer.* Weak excitatory links between similar items facilitate priming.
3. *Task Filtered layer.* This layer implements the salience filter. Task demand selects categories of items. The layer also sends feed-forward excitation to gate nodes of binding units and triggers the blaster.

- Intermediary Stages:

1. *Binding Pool.* This provides a binding resource, which enables types to be bound to tokens. Binding units are implemented as gate-trace pairs, yielding a gated active memory. An inhibitory projection from binding gates to the blaster maintains the blaster offline during tokenization.

- Second Stage:

1. *Tokens.* These are also implemented as gate-trace pairs (called WM gate - traces), yielding another gated active memory.
2. *WM Gate nodes.* These gate token access and are also the origin of (binding pool encoded) associations from tokens to TFL items. Strong lateral inhibition ensures that only one token is ever available.

3. *WM Trace nodes.* These actively maintain tokens in a bound state. Thus, at encoding, they ensure that tokens can only be bound once per trial and at recall, they denote that a token has been bound.

Blaster. On detection of a target, this mechanism causes a rapid and temporary increase in activation levels across Item and Task Filtered layers.

4.8 Modeling the Empirical Landscape

The model reproduces a spectrum of AB data. We consider these results in turn.

The Basic Blink

Suppression of the blaster during tokenization is the mechanism that generates the blink. That is, the binding of T1 to a token can take long enough that a weak T2 has subsided in both item and task filtered layers before the blaster is available again. Targets at lags 2 and 3 fall at the point of maximum impairment (see Figure 15(a)). However, the impairment decreases linearly through lags 4, 5 and 6, as it becomes more likely that tokenization of the T1 finishes before T2 has decayed.

Insert Figure 15 Here

Lag-1 sparing (see Figure 15(a)) results from the T2 being close enough in time to T1 to take advantage of the (T1 initiated) blaster firing, allowing it to be encoded into the first token alongside T1. The loss in T1 accuracy at lag 1 (see Figure 15(c)) arises since, when T2 is very strong and T1 weak, binding can complete before T1 is strongly active, yielding a successful binding from token 1 to T2 and no binding to T1. As can be seen in Figure 15(a), the results of the model largely conform to the conventional shape of the AB curve. Manipulating parameters can change the duration, depth and onset of the blink, but the U shape is present whenever the model is operating in accord with its theoretical underpinnings.

Insert Figure 16 Here

Blanks

Placing a blank in the T1+1 slot attenuates the blink; see Figure 15(a). An unmasked item generates a strong activation trace, a situation that greatly shortens tokenization. As a result, tokenization is more likely to have finished releasing the blaster, before the T2 has decayed at the TFL. In contrast, T2+1 blank produces strong T2 traces that are capable of outliving the blink. Human data on T2+1 blanks does not exist, but we were able to compare the model to data from Giesbrecht and Di Lollo (1998), who examined the effect of placing T2 at the end of the stream; see Figure 15(a). T1+2 and T2+2 blanks, on the other hand, have minimal affects on the trace strength of targets, and therefore have little affect on the blink, see Figure 16.

Swaps

In accordance with human data, at lag 1, the model is inaccurate at determining the order of the two targets; see Figure 15(c). At lag 1, typically, both T1 and T2 are bound to the first token and in fact, to a small extent also to the second token. However, T2 is more often uniquely bound to the second token than T1. It is this binding profile that allows the recall process to successfully disambiguate the order of presentation of the two targets on at least some of the lag-1 trials, explaining why swaps are better than chance.

Missed T2 Processed Extensively

We adopt a compromise position on the fate of blinked T2s. None of the type layers in Neural-ST^2 (masking, item or task filtered) are strongly suppressed during the blink. As a result, the model is capable of exhibiting priming between semantically related items even during the blink. That is, types are always, at least to some extent, extracted from a target, regardless of whether it is blinked. However, there are notable differences in the traces of seen and unseen T2s. A T2 that is seen during the blink has the virtue of having a stronger

trace than a missed T2, either by variation of input strength, or by being excited by the blaster, see Figure 17.

Insert Figure 17 Here

The difference in activation strength, between missed and seen T2s increases as we progress through the stage 1 layers. This is apparent in Figure 17, where the difference between masking layer missed and seen traces is very small, but is slightly increased at the item layer and then further accentuated at the TFL. This is broadly consistent with Luck et al (1996) and Vogel et al (1998), where no discernible difference between early ERP components (N1 and P1) for lag 3 vs lag 7 T2s was identified. Such early ERP components would arise from ST^2's masking layer activations. They also found little discernible difference between the N400s elicited by lag 3 and lag 7 T2s, which would emerge from ST^2's item layer. However, Vogel et al (1998) did observe a marked impairment in the P3s generated by lag 3 vs lag 7 T2s, which is likely to correspond to a combination of our TFL and token activations.

The only slight inconsistency with Vogel et al's (1998) results is that at our item layer, the difference in activation between seen and missed T2s is perhaps somewhat larger than the ERP N400 finding. As supported by Rolke, Heil, Streb and Hennighausen (2001), we suggest that, although missed T2s will prime T3s, their priming will not be as strong as for seen T2s, perhaps particularly in respect of the duration of the priming (Rolke et al., 2001). Furthermore, the comparison to Vogel et al's ERP profiles is somewhat mudded by two issues. Firstly, the precise way to relate ERPs and neural network activation (especially those generated by rate coded models) is not well understood. Indeed there is uncertainty about what exactly, neurobiologically speaking, is being recorded in ERP, especially from later ERP components (such as the N400 and P3) (Fell et al., 2004). Secondly, Vogel et al (1998) compare lag 3 with lag 7 N400s, while in Figure 17 we are comparing missed vs seen T2 traces at lag 3, which, in a sense, is a purer measure of the effect of the blink. This is because,

the lag 3 ERPs will contain a proportion of seen T2 trials (since T2 accuracy is never zero) and the lag 7 ERPs will contain missed T2 trials (since T2 accuracy is never 100%). Thus, it may be more difficult to observe N400 differences between missed and seen T2s (which is the point of interest) using Vogel et al's (1998) lag-based differentiation.

Insert Figure 18 Here

Enhanced Processing of the T1+1 slot

The blaster enhances the activation of the T1 and T1+1 slots; see section 4.5. Accordingly, the blink is attenuated by the insertion, at the T1+1 position, of a semantic prime for the T2 (Chua et al., 2001); see Figure 18. This arises because the T1+1 distracter weakly excites its semantic associates (one of which is T2) via the lateral excitatory links in the Item layer. We do not provide a quantitative fit to Chua et al's (2001) data since they were working with a color-marked task (and in fact, identity rather than semantic priming), which is beyond the scope of Neural-ST^2.

For comparison purposes, Figure 18 includes the performance of the model on the basic blink and T1+1 blank conditions. This makes clear that the priming benefit arising from a T1+1 distracter is relatively short-lived. Thus, if the (primed) T2 arrives at lags 3 or 4, it benefits from its proximity to the semantically related T1+1 distracter. However, at later lags (e.g. 5 or 6), Item layer activation of the T1+1 distracter has decayed sufficiently that the (primed) T2 receives little, if any, benefit.

4.9 Neurophysiological Correlates

There are three functional elements that we will discuss in relation to proposed brain regions: Stage 1 (type processing), Stage 2 (token and working memory maintenance) and the Blaster (transient attention). Firstly, there is a strong correspondence between the levels of ST^2's stage 1 and sequential steps of the ventral visual processing pathway, comprising a cascaded

highly parallel system (see, Rousselet, Thorpe, & Fabre-Thorpe, 2004 for a review). In the visual system of the macaque, early stages (V1-V2) are thought to have small receptive fields, making them susceptible to spatially overlapped masking, much like the masking layer within ST^2. As one ascends the hierarchy of visual areas, to V4 and TE cortex (the latter being analogues to late stages of temporal cortex in humans), spatial specificity decreases as the processing gives way to progressively more identity specific information (Rousselet et al., 2004). The same shift occurs within our model in progressing from the masking layer, with its strong spatial masking, to the item layer, which represents the semantic relationships between items.

Type nodes in the TFL correspond well to cells in the late stages of the ventral visual processing stream, a position supported by recordings from IT (Inferior Temporal) cortex in macaques performing a WM task (Miller, Erickson, & Desimone, 1996). In their data, IT neurons selective for stimuli encoded into WM did not sustain activity during the presentation of other items. This is exactly the profile of activity produced by type nodes in ST^2, as activation of a type is only necessary during encoding and retrieval.

Sustained activity during maintenance is the province of tokens and the binding pool. In Prefrontal Cortex (PFC), Miller et al (1996) found cells that have sustained activity and are insensitive to stimulus identity, as would arise from trace nodes of tokens. In contrast, other cells in PFC sustained activity and were stimulus selective. These cells could correspond to trace nodes within the binding pool.

Neurophysiological support for trace nodes in humans can be seen in a recent fMRI study of the AB (Marois, Do-Joon, & Chun, 2004). Of particular note, Marois et al (2004) identified a notably greater response in a lateral frontal region for seen vs missed T2s, which remained pronounced even at the end of the recording period. This may be the marker of sustained activation of two active tokens (in the seen case) vs one (in the missed case). The

work of Kranczioch, Debener, Schwarzbach, Goebel and Engel (2005) made a similar point. They demonstrated that in an AB setting, when participants saw the T2, activity in frontal and parietal areas (Inferior Frontal Gyrus, Lateral Frontal Cortex and Inferior Parietal Lobe) was elevated in a sustained manner, while there was no sustained elevation in the ventral stream (Lateral Occipital Complex and Fusiform Gyrus). In our model, WM trace nodes would perform such sustained activation. In contrast, areas of the brain which would correspond more closely to type processing (Parahippocampal Place Area (PPA) in Marois et al (2004) and Fusiform Gyrus / Lateral Occipital Complex in Kranczioch et al (2005)) showed temporary differences between T2 seen and T2 missed trials. The localisation of the token system to prefrontal areas is also consistent with an extensive literature of imaging and lesion studies that demonstrate how prefrontal areas contribute to WM encoding, maintenance and retrieval (Desimone, 1996; Passingham & Saka, 2004). In addition, Petrides (1995) demonstrated that monkeys with lesions of the mid-dorsal part of lateral frontal cortex have a permanent deficit in their ability to recall the temporal order of items.

The neurophysiological implementation of transient attention may correspond to a putative attentional mechanism comprising right-lateralized areas of the Temporo-Parietal Junction (TPJ) and Ventral Frontal Cortex (VFC), for a review, see Corbetta and Shulman (2002). These areas of the brain are thought to detect both exogenously and endogenously salient targets and rapidly orient attention to them. For example, in a recent imaging study by Serences et al. (2005), participants monitored a central RSVP stream for red letters while ignoring two flanking streams. When flanking distracters switched from gray to a non-red color, they stood out strongly but failed to capture attention. Only when the flanking items were the same color as the target was a contingent capture of attention observable as a reduction in target detection. On presentation of red (target colored) flanking distracters, the right TPJ and VFC were strongly activated, in contrast to when non-target colored distracters

were presented. This cortical circuitry may work in concert with subcortical mechanisms, such as the Locus Coeruleus, see section 6.8. As described by Aston-Jones, Rajkowski and Cohen (2000), within a visual discrimination paradigm, monkey LC neurons are activated with a temporal profile that matches closely what our blaster would predict. Namely, in such a paradigm, LC neurons were strongly activated for a period of 100 ms, peaking about 150 ms after a target was presented. Following this activation, LC neurons appeared to be inhibited for a further 100 ms. These data suggest that the LC and the TPJ-VFC networks may cooperate to rapidly deploy transient attention in response to detected targets with cortical mechanisms filtering stimuli through the frontally mediated task set and triggering the LC to coordinate the deployment of attention.

Validation of Model

5 Predictions and their Empirical Validation

Section 4.8 has shown that Neural-ST^2 does a good job of modeling existing AB findings. Here we generate predictions from ST^2 and validate these predictions empirically.

5.1 Prediction 1

Within the ST^2 theory, the temporal dynamics of AB sparing are dictated by the TAE, which fires for a (broadly) fixed window of time. The most direct prediction of this mechanism is that sparing is primarily dictated by the temporal gap between targets, rather than their sequential separation. By doubling the rate of presentation from 100 to 50 ms per item, we should observe that sparing extends out to lag-2. That is, items are spared at short lags not because they are directly adjacent to the T1, but because they occur within 100 ms of it.

Figure 19 (on the left) demonstrates the model's performance at 100 and 50 ms SOAs. These data are sorted by the time between targets (i.e. Target Onset Asynchrony), rather than

the number of items between targets (i.e. lag). The 100 ms/item data is the basic blink condition of Figure 15. For the faster presentation rate, data are shown for Target Onset Asynchronies (TOAs) of 100-800ms, corresponding to lags 2, 4, 6, 8, 10, 12, 14 and 16.[viii] Importantly, there is strong sparing even for lag-2, and the general time course of the blink is similar to that for the slower presentation rate. The reduced baseline accuracy (easily seen at lag-8) is the result of the weaker traces produced by rapid presentation rates.

Insert Figure 19 Here

To test this prediction we ran an experiment, described in the 54ms SOA Experiment subsection in section A.3 in the appendix, which presents participants with letter targets in a stream of digit distracters at 54 ms/item. As predicted, the AB conformed to the conventional temporal characteristics, with sparing at 100 ms TOA and recovery well under way by a 500 ms TOA; see Figure 19 on right. These data are directly compared to the results of a second experiment (see the 94ms SOA Experiment subsection in section A.3 in the appendix) that used a 94 ms SOA. Comparing these results with a 2x8 ANOVA provides a main effect for lag ($F(7, 176) = 12.34$, $p < 0.00001$), and for presentation speed ($F(1,176) = 76.69$, $p < 0.00001$), but no interaction ($F(7,176) = 0.39$, $p > 0.9$). The lack of an interaction indicates the absence of an effect of presentation speed on the time course of the AB. These results stand counter to models of the AB which place primary importance on the T1+1 distracter as an initiator of the blink, including the interference theory (Shapiro, Arnell, & Raymond, 1997) and the Temporary Loss of Control theory (Di Lollo, Kawahara, Shahab Ghorashi, & Enns, 2005).

5.2 Prediction 2

The onset of transient attention is delayed relative to T1 and extends for some time thereafter. Typically, this time course ensures that a T2 at lag-1 is perceived more accurately than T2's that themselves trigger the blaster (i.e. at lag 8). In effect, T1 acts as a cue for T2 in the same

way that cues in Nakayama and Mackeben (1989) improved performance for the cued item. We observe this pattern within the model. If we switch to a non-conditional analysis (T1 may fire the blaster even when it is too weak to be reported), we see an improvement in T2 at lag-1 relative to all other lags. In the model, unconditional T2 accuracy stabilizes at 86% for lags 7 and 8, but is 89% at lag-1. This effect is even more pronounced for a faster presentation rate as performance is farther from ceiling. In the model, the baseline accuracy of 58% at lags 14 and 16 increases to 75% at lag-2 (100 ms TOA), for a 50 ms SOA.

This prediction is supported by a reanalysis of the data collected for prediction 1 (methods reported in the first two subsections of section A.3), which examines the raw (i.e. non-conditional) accuracy of T2. Using a planned series of T tests for the slow condition we see a marginal, but significant, improvement of T2 during sparing. T2 is 90% (std error 2%) at lag-1 and 83% (std error 3%) at lag-8. This difference is significant (T of 2.74, df=13 gives $p < .02$, two-tailed).

Turning to the fast presentation rate, where we expect an even more pronounced effect, T2 at lag-2 (equivalent to lag-1 at the slower rate) is 75% (std error 5%), and at lag 16 (equivalent of lag-8) is 59% (std error 4%). This difference is significant (T of 3.06, df = 9 gives $p < .015$ two-tailed). These results confirm that, not only is T2 spared at lag-1, but it can also exhibit a net accuracy improvement. The results from predictions 1 and 2 support the TAE hypothesis. Indeed, it is hard to see how raw T2 could be superior at early lags than post blink recovery without a T1 initiated transient enhancement.

5.3 Prediction 3

As discussed in the Lag 1 Sparing subsection of section 3.3, we claim that sparing may be indicative of a failure of selective attention. That is, the processing mechanisms revealed by the AB may exist to exclude T2 when it threatens to interfere with T1, but, at lag-1, there is insufficient temporal resolution to exclude the second target. Putative failure of attention

appears as sparing at lag 1 for two reasons. First, in the majority of AB studies, temporal order of T1 and T2 is either implicit in the task-switch between the targets (e.g. Raymond et al., 1992) or order report is not requested of participants (e.g. Chun & Potter, 1995). Second, targets in AB studies are almost always unique, individually discriminable items that cannot be mis-combined into other possibly valid targets. What this means is that T1 and T2, even when bound into the same token at lag-1, can be (unambiguously) separated in a post-hoc retrieval.

We predict that an AB experiment that requires participants to maintain both the temporal order of T1 and T2, as well as the binding of sub-elements within targets, will produce accuracy at lag-1 that is worse than at any other lag. The experiment to test this prediction demands that participants report the identity of both T1 and T2, in the correct order. Furthermore, T1 and T2 each consist of a letter pair, which must be accurately paired at recall. Thus, given a T1 of "AB" and a T2 of "CD", responses of "CD" "AB" would be counted as errors, as would "AD" "CB". We ran this experiment, as described in the Letter Pairs subsection in section A.3 in the appendix.

Insert Figure 20 Here

Figure 20 shows that at lag-1, T2 exhibits no appreciable sparing and T1 performance is at just 20% having dropped from a baseline of about 65%. By combining these (the (T1+T2)/2 curve), we index the efficiency with which participants extract correctly bound targets (T1 or T2), in the proper temporal order. There is a sharp reduction in (T1+T2)/2 at lag-1 that is more severe than the decline at any other lag; see Figure 20. In a planned T-test, lag 1 was significantly lower than lag 2 ($T = 7.5$, $df = 11$ gives $p < .00001$).

These data strongly support our position that use of the term sparing to describe the lag 1 case, is, in many respects, an artifact of the form of classic AB experiments. In the sense that, in these experiments, 1) the order in which targets are detected is not assessed and 2)

component features of T1 and T2 are not confusable. Our results show that when these constraints are removed, sparing disappears at lag 1; indeed it is the lag at which the least information is extracted. We thus have obtained evidence to support our position that the blink is a mechanism to prevent the arrival of T2 from interfering with the ongoing T1 binding. Furthermore, lag 1 is a breakdown of the system in the sense that the T2 arrives so rapidly after T1 that the system is unable to withhold attentional resources from T2.

Comparison with Other Models

There are a number of AB models, which we compare with ST^2 in this section. We begin by consider the applicability of three general neural mechanisms. Then we consider five formal models and finally, we discuss four informal theories of the AB.

6 Neural Models

General Neural Mechanisms

We consider the applicability of three neural mechanisms for modeling the AB.

6.1 Active Working Memory

WM is often modeled as sustained neural activation (Davelaar et al., 2005; Davelaar & Usher, 2004). In a localist approach, this is realized by a self-sustaining node; see Figure 7(a). Our TFL uses aspects of localist active memories, such as self-loops. However, we also include off nodes with relatively rapid activation dynamics (certainly in relation to the decay rate of WM). As a result, TFL nodes only self-sustain for brief periods (a few hundred ms) before being suppressed by their off nodes. So, why not do away with these off nodes and implement an active WM at our TFL?

Unfortunately, pure active memories do not reflect the types-tokens distinction. For example, since type representations are not freed up, pure active memories have difficulties representing multiple instances of the same item, i.e. they are permanently repetition blind. In addition, there is no mechanism available to represent temporal order. (These issues were discussed in section 2.3.) Extra mechanisms could no doubt be added in response to these limitations. However, such an approach would not be obviously simpler or more theoretically appealing than the tokenization approach we are advocating.

6.2 Winner-take-all

A simple way to generate an AB would be to enforce lateral inhibition between task relevant items at, what would become, a winner-take-all layer. For example, in our model, we could impose strong lateral inhibition at the TFL and view that as the "output" from the model, i.e. remove the WM token system and simplify TFL nodes by removing their self-loops and shut-off systems. At such a winner-take-all layer, T1 activation would suppress a T2 arriving during the blink. In order to explain lag 1 sparing, this approach would have to somehow argue for a delayed onset of T1 inhibitory pressure. In addition, the approach would explain blink recovery through T1 activation decay, which would eventually release the layer from inhibition. While pleasingly simple, this approach is unsatisfactory for a number of reasons.

1. *WM Encoding*. Such an approach would not explain how items are *encoded (and maintained)* in WM. Rather it would just generate short-lived activation traces at the (final) winner-take-all layer. One would then have to argue that activation that crosses a particular threshold at the winner-take-all layer would be interpreted as reported. But, there would be no mechanism provided to explain how those items would be maintained until the point at which report occurs. In particular, by the nature of this approach, item nodes cannot self-sustain at the winner-take-all layer, because if they did, there would be no blink recovery, since T2s would be permanently suppressed by T1 activation.

2. *Late AB Bottleneck.* In addition, it is unclear how to reconcile such an approach with the known late locus of the AB, see section 3.2. It is difficult to be unequivocal about this point, since it is dependent upon the interpretation imposed upon layers of the model. However, if the winner-take-all layer was interpreted as a component of the representation of a type, then a missed T2 would have a greatly reduced (compared to seen T2s), and probably non-existent, type activation. This may well imply absence of a semantic representation of missed T2s.
3. *Types-tokens Distinction.* A winner-take-all approach would also fail to code temporal order. Although, the approach could potentially avoid the problem possessed by active memories of permanent repetition blindness. This is because target activation would eventually decay (this is required to ensure blink recovery) and thus, a (sufficiently delayed) repetition could be represented. However, avoidance of permanent repetition blindness is only obtained because a temporally sustained representation is not generated, see point 1 above.
4. *Lag 1 Sparing.* Finally, a winner-take-all approach would not give an illuminating explanation of lag 1 sparing. That is, if one views the lateral inhibition as reflecting interference between T1 and T2, then one is still left with the problem of explaining why T1 and T2 do not interfere at lag 1.

6.3 TAE Unavailability as a Refractory Period

A number of models use approaches similar to our Transient Attentional Enhancement, e.g. see section 6.8; that is, a generalized (in terms of types) and short-lived enhancement that is initiated by the detection of an AB target. Furthermore, they all generate a blink by the TAE being offline when the T2 arrives. There is though an important issue concerning how the TAE is held offline. In ST^2, two mechanisms contribute to the TAE being held offline, 1) a blaster refractory period and 2) ongoing tokenization. The first of these is common to all the

TAE based blink models. However, it is only ST^2 that also includes the second. This is important since, refractory periods have a fixed time-course, which is locally controlled and not tied to ongoing processes elsewhere in the model.

However, tying TAE unavailability to tokenization fits with the position that the system tries to prevent a second target from interfering with ongoing tokenization. This, in turn, fits with the ST^2 principle of ascribing discrete episodic contexts. In more practical terms, ongoing tokenization causing the TAE to be withheld explains why blink length/depth varies with bottom-up trace strength. That is, the shorter the time to tokenize, the shorter TAE unavailability and thus, the shorter the blink. Approaches that just tie TAE unavailability to its refractory period lack this link between tokenization and blink length. As a result, they typically find it more difficult to explain how unmasking T1 attenuates the blink.

Formal Neural Models

We now discuss the available formal neural models of the AB.

6.4 The Gated Auto-associator

The Model

Chartier, Cousineau and Charbonneau (2004) present a connectionist model of a color marked digits AB task, see Figure 21. It contains an input layer, which feeds into two competitive networks, performing number and color identification respectively. The number identification system feeds into an auto-associator, which acts as a WM. When the representation of a digit becomes active in this auto-associator, a learning rule is used to adjust the weights to auto-encode (into WM) the activation pattern.

Gating mechanism. A gating mechanism sits between the number identification system and the auto-associator. The opening of this gate is regulated by a comparison system, which compares the color being identified with the target color. If the colors match, the gate is

opened and the digit currently activate at the identification network is encoded into WM. However, opening the gate starts an inhibitory process, which ensures that the comparison generated by the T1+1 item takes significantly longer than for the T1. Comparison efficiency progressively improves from the T1+2 item onwards until it eventually returns to its original level.

Insert Figure 21 Here

Auto-association. Items are maintained in WM via auto-associator weights. The strength of these weights, i.e. the strength of encoding, is a function of the extent to which the gate is open while the item is represented at the identification layer. The system "recalls" the two items that are most strongly encoded into WM, i.e. possess the strongest weights.

How the Blink arises. The key mechanism that generates the blink is the inhibitory process. Thus, the increase in comparison time resulting from the inhibitory process ensures that the gate opens later for a T2 presented during the blink. Such T2s are thus only weakly encoded into WM and it is likely that a distracter is more strongly encoded, which would be recalled in preference to the T2.

Blink recovery is governed by the length of time it takes this inhibitory process to subside. Lag 1 sparing is obtained since a T2 at lag 1 benefits from the gate remaining open following T1 encoding. Thus, in this case, T2 encoding is largely independent of the T2 comparison process.

Assessment and Comparison to ST^2

Chartier et al provide a valuable AB model. However, it has a somewhat different intent from Neural-ST^2, since it is targeted at color marked data, rather than letters-in-digits data. Thus, it is difficult to make a direct comparison, but we offer the following observations.

- Firstly, the approach only reproduces a small set of AB phenomena when compared with Neural-ST^2. It reproduces a blink curve with lag 1 sparing.
- Although not a strictly active memory approach, since items are stored as weights rather than as sustained activation, the Gated Auto-associator does suffer some of the problems of such approaches. For example, it seems that the model would also be permanently repetition blind. This is because both instances of an item would generate identical auto-associator activation. Thus, a repeated item would be extremely strongly encoded (in the sense of weight strength), but there would be no representation of multiple tokens.
- In addition, it seems that the Gated Auto-associator cannot represent serial order. Although, there is an indication to the contrary in Chartier et al (2004), in which they claim that their model generates inversion errors (i.e. swaps), but no supporting mechanism or data is provided.
- The Gated Auto-associator relies upon a very fast weight change process, which would be somewhere in the 100 ms time frame. This is difficult to support from a neurobiological standpoint, see section 2.5.
- The Gated Auto-associator does not contain an analogue of our masking system. Thus, it is difficult to see how the model could reproduce blink attenuation with target + 1 blank.
- Finally, the approach seems to be consistent with a late bottleneck view of the AB. Specifically, assuming that the competitive layers perform all processing prior to WM encoding, then missed T2s should obtain identical pre-encoding activation profiles as reported T2s.

6.5 The Global Workspace Model

The Model

The Global Workspace model provides an elegant and compelling theory of conscious perception and attentional control (Dehaene, Sergent, & Changeux, 2003). The model includes neural processing pathways (see Figure 22) from early sensory regions (areas A and B), e.g. visual cortex, through to higher association areas of temporal, parietal, frontal and cingulate cortex (areas C and D) (Dehaene et al., 2003). Stimuli compete to recruit a global workspace sitting at the top of these pathways and which, once recruited, affords global exclusive access. That is, the workspace,

"… mobilises excitatory neurons with long distance axons, capable of interconnecting sensory and high-level areas into global brain-scale states". (Dehaene et al., 2003)

Exclusivity of brain-scale states (i.e. that consciousness is unitary) arises since neurons in the workspace inhibit surrounding neurons (the lateral inhibition in Figure 22). Furthermore, reverberating activation between constituent neurons generates a pressure for brain-scale states to be self-sustaining. The model is biologically detailed; for example, it employs spiking neurons and is anatomically prescribed. In particular, cortico-thalamic columns act as building blocks for the model, see Figure 22.

Insert Figure 22 Here

How the Model Blinks

The AB version of the Global Workspace model contains processing pathways from visual input for both the target stimuli: T1 and T2; see Figure 22. The firing rate of pyramidal cells coding T2 in the highest two areas (C and D in Figure 22) is taken as an index of T2 reportability. The blink arises because the brain-scale state generated by the T1 suppresses

competing stimuli and thus, reduces their reportability index. Recovery from the blink arises when the T1 brain-scale state has subsided, freeing the workspace for late lag T2s.

Assessment and Comparison with ST^2

We now discuss the contribution of the Global Workspace model in the context of the AB.

- The model does not exhibit as broad a spectrum of data as ST^2; it generates a blink curve and a fit with electrophysiological data is highlighted. We return to both these issues in later points.
- Dehaene et al (2003) generate a U shaped blink curve. However, the blink curve in figure 3A of Dehaene et al (2003) is rather short, with accuracy returning to baseline by a 200ms T1-T2 lag, while human blink curves tend to return to baseline by lag-5, i.e. a T1-T2 lag of around 500ms, e.g. see Figure 15.
- Figure 3A and the results section of Dehaene et al (2003) indicate that they have obtained lag 0 sparing, i.e. that T2 is spared if it is presented simultaneously with T1, which is not a typical AB data point in which items are presented in sequence. In addition, the AB curve in Figure 3A is at its deepest point at T1-T2 lags of 50ms and 100ms, which is the region at which lag 1 sparing is obtained in human data, see for example Figure 15. However, the parameter settings used have been drawn from the monkey literature and, as indicated by the authors, they may not be consistent with the human system.
- Dehaene et al (2003) acknowledge the limitations that arise from not implementing masking. This means that this model cannot address the impact of blanks in the stream. In addition, since RSVP distracters are not represented at all in the Global Workspace simulation, the model is unable to address the capacity of T1+1 distracters to prime future targets (Chua et al., 2001).

As previously highlighted, the key mechanism in the Global Workspace model that enables it to blink, is the lateral inhibition between thalamo-cortical columns in areas C and D. Effectively, this means that the model implements a winner-take-all solution. Thus, the model has many of the characteristics discussed in section 6.2.

- Finally, there is no representation of serial order in the model.

6.6 The Conflict Monitoring Model

Elements of the Model

Battye's (2003) conflict monitoring model employs a stimulus-response pathway, which includes localist input and output layers connected in a one-to-one fashion (see Figure 23). Attentional control arises from an attention layer, which contains two nodes, denoted location and letter. The former of these gives an attentional boost to stimuli presented in a particular spatial location, while the latter foregrounds the task set. This foregrounding is realized through projections from each attention node to the relevant input nodes. In fact, since all items are presented in the same spatial location in RSVP, the location node projects to all input nodes. Importantly, the attention to input layer projections are bi-directional. Thus, in addition to top-down (endogenous) attentional control, the model also supports bottom-up (exogenous) control (from the input to the attention layer). This bi-directionality generates a Transient Attentional Response (TAR), similar to ST^2's TAE.

The final mechanism implemented is conflict monitoring, which takes inspiration from theories of cognitive control (Botvinick et al., 2001). Conflict is a measure of the extent to which multiple responses are co-active and are thus interfering. Importantly, the strength of lateral inhibition at the output layer is adjusted according to conflict. Thus, when there is response conflict, competitive pressures are increased, in order that the most active response can more easily dominate.

Insert Figure 23 Here

How the Battye Conflict Monitoring Model Blinks

There are two mechanisms (discussed below) that cause the blink,

1. adjustment of output layer lateral inhibition due to conflict; and
2. a Transient Attentional Response (TAR).

Lateral Inhibition and Conflict. As previously highlighted, the value of output layer lateral inhibition is proportional to the current value of conflict. Thus, co-activation, at the output layer, of T1 and the T1+1 distracter generates a short period of high conflict, which pushes output layer lateral inhibition up, closing the door on the T2. The door stays closed while the T1 and T1+1 distracter responses are co-active.

Transient Attentional Response (TAR). In the Battye model, the T1+1 distracter is only strong enough to cause conflict at the output layer because it benefits from the T1 initiated TAR. When T1 is active at the input layer, it excites the location node in the attention layer; but since this node projects to the whole of the input layer a window of recurrent activity boosts the T1+1 distracter. However, once it has fired the TAR enters a refractory period, which contributes to the blink.

Resulting Behavior. Thus, in Battye's model, the blink arises from the following sequence of events. 1) The T1 initiates a brief window in which the input layer is boosted (the TAR); 2) the T1+1 item benefits from this boost; 3) co-activation of T1 and T1+1 at the output layer generates conflict; 4) this causes increased lateral inhibition at the output layer; 5) the now strong lateral inhibition causes T2s to be suppressed by the already strongly active T1 and (to a lesser extent) T1+1; and thus, 6) the T2 only yields a weak output layer representation. Blink recovery arises because the T1 and T1+1 output layer representations eventually subside, which releases the system from conflict and resets lateral inhibition to zero. In

respect of lag-1 sparing, a T2 at lag 1 will benefit from the TAR and will yield a strong output layer representation before conflict is registered.

Assessment and Comparison to ST^2

A strength of Battye's model is its simplicity and that it fits within an accepted connectionist-modeling framework: GRAIN (McClelland, 1993). A number of aspects of Battye's model have similarities to ST^2 mechanisms. Perhaps most significantly, Battye's attention layer could be viewed as a conflation of our task demand and TAE. Specifically, Battye's letter node is an almost direct analogue of Neural-ST^2's task demand unit, while his location node is a TAE. This said, an important difference with ST^2 is that, effectively, Battye's TAR has a fixed refractory period. As a result, Battye's approach suffers the problems we discussed in section 6.3 associated with not tying TAE unavailability to ongoing tokenization.

This difference is reflected in the manner in which Battye's model generates blink attenuation with T1+1 blank. The ST^2 model explanation is that increased bottom-up trace strength of an unmasked T1 speeds up T1 tokenization. In contrast, in Battye's model, a T1 followed by a blank will generate less conflict. This is because, in the T1+1 blank case, the distracter after T1 (at T1+2) obtains a reduced benefit from the TAR and, in any case, arrives a little too late to generate substantial conflict with the T1 at the output layer.

In addition, there is more of a serial (by-item) element to the blink in Battye's model. This arises because in Battye's model, the blink is initiated by the increase in conflict caused by the arrival of the T1+1 item. Thus, a delay in the onset of this distractor, which arises in the T1+1 blank condition, will itself delay the onset of increased output layer conflict. Battye justifies the approach on the grounds that in Chun and Potter (1995) there is evidence of lag 2 sparing in the T1+1 blank condition. Although, it should be noted that the ST^2 model obtains a degree of lag 2 sparing in the T1+1 blank condition (see Figure 15(a)) without subscribing to this theoretical position. Also, our experimental verification of prediction 1 (see section

5.1) stands against the serial aspect of Battye's model. Although Battye's conflict monitoring approach certainly makes an important contribution, there are a number of limitations.

1. Battye's model does not generate as broad a spectrum of data as the ST^2 model. In particular, there is no consideration of impaired T1 performance and increased swaps at lag 1 and blink attenuation with T2 unmasked. However, inclusion of the TAR, suggests that Battye's model is theoretically consistent with increased processing of the T1+1 distracter (Chua et al., 2001).
2. The theoretical justification for using a conflict monitoring approach in the context of the AB is not as clean as it may initially seem. In particular, Battye is hypothesizing a much faster adjustment of cognitive control than that previously proposed. For example, in Botvinick et al (2001), cognitive control is updated per trial and not at the 100ms conflict detection to adjustment of control speed that Battye requires.
3. Probably most significantly, Battye's approach seems incompatible with a late bottleneck. This is because, in Battye's model, the key to the blink is the enforcement of strong lateral inhibition in order to suppress the T2. Thus, the model effectively employs winner-take-all and thus, it suffers the problems discussed in section 6.2, which includes incompatibility with a late bottleneck
4. As was identified in section 6.2 as a general problem of winner-take-all models, Battye's conflict monitoring model also does not provide an explicit model of how items are encoded, maintained and retrieved from WM.

Insert Figure 24 Here

6.7 CODAM

The Model

Fragopanagos, Kockelkoren and Taylor (2005) model the AB in the context of Taylor's COrollary Discharge of Attention Movement (CODAM) model (Taylor, 2002; Taylor & Rogers, 2002), which has made an important contribution to the modeling of attentional control. The model contains a bottom-up pathway, comprising the Input and Object Map modules, along which input stimuli have to pass in order to reach the Working Memory module, see Figure 24. Items are encoded into WM in an active memory fashion (see, section 6.1). Critically, items in the Object Map require an attentional boost in order to successfully progress to the Working Memory. This boost is provided by an attentional control signal generated by the IMC (Inverse Model Controller).

How the CODAM Model Blinks

In an AB setting, the attentional control signal is withheld for the T2, in order to prevent its encoding interfering with the encoding of the T1. Understanding this mechanism requires us to consider the Monitor, Endogenous Goals (part of the Goals module) and Corollary Discharge, see Figure 24. The Monitor has the capacity to suppress items in the IMC that are interfering with ongoing target encoding. The mechanism by which the Monitor determines when it should selectively inhibit the IMC is slightly more involved, but the key aspect is that the Monitor is continuously computing a comparison between the activity in the Endogenous Goals and the Corollary Discharge. The former of these maintains a representation of the current target and the latter a predictor of the attended stimulus by taking a copy of the attention control signal, i.e. the IMC.

The blink arises because T1 remains the current target until its encoding at the WM layer is complete. Thus, if the system attempts to boost the T2, through the IMC, before T1

encoding is complete, the Monitor will detect an error, since the current endogenous goal is T1 and the corollary discharge is representing T2. As a result, the Monitor will suppress the T2 in the IMC and thus, retract its attentional boost, which will in turn prevent T2 from being successfully encoded into WM. When T1 working memory encoding is complete the system recovers from the blink by handing the endogenous goal from T1 to T2. Consequently, the Monitor will not register an error when comparing the current endogenous goal (now T2) with the Corollary Discharge representation of T2.

Assessment and Comparison to ST^2

There are some similarities between the CODAM approach and the ST^2 model. Firstly, the activation profiles shown in Fragopanagos et al (2005) suggest that the CODAM approach is consistent with a late bottleneck. Secondly, both hypothesize that this bottleneck arises to prevent interference during WM encoding. However, in instantiating this theoretical position, the two are rather different.

A strength of the CODAM approach is that it simulates the AB within the context of a broad scope and mature model of attentional control, which also has ties to neurobiology. In addition, it produces a larger spectrum of data than the other approaches discussed in this section. It generates a basic (U-shaped) blink curve with lag-1 sparing, as well as blink attenuation with T1+1 blank. However, the CODAM approach does not generate as broad a spectrum of data as the ST^2 model. In particular, it does not reproduce blink attenuation with T2 unmasked, decline in T1 performance at lag-1, increased temporal order confusion at lag-1 and increased processing of the T1+1 distracter.

In some senses, CODAM and the ST^2 model have different ambitions: the starting point for CODAM is neurobiology, whereas ST^2 began as a cognitive-level explanation of the AB and temporal attention in general. Thus, ST^2 has been driven by a desire to reflect, for example, the types – tokens distinction. In addition, the Fragopanagos et al approach is

subject to the characteristics of pure active memory approaches (see section 6.1): permanent repetition and temporal order blindness.

In the CODAM approach, recovery from the blink is tied to task hand-over: from the T1 to the T2 task. This may have some degree of plausibility in the context of blink paradigms with a task switch (e.g. Raymond et al., 1992). However, other AB paradigms, such as the letters-in-digits task, do not possess a task switch between the two targets, yet there is still a blink and a blink recovery. In addition, it has been argued that a task switch confounds the AB (Chun & Potter, 2000; Potter, Chun, Banks, & Muckenhoupt, 1998). Thus, as implemented, the CODAM approach has difficulty fully explaining blink recovery, certainly in the context of experimental paradigms without a task switch.

6.8 The Locus Coeruleus Model

The Model

This recently proposed model explains the AB according to the functioning of the Locus Coeruleus (LC) (Nieuwenhuis, Gilzenrat, Holmes, & Cohen, 2006), a minute structure (German et al., 1988), which projects to almost every region of the brain, with a special emphasis on areas involving attentional processing (Aston-Jones et al., 2000). The LC seems to be activated in response to the detection of a salient item. In addition, the effect of LC innervation includes amplification of excitatory responses for feature selective cells and, in a visual discrimination paradigm, monkey LC neurons were activated with a temporal profile that seems to match the AB. This observation prompted the Nieuwenhuis et al (2006) model; see Figure 25. The model can be divided into two components: the behavior network (comprising the input, decision and detection layers) and the LC.

Insert Figure 25 Here

The behavior network is a simple feed-forward network, with one-to-one inter-layer connections. Crosstalk connections are also included between input and decision layer nodes, reflecting feature similarity between stimuli. In addition, excitatory self-loops are included to sustain activation at decision and detection nodes. However, these loops are not strong enough to yield an active memory. Finally, as a reflection of its role, nodes in the decision layer compete through lateral inhibition.

The LC circuit modulates activity in the behavior network. Specifically, the LC is excited by detection of a salient stimulus, since target nodes in the decision layer project to the LC node. LC activity has a modulatory effect on the behavior network, simulating the release of norepinephrine. This release multiplicatively scales the afferent signals to network units, transiently adjusting their gain. Importantly though, after firing, the LC enters a refractory period. It is unavailability of the LC during this period that causes the blink. At lag-1, T2 benefits from the LC firing initiated by the T1, generating sparing.

Assessment and Comparison to ST^2

Nieuwenhuis et al's LC model makes an important contribution to understanding the AB. A strength is that the model is framed within the context of a broad neurophysiological theory of attentional function. There are also important similarities between the approach and the ST^2 theory, most notably, between the LC and ST^2's TAE. Both are initiated by detection of a salient stimulus and their temporal characteristics are similar.

Nieuwenhuis et al's model is also consistent with a temporal, rather than sequential (by-item) interpretation of blink onset. That is, it suggests that the blink has a fixed temporal dynamic, the onset of which is not regulated by intervening distracters, as explored in prediction 1 (see section 5.1)[ix]. This aspect of Nieuwenhuis et al's model suggests that it is also consistent with increased processing of the T1+1 slot (Chua et al., 2001). There are though differences between the approaches.

1) Perhaps most importantly, Nieuwenhuis et al's model suffers the problem we discussed in section 6.3. In fact, in Nieuwenhuis et al's model the length of the LC refractory period is not fixed. Rather, stronger LC firings lengthen the LC refractory period, which, in turn, means that greater bottom-up trace strength leads to a longer refractory period. This facet of Nieuwenhuis et al's model has the consequence that unmasking T1s (i.e. T1+1 blank conditions) would lengthen the blink. This stands contrary to the data, which, as previously discussed, shows unmasking attenuates the blink.[x]
2) ST^2's TAE is an additive enhancement, while the LC enhances by increasing the signal to noise ratio. The relative benefit of these two approaches is not yet clear. However, it is important to note that there is a gating aspect to Nieuwenhuis et al's mechanism, which is not present with ST^2's TAE.
3) In addition, while the TAE is assumed to be location specific, this is not the interpretation associated with the LC. This is an important point of comparison that should yield experiments to differentiate the two methods.
4) As acknowledged by Nieuwenhuis et al (2006), their model suffers the common problem that (unlike ST^2) nothing is actually sustained to retrieval. That is, T1 and T2 activations rise and fall at the detection layer with a time-course in the range of a few hundred milliseconds of simulated time.
5) Finally, Nieuwenhuis et al's model does not reproduce as extensive a set of AB data as ST^2 and there is no handling of temporal order.

6.9 Synthesis Across Models

We now concisely summarize the AB modeling landscape. For this purpose, Table 1 shows commonalities across models. There is a further AB model (Barnard & Bowman, 2004), which we do not discuss, since it is not a neural network.

Insert Table 1 Here

7 ST^2 and Informal theories of the Attentional Blink

We now relate ST^2 to informal theories of the blink.

7.1 The Interference Theory

The interference theory (Shapiro, Arnell et al., 1997) is a prominent explanation of the AB. We discuss how our model relates to this theory now.

Serial Order. Current formulations of the interference theory do not consider serial order. It is assumed that the four relevant items (T1, T2, T1+1 and T2+1) are all placed in a buffer and weights are associated with these items. However, these weights bias the competition between items in the buffer, rather than enabling serial order to be retrieved. It is though clear that some order information can be extracted from RSVP streams and thus, that such information must be recorded somewhere and must somehow impact upon retrieval. [xi]

However, in respect of ordering, perhaps the most critical difference between the ST^2 model and interference theory is that the latter seems to imply a sequential rather than temporal initiation of the blink. That is, that the arrival of the T1+1 item is the key initiator of the blink, as, for example, encapsulated in Battye's conflict monitoring model. Thus, if the T1+1 distracter is delayed it is suggested that the blink profile will be slowed. This is in contrast to the ST^2 model, which, as previously emphasized, assumes that the blink profile is time rather than item dependent. Verification of prediction 1 in section 5.1 (and also the experimental findings in Nieuwenhuis et al. (2006)) are consistent with the ST^2 approach in this respect.

Categorical Interference vs Low-level Masking. The manner in which the T1+1 and T2+1 items disrupt processing of their respective targets is a hotly debated topic. There are two extreme positions:

1. *Low Level Masking.* The disruption is purely at the low level of visual features, i.e., distracters following targets visually mask targets (Seiffert & Di Lollo, 1997).
2. *Categorical Interference.* The disruption occurs at a categorical level, perhaps even at retrieval (Isaak, Shapiro, & Martin, 1999). Thus, the more categorically similar targets and target+1 distracters are, the more confusable they are and the more difficult it is to encode / retrieve the correct targets.

Our model selects (1) over (2). That is, it is assumed that weaker visual traces are the key cause of the AB. Masking weakens neural activation traces, thereby requiring more time to bind tokens at later stages. Recent data are generally consistent with the selection of (1) over (2) (Giesbrecht, Bischof, & Kingstone, 2004). Efforts to adjust blink depth by manipulating semantic similarity between T1 and the T1+1 distracter have generally not been successful.

Chun and Potter used mathematical symbols and found that categorically similar distracters caused deeper blinks than dissimilar distracters (Chun & Potter, 1995). However, it seems likely that the mathematical symbols used were not as effective visual masks as digits. In particular, Maki, Bussard, Lopez and Digby (2003) attempted to replicate Chun and Potter's findings using false fonts that were matched for pixel density with digits and letters. From a categorical perspective, these false fonts should have been less similar to digits or letters than the symbols, yet no blink attenuation was observed. Their own experiments, using symbols as masks, did replicate the Chun and Potter findings by strongly attenuating the blink. The converse of this experiment (i.e. using letters as distracters and symbols as targets) failed to find an attenuated blink. These data stand against the categorical interference hypothesis, suggesting that visual featural masking properties of distracters are most important in determining blink depth.

7.2 Temporary Loss of Control

An important new theory of the blink is based upon the idea that T1 processing generates a period in which the system is vulnerable to a Temporary Loss of Control (TLC) (Di Lollo et al., 2005). Thus, a central processor is initially allocated to maintaining an input filter (similar to our salience filter) in a state in which it filters out distracters; only allowing targets to pass. However, T1 detection initiates a period in which the central processor is occupied with T1 stimulus processing and response planning. As a result, the central processor is unable to maintain the input filter and the system becomes vulnerable to exogenous control. The arrival of a distracter item will then trigger a change in the system's configuration, causing it no longer to be optimally tuned to the target category. It is this reconfiguration that makes the system vulnerable to missing T2s.

The key findings that motivate the TLC theory are that sequences of three or even four consecutive RSVP targets seem immune to the blink (Di Lollo et al., 2005; Olivers, van der Stigchel, & Hulleman, 2005). In terms of the ST^2 model, this intriguing data suggests that TAE unavailability during tokenization is not absolute and that it may re-fire in response to sustained TFL activation. However, there are a number of empirical questions that need to be addressed before informed modeling of these data can be undertaken. Firstly, Weichselgartner and Sperling (1987) in their RSVP study found two glimpses in a consecutive target paradigm; see section 8.3. Why Weichselgartner and Sperling identified a blink-like effect when their participants had to encode four items in a row, while Di Lollo et al and Olivers et al did not find such an impairment needs to be explored. The Weichselgartner and Sperling task is indeed somewhat different to Di Lollo et al and Olivers et al's, but exactly why that changed the profile of data needs to be clarified.

In addition, in order to model Di Lollo et al and Olivers et al's data the profile of order errors amongst the three / four targets needs to be known. This would indicate which items

are bound to which tokens. Nobody has reported order information in three or four target experiments to date. Finally, there is existing data that current formulations of the TLC find difficult to explain. For example, the findings of our prediction 1 (see section 5.1) and the experiments contained in Nieuwenhuis et al. (2006), suggest that the blink is temporal, rather than sequential (i.e. by-item). In contrast, the TLC theory would suggest that the first distracter after the T1 should initiate the blink; this is not what we find in our 50ms SOA data.

7.3 Central Interference Theory

ST^2 sits relatively well with Jolicoeur (1998) and co-worker's Central Interference Theory (CIT). This is not surprising since the CIT has a good deal in common with Chun and Potter's (1995) 2-stage model; for example, it assumes that a parallel first stage meets a sequential second stage, which is involved in encoding into short term memory. However, Jolicoeur and co-workers emphasize the spectrum of tasks that can interfere with such short-term memory encoding, e.g. response selection, ongoing short term memory encoding of a different item and task switching (Jolicoeur, 1998). Furthermore, as a reflection of the claimed central nature of the bottleneck, it is asserted that the process of encoding into short-term memory is essentially amodal (Arnell & Jolicoeur, 1999).

A particularly strong commonality with ST^2 is that the CIT emphasizes how the ease of the T1 task affects the length of the blink, since the longer encoding of T1 into short-term memory takes, the more decayed the waiting T2 will be by the time it can be encoded. ST^2 provides a concrete realization of this position[xii]. With regard to Jolicoeur and co-workers claim that the bottleneck is central and essentially amodal, ST^2 is largely neutral. ST^2 is currently formulated in the visual modality, as a reflection of the data that we have been attempting to reproduce. However, a reformulation to accommodate an amodal locus of interference could be explored.

Discussion and Concluding Remarks

8 ST^2 and Temporal Attention in General

8.1 Basic Principles

The ST^2 theory is a general framework for considering visual temporal attention and working memory encoding. It is based upon five principles.

1. *Two Stages.* A cascaded parallel first stage is followed by a serial second stage that is closely tied to working memory encoding.
2. *Salience Filter.* A filter emphasizes salient and de-emphasizes non-salient items.
3. *Types-tokens.* The association of a token with an active type is the key mechanism involved in ST^2 WM encoding.
4. *Transient Attentional Enhancement (TAE).* In response to the detection of a salient stimulus, the TAE provides a brief (but spatially specific) enhancement of type representations.
5. *Binding Pool.* Through the binding pool, types and tokens can be arbitrarily associated, order information can be recorded and degenerate type - token bindings can be represented, e.g. multiple types can be bound into the same token. In addition, since type resources are freed-up once an item has been tokenized, further instances of the same item can be perceived and encoded.

As evidence of the success of the ST^2 framework, we can see how it satisfies the five functional requirements we set out for our theory in section 1.2.

1. *Salience Detection.* ST^2's salience filter enables rapid detection of salient stimuli when briefly presented in demanding visual environments.

2. *Sustain Representations.* The TAE sustains fleeting, but salient, representations, enabling encoding into working memory.
3. *Ascribe Episodic Context.* The types-tokens system enables discrete episodic contexts to be ascribed to items.
4. *WM Maintenance.* Since multiple tokens are available, many items can be maintained together in WM.
5. *Repetitions.* Multiple tokens can be bound to the same type, compactly maintaining multiple instances of the same item in WM.

We have also provided Neural-ST^2, a connectionist realization that is faithful to the principles of the ST^2 theory. In respect of neural network modeling, our approach is relatively high-level. In particular, we have not modeled individual biological neurons, but rather, nodes in our model correspond to assemblies of (perhaps many thousands of) biological neurons. This is an appropriate level of abstraction to work at for a number of reasons. Firstly, the main constraints on our modeling are cognitive-level behavioral data; in particular, single cell recordings in such a setting are not available. Secondly, our theoretical position is most naturally elaborated in terms of high-level neural dynamics; it is unclear that low-level details of biological implementation would be illuminating. In particular, the systems-level at which we have worked has enabled us to develop a relatively simple and computationally tractable instantiation of our theory. Finally, the level of modeling we work at corresponds to that employed in a large number of connectionist models and has been convincingly justified (O'Reilly & Munakata, 2000).

In terms of theories of attention, it is also interesting to note that we have not required any direct analogue of biased competition (Desimone & Duncan, 1995) in ST^2. This could be because, in the temporal attention setting, items compete in time, rather than space. Biased competition usually has a spatial focus.

8.2 Tokens, Object Continuity and Object Files

The ST^2 framework takes a good deal of inspiration from Chun's (1997) paper. By exploring AB and Repetition Blindness effects within a single RSVP set-up, Chun articulated an informal types-token theory of the AB. This theory made the important link between T1 tokenization and the blink, i.e., that during the blink, the T2 type was locked out of WM encoding, while the T1 type was being associated with a token. However, there are a number of respects in which ST^2 extends and diverges from Chun's theory. Firstly, ST^2 makes the strong claim that the central role of distracters in generating the blink is to weaken target traces through masking. In contrast, Chun argued for higher-level categorical similarity effects. See section 7.1 for a justification of the position we have taken on this topic. Secondly, Chun made a distinction between object tokens and spatio-temporal tokens; this distinction has not been required in the current formulation of ST^2. Thirdly, and most importantly, the link between TAE availability and tokenization that, in ST^2, is actually the direct mechanism for generating the blink, was not present in Chun (1997). In this sense, ST^2 elaborates an exact mechanism by which the gate can be closed on T2 while T1 is being tokenized.

A number of recent studies have revealed a tantalizing effect of object constancy on the AB (Kellie & Shapiro, 2004; Raymond, 2003). The central message here is that, if there is object constancy across the items of an RSVP stream, e.g. the items represent the gradual morphing of one object into another (Kellie & Shapiro, 2004), then targets marked by small featural changes within the stream will not generate a blink. These findings suggest a link between the AB and object file updating. That is, in a stream with object constancy, a single object file is set-up, which is continually updated as the object changes. However, importantly, these object updates, even when target discriminating features are updated, do

not generate a blink. Rather, it is the creation of new object representations that generates a blink[xiii].

Although ST^2's tokens are not currently as representationally rich as object files, the two concepts clearly have similarities. For example, modulo the lag-1 case, when presented with an AB stream, ST^2's tokens will come to represent distinct objects, with the featural attributes of those objects associated with tokens through the binding pool. Thus, it could be that the token binding and handover mechanism implemented in ST^2 should be viewed as the creation of a new WM object. This is the subject of ongoing research.

8.3 Sperling & Co-workers' Theory of Temporal Attention

A prominent and powerful theory of temporal attention is that developed by Sperling and co-workers (Shih & Sperling, 2002; Sperling & Weichselgartner, 1995; Weichselgartner & Sperling, 1987). They investigated tasks where a window of attentional engagement is cued in which a sequence of (usually four) items presented in RSVP had to be reported. The bimodal profile of temporal performance obtained by Weichselgartner and Sperling's (1987) procedure two is particularly noteworthy. They identified a first glimpse in which the 1st and 2nd items after the cue were well reported; this was followed by significantly impaired performance for the 3rd item and then a second glimpse in which performance on the 4th and 5th items recovered. The ST^2 theory would explain this profile in terms of two token/blast episodes separated by a window of blaster unavailability[xiv].

Sperling and co-worker's theory seeks to explain trajectories between location specific attentional episodes. Although it does not explicitly address spatial relocation of attention, ST^2 explains temporal attention in terms of discrete token allocation events, which resonates with Sperling and co-worker's emphasis on discrete attentional episodes. In addition, ST^2's TAE plays a similar role to Sperling and co-worker's attentional gate, which opens in response to detection of a cue. However, Sperling and co-worker's attentional gate is

multiplicative, while ST^2's TAE is additive. Most significantly though, the ST^2 theory assumes a close coordination between tokenization and TAE availability, whereby the TAE fires once and then is held offline during tokenization. This TAE - tokenization interaction is central to ST^2's explanation of the blink and is not present in Sperling and co-worker's theory.

Finally, Sperling and co-worker's models propose that temporal order can be extracted from the activation strengths with which items are encoded into working memory. This approach would suggest that at lag-1, T1 - T2 swaps would be higher than chance, since T2 is typically reported better than T1 at this lag (Potter et al., 2002). However, swaps are below chance at lag-1; see Figure 15-d. Such data has motivated our token-based approach to temporal order encoding.

8.4 Concluding Remark

As is in fact the case with all models, there is no sense to which the Simultaneous Type Serial Token model is in an absolute sense the right model of temporal attention. However, we do believe that the ST^2 model illustrates a set of issues that would need to be addressed by any competitor model. Most prominent amongst these is the types-tokens distinction. Some mechanism is needed by which an instance specific marker of the occurrence of a type is held in WM. Furthermore, such a marker needs to record order information, accommodate repetition of types (beyond the repetition blindness window) and enable types to be regenerated during retrieval. Our binding pool satisfies these criteria in a neurally plausible manner.

References

Amit, D. J., Bernacchia, A., & Yakovlev, V. (2003). Multiple-object Working Memory-A Model for Behavioral Performance. *Cerebral Cortex, 13*(5), 435-443.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411*(6835), 305-309.

Arnell, K. M., & Jolicoeur, P. (1999). The Attentional Blink Across Stimulus Modalities: Evidence for Central Processing Limitations. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 630-648.

Aston-Jones, G., Rajkowski, J., & Cohen, J. (2000). Locus coeruleus and regulation of behavioral flexibility and attention. *Progress in Brain Research, 126*, 165-182.

Bal, T., & McCormick, D. A. (1993). Mechanisms of oscillatory activity in guinea-pig nucleus reticularis thalami in vitro: a mammalian pacemaker. *The Journal of Physiology, 468*, 669-691.

Barnard, P. J., & Bowman, H. (2004). Rendering information processing models of cognition and affect computationally explicit: Distributed executive control and the deployment of attention. *Cognitive Science Quarterly, 3*(3), 297-328.

Battye, G. (2003). *Connectionist Modelling of Attention and Anxiety.* Unpublished PhD, The Medical Research Council's Cognition and Brain Sciences unit, Cambridge.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652.

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

Chartier, S., Cousineau, D., & Charbonneau, D. (2004). A Connectionist Model of the Attentional Blink Effect During a Rapid Serial Visual Task. In *ICCM 2004, International Conference on Cognitive Modelling*.

Chua, F. K., Goh, J., & Hon, N. (2001). Nature of codes extracted during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 27*(5), 1229-1242.

Chun, M. M. (1997). Types and tokens in visual processing: a double dissociation between the attentional blink and repetition blindness. *Journal of Experimental Psychology: Human Perception & Performance, 23*(3), 738-755.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance, 21*(1), 109-127.

Chun, M. M., & Potter, M. C. (2000). The attentional blink and task switching within and across modalities. In K. Shapiro (Ed.), *Temporal Constraints in Human Information Processing*. Oxford: Oxford University Press.

Cohen, J. D., & Huston, T. A. (1994). Progress in the Use of Interactive Models for Understanding Attention and Performance. In *Attention and Performance XV* (Vol. 15, pp. 453-476).

Coltheart, M. (1983). Iconic memory. *Philosophical Transactions of the Royal Society London B Biological Sciences, 302*(1110), 283-294.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201-215.

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

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review, 112*(1), 3-42.

Davelaar, E. J., & Usher, M. (2004). An extended buffer model for active maintenance and selective updating. In H. Bowman & C. Labiouse (Eds.), *Proceedings of the 8th Neural Computation and Psychology Workshop* (Vol. 15, pp. 3-14). Singapore: World Scientific.

Deco, G., Rolls, E. T., & Horowitz, B. (2004). "What" and "where" in Visual Working Memory: A Computational Neurodynamical Perspective for Integrating fMRI and Single-Neuron Data. *Journal of Cognitive Neuroscience, 16*(4), 683-701.

Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences of the USA, 100*(14), 8520-8525.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the USA, 93*, 13494-13499.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Di Lollo, V., Kawahara, J., Shahab Ghorashi, S. M., & Enns, J. T. (2005). The attentional blink: resource depletion or temporary loss of control? *Psychological Research, 69*(3), 191-200.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*(1), 53-78.

Duncan, J. (1981). Directing attention in the visual field. *Perception and Psychophysics, 30*(1), 90-93.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise-letters on identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143-149.

Fell, J., Dietl, T., Grunwald, T., Kurthen, M., Klaver, P., Trautner, P., et al. (2004). Neural Bases of Cognitive ERPs: More than Phase Reset. *Journal of Cognitive Neuroscience, 16*(9), 1595-1604.

Fragopanagos, N., Kockelkoren, S., & Taylor, J. G. (2005). A Neurodynamic Model of the Attentional Blink. *Cognitve Brain Research, 24*(3), 568-586.

German, D. C., Walker, B. S., Manaye, K., Smith, W. K., Woodward, D. J., & AJ., N. (1988). The human locus coeruleus: computer reconstruction of cellular distribution. *Journal of Neuroscience, 8.*(5), 1776-1788.

Giesbrecht, B., Bischof, W. F., & Kingstone, A. (2003). Visual masking during the attentional blink: tests of the object substitution hypothesis. *Journal of Experimental Psychology: Human Perception & Performance, 29*(1), 238-258.

Giesbrecht, B., Bischof, W. F., & Kingstone, A. (2004). Seeing the light: Adapting luminance reveals low-level visual processes in the attentional blink. *Brain and Cognition, 55*(2), 307-309.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: visual masking by object substitution. *Journal of Experimental Psychology: Human Perception & Performance, 24*(5), 1454-1466.

Hebb, D. O. (1949). *The Organization of Behaviour*. New York: John Wiley and Sons.

Hotson, J. R., & Prince, D. A. (1980). A calcium-activated hyperpolarization follows repetitive firing in hippocampal neurons. *Journal of Neurophysiology, 43*(2), 409-419.

Hummel, J. E., & Biederman, L. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*(3), 480-517.

Isaak, M. I., Shapiro, K. L., & Martin, J. (1999). The attentional blink reflects retrieval competition among multiple rapid serial visual presentation items: tests of an

interference model. *Journal of Experimental Psychology: Human Perception & Performance, 25*(6), 1774-1792.

Jolicoeur, P. (1998). Modulation of the Attentional Blink by On-line Response Selection: Evidence from Speeded and Unspeeded $Task_1$ Decisions. *Memory and Cognition, 26*(5), 1014-1032.

Kanwisher, N. G. (1987). Repetition blindness: type recognition without token individuation. *Cognition, 27*(2), 117-143.

Kanwisher, N. G. (1991). Repetition blindness and illusory conjunctions: errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception & Performance, 17*(2), 404-421.

Kellie, F. J., & Shapiro, K. L. (2004). Object File Continuity Predicts Attentional Blink Magnitude. *Perception and Psychophysics, 66*(n), 692-712.

Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences, 6*(3), 120-125.

Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons*: Oxford University Press.

Kovacs, G., Vogels, R., & Orban, G. A. (1995). Cortical Correlate of pattern backward masking. *Proceedings of the National Academy of Sciences of the U.S.A (Neurobiology), 92*, 5587-5591.

Kranczioch, C., Debener, S., Schwarzbach, J., Goebel, R., & Engel, A. K. (2005). Neural correlates of conscious perception in the attentional blink. *Neuroimage, 24.*(3), 704-714.

Kristjansson, A., Mackeben, M., & Nakayama, K. (2001). Rapid, object-based learning in the deployment of transient attention. *Perception, 30*(11), 1375-1387.

Kristjansson, A., & Nakayama, K. (2003). A primitive memory system for the deployment of transient attention. *Perception and Psychophysics, 65*(5), 711-724.

Liaw, J. S., & Berger, T. W. (1996). Dynamic synapse: a new concept of neural representation and computation. *Hippocampus, 6*(6), 591-600.

Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55*, 207-234.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature, 383*, 616-618.

Maki, W. S., Bussard, G., Lopez, K., & Digby, B. (2003). Sources of interference in the attentional blink: target-distractor similarity revisited. *Perception and Psychophysics, 65*(2), 188-201.

Maki, W. S., Couture, T., Frigen, K., & Lien, D. (1997). Sources of the attentional blink during rapid serial visual presentation: perceptual interference and retrieval competition. *Journal of Experimental Psychology: Human Perception & Performance, 23*(5), 1393-1411.

Marois, R., Do-Joon, Y., & Chun, M. M. (2004). The Neural Fate of Consiously Perceived and Missed Events in the Attentional Blink. *Neuron, 41*, 465-472.

McClelland, J. L. (1993). Toward a theory of Information-processing in graded, random and interactive networks. In *Attention and Performance* (pp. 655-688).

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience, 16*(16), 5154-5167.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140.

Muller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and Voluntary Orienting of Visual Attention: Time Course of Activation and Resistance to Interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15*(2), 315-330.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*(11), 1631-1647.

Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., & Cohen, J. D. (2006). The Role of the Locus Coeruleus in Mediating the Attentional Blink: A Neurocomputational Theory. *Journal of Experimental Psychology: General (in Press)*.

Nieuwenstein, M. R., Chun, M. M., van der Lubbe, R. H. J., & Hooge, I. T. C. (2005). Delayed Attentional Engagement in the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1463-1475.

Olivers, C. N., van der Stigchel, S., & Hulleman, J. (2005). Spreading the sparing: against a limited-capacity account of the attentional blink. *Psychological Research, 8*, 1-14.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain.*: MIT Press.

Pantic, L., Torres, J. J., Kappen, H. J., & Gielen, S. C. (2002). Associative memory with dynamic synapses. *Neural Computation, 14*(12), 2903-2923.

Passingham, D., & Saka, K. (2004). The prefrontal cortex and working memory: physiology and brain imaging. *Current Opinion in Neurobiology, 14*(2), 163-168.

Petrides, M. (1995). Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience, 15.*(1 Pt 1), 359-375.

Potter, M. C. (1993). Very short-term conceptual memory. *Memory and Cognition, 21*(2), 156-161.

Potter, M. C., Chun, M. M., Banks, B. S., & Muckenhoupt, M. (1998). Two attentional deficits in serial target search: the visual attentional blink and an amodal task-switch deficit. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 979-992.

Potter, M. C., Staub, A., & O'Conner, D. H. (2002). The time course for competition for attention: attention is initially labile. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1149-1162.

Raymond, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science, 14*(1), 54-59.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception & Performance, 18*(3), 849-860.

Rolke, B., Heil, M., Streb, J., & Hennighausen, E. (2001). Missed prime words within the attentional blink evoke an N400 semantic priming effect. *Psychophysiology, 38*(2), 165-174.

Rolls, E. T., Tovee, M. J., & Panzeri, S. (1999). The Neurophysiology of Backward Visual Masking: Information Analysis. *Journal of Cognitive Neuroscience, 11*(3), 300-311.

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends in Cognitive Sciences, 8*(8), 363-370.

Seiffert, A. E., & Di Lollo, V. (1997). Low-Level Masking in the Attentional Blink. *Journal of Experimental Psychology: Human Perception & Performance, 23*(4), 1061-1073.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of Voluntary and Stimulus-Driven Attentional Control in Human Cortex. *Psychological Science, 16*(2), 114-122.

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The Attentional Blink. *Trends in Cognitive Sciences, 1*(8), 291-297.

Shapiro, K. L., Caldwell, J., & Sorensen, R. E. (1997). Personal names and the attentional blink: a visual "cocktail party" effect. *Journal of Experimental Psychology: Human Perception & Performance, 23*(2), 504-514.

Shapiro, K. L., Driver, J., Ward, R., & Sorensen, R. E. (1997). Priming from the Attentional Blink: A Failure to Extract Visual Tokens but Not Visual Types. *Psychological Science, 8*(2), 95-102.

Sheppard, D. M., Duncan, J., Shapiro, K. L., & Hillstrom, A. P. (2002). Objects and events in the attentional blink. *Psychological Science, 13*(5), 410-415.

Shih, S. I., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review, 109*(2), 260-305.

Sperling, G., & Weichselgartner, E. (1995). Episodic Theory of the Dynamics of Spatial Attention. *Psychological Review, 102*(3), 503-532.

Taylor, J. G. (2002). Paying attention to consciousness. *Trends in Cognitive Sciences, 6*(5), 206-210.

Taylor, J. G., & Rogers, M. (2002). A control model of the movement of attention. *Neural Networks, 15*(3), 309-326.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136.

Triesman, A. M. (1998). Feature Binding, Attention and Object Perception. *Philosophical Transactions of the Royal Society London B., 353*(1373), 1295-1306.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance, 24*(6), 1656-1674.

von der Malsburg, C. (1981). *The Correlation Theory of Brain Function (Technical Report)* (No. 81-2). Gottingen, West Germany: Max-Planck Institute for Biophysical Chemistry.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of Automatic and Controlled Visual Attention. *Science, 238*, 778-780.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12), 1120-1135.

Wyble, B., & Bowman, H. (2005). Computational and experimental evaluation of the attentional blink: Testing the simultaneous type serial token model. In B. G. Bara, L. W. Barsalou & M. Bucciarelli (Eds.), *CogSci 2005, XXVII Annual Conference of the Cognitive Science Society* (pp. 2371-2376): Cognitive Science Society through Lawrence Erlbaum.

Wyble, B., & Bowman, H. (2006). A neural network account of binding discrete items into working memory using a distributed pool of flexible resources. [abstract & online presentation]. *Journal of Vision, 6*(6), 33a.

Tables

	ST^2	Chartier et al (Gated Autoass.)	Dehaene et al (Global Worksp.)	Battye (Conf. Monit.)	Fragopanagos & Taylor (CODAM)	Nieuwenhuis et al (Locus Coeruleus)
Lateral Inhib between Type Representations	×	×	✓	✓ [a]	✓ / × [b]	✓
TAE with Refractory Period	✓	×	×	✓	×	✓
TAE is Location Specific	✓	NA	NA	✓	NA	×
TAE withheld during tokenization	✓	NA	NA	×	NA	×
Target Reps. Maintained to Retrieval	✓	✓	×	×	✓ / × [c]	×
Active Maintenance of Entire Type	×	× [d]	×	×	✓ / × [e]	×
Maintenance of Pointer to Type (Tokens)	✓	×	×	×	×	×
Uses Discrete Episodic Contexts (Tks.)	✓	×	×	×	×	×
Extraction of Temporal Order	✓	×	×	×	×	×

Table 1: Commonalities between AB models. Note the following, [a]) lateral inhibition only enforced when conflict is high; [b]) depends upon what are interpreted as types (Object Map or WM nodes); [c]) there is no reason why WM nodes could not sustain until retrieval, although they do not seem to do so in the current implementation; [d]) weights are maintained, but activation is not; and [e]) same point as [c]). NA denotes Not Applicable.

Layer	*Leak*	*Number of nodes*
1. Input	.07	12
2. Masking	.01	12
3. Item	.04	12
4. Item Off	.02	12
5. TFL	.02	12
6. TFL Off	.07	12
7. Binder Gate	.07	6
8. Binder Trace	.01	6
9. WM Gate	.07	3
10. WM Trace	.01	3
11. Blaster Input ON	.07	1
12. Blaster Input OFF	.07	1
13. Blaster Output ON	.07	1
14. Blaster Output OFF	.07	1

Table 2: Parameter settings per layer.

	Value	*Type*	*Threshold*
Excitatory Connections			
1->2	.022	1	0.15
2->3	.014	1	0.15
3->5 *	.015	1	0.15
3->3	.0095	1	0.15
4->4	.0095	1	0.4
6->6	.01	1	0.4
5->5	.022	1	0.4
5->6	.0075	1	0.52
3->4	.02	1	0.52
9->10	.0055	1	0.4
10->10	.06	1	0.15
5->11	.018	2	0.15
13->5 **	.8	2	0.48
13->3 **	.3	2	0.48
11->13 **	4	1	0.15
13->14	.5	1	0.7
7->12	.01	2	0.4
7->8	.0039	1	0.52
8->8	.01	1	0.15
11->12 **	5	1	0.15
13->13	.04	1	0.15
12->12	.01	1	0.15
Inhibitory Connections			
1->2	-.105	3	0.15
2->2	-.06	3	0.15
6->5	-.12	1	0.4
4->3	-.15	1	0.52
10->9	-.6	1	0.4
12->11	-1	1	0.15
14->13	-1	1	0.15

* Saturating connection: output stops at .2, limiting output dynamic range.

** Saturating connection: output stops at .01, limiting output dynamic range.

Table 3: Connectivity between Layers. Value is the weight setting, type the connectivity type (see Figure 26) and threshold the value of θ used in the pre-synaptic output equation. Note, the Item Off, TFL Off nodes, are

weakly self-excitatory (4->4, 6->6 & 12->12); which was not discussed in the Type Processing subsection of section 4.3 or in section A.2.

Appendix

A Implementation and Experimental Methods

A.1 Parameters Settings

The following parameters are fixed across all Neural-ST^2 layers: (from Equation 1) DT_VM = 1.2, EE = 3, EI = -0.2, EL = 0 and (from Equation 2) γ = 5. Table 2 gives parameter settings per layer. Different connectivity types arise in Neural-ST^2; see Figure 26(A). Table 3 uses these types to detail inter layer connectivity, where layer numbering is as per Table 2. We illustrate connections between elements of the token binding process directly, with the value of weights explicitly shown; see Figure 26(B).

Insert Table 2 Here

Insert Figure 26 Here

Insert Table 3 Here

A.2 Blaster Implementation

The blaster implementation is shown in Figure 27. All connections in this module saturate at a very low level, so the value of nodes, once a threshold is crossed, is essentially constant with increasing activation. The blaster implementation comprises two on-off circuits, distinguished as Blaster Input, which governs initiation of the blaster, and Blaster Output, which governs the blaster's output effect. Thus, the blaster is initiated by TFL excitation of the Blaster Input on node (via projection (a)). However, the resulting activation is short lived and rapidly curtailed when the Blaster Input off node feeds inhibition back onto the Blaster Input on node (generating the blaster refractory period). Furthermore, this suppression is maintained by ongoing binding, via the projection marked (c).[xv]

The resulting brief on node activation of the Blaster Input circuit has the role of instigating Blaster Output activation through the link marked (d). This Output circuit generates a temporally fixed output to the TFL and Item layer (via the projection marked (b)). This effect is obtained through the Blaster Output on node having different thresholds for outputting along different connections, where threshold (1) is less than (2) is less than (3).

When the Blaster Output on node is excited, it begins its self-excitatory cycle with a low threshold (1), causing it to ramp up with a predictable time course. When activation crosses threshold (2), it begins outputting at a fixed level over projection (b). On crossing the third threshold (3), it strongly excites the Blaster Output off node, which in turn inhibits it, ending the blaster's output at the appropriate time.

Insert Figure 27 Here

A.3 Experimental Methods

54 ms SOA Experiment

Prediction 1 (see section 5.1) involves perception of letter targets in a stream of digit distracters. MATLAB v6.5 and the psychophysics toolbox[xvi] were used to present RSVP streams to participants on a Windows 2000 computer. Display timing rates and screen refresh synchrony were confirmed with photodiodes. Participants were positioned in front of a 17" computer screen, displaying a white background, upon which black letters in 180 point Arial font were used as targets. Black digits in 220 point Arial were used as distracters. The letters I, M, O, Q, S, W, X and Z, as well as digits 1 and 0, were excluded from all sets. Digits were 5 cm tall and approximately 3 cm wide, resulting in a viewing angle of 7.1 by 4.3 degrees. Letters were 82% of this size.

Participants were University of Kent students, with normal or corrected to normal vision and who spoke English fluently. 10 participants undertook this experiment. Items were

presented at 54 ms SOA, followed immediately by the next item. RSVP streams were 40 items long. Each trial was prefaced by a 500 ms fixation cross in the location of the stream. T1 could appear in slots 9 to 17, while T2 could appear at even lags 2-16 afterwards. At least 8 distracters followed T2. Participants saw 96 two-target trials and 10 catch trials (with only one target) per block, for 3 blocks for a total of 288 critical trials. Instructions presented on each trial asked participants to "Enter all of the letters you saw in order, then press Enter". Participants were allowed to correct their input with backspace and were not given feedback. They were encouraged to guess if they were not sure, but not to guess blindly. Trials were considered correct if participants got the correct identity, without regard to temporal order.

94 ms SOA Experiment

The comparison data in Figure 19 was a subset of the results collected in a second experiment that was similar to the above. 14 University of Kent students participated. Items were presented at 94 ms SOA, followed immediately by the next item. RSVP streams were 18 items in length. T1 could appear in slots 5 to 8, while T2 could appear from lags 1-8 afterwards. At least 4 distracters followed T2.

With equal probability, items could have a blank in the T1 +1 slot, the T2 +1 slot, or no blank at all. Trial types were crossed in an 8x3 paradigm, with lag and blanks as primary factors. There were 144 trials per block, with three blocks per participant and 14 catch trials with no T2. The T1+1 blank and T2+1 blank trials were discarded for this comparison.

Letter Pairs Experiment

On each trial, an RSVP stream was presented at 115 ms/item. Each stream consisted primarily of a sequence of digit trios (147, 258, 369, 470, 581, 692, 814, 703, 925), presented in 100 point Arial font. Embedded within this stream were one or two targets selected from a prepared set of almost 12,000 possible four-tuples of letters.[xvii] Characters were 2.4 cm in

height. Digit trios were approximately 5.5 cm in width and letter pairs approximately 4.4 cm. Viewing angles were approximately 3.4 degrees in height by 7 degrees in width. 12 University of Kent students participated in this experiment.

For the two targets in the stream, no repeats of individual letters were permitted. T1 could appear from lag 4 to 9, T2, if present, would follow T1 by lags 1-8 and there were at least 3 distracters following T2. Each participant viewed 3 blocks consisting of 96 two-target trials and 10 catch trials for a total of 288 critical trials. After each trial, participants were asked first to “Enter the first pair of letters you saw in order, then press Enter” and the same for the second pair.

Acknowledgments

The UK Engineering and Physical Sciences Research Council supported this research (grant number GR/S15075/01). We would also like to thank Phil Barnard, Mary Potter, Patrick Craston, Mark Nieuwenstein, Dinkar Sharma, Marvin Chun, David Huber, Kyle Cave, Randall O'Reilly and an anonymous reviewer for discussions and input that have contributed to the development of the ST^2 model.

The Matlab code that implements the Neural-ST^2 model is available on the following page, http://www.cs.kent.ac.uk/people/staff/hb5.

Figure Captions

Figure 1: Depiction of ST^2 processing. Items 1 and 2 are progressing simultaneously and independently through stage 1. The salience filter is enhancing the i1 channel and suppressing the i2 channel, shown as a change in the darkness of the rectangles depicting items. The system is in the process of binding a stage 1 item / type into WM.

Figure 2: i1 bound to token 1.

Figure 3: ST^2 binding pool, pre any tokenization. Note, the connectivity shown would be duplicated across type and token layers.

Figure 4: ST^2 binding pool in a state arising from co-activation of token 1 and the i1 type. Fill denotes what the binding unit codes, i.e. neither i1 nor tk1, just tk1, just i1 or both i1 and tk1.

Figure 5: Serial position curves for letters-in-digits tasks, with a 100ms SOA. Conditions were, no blank (i.e. the basic blink), a blank at T1+1, a blank at T1+2 and T2 as end of stream. Due to the insertion of blanks, some data points do not exist. The Basic Blink, T1+1 and T1+2 blank are adapted with permission from fig. 4 of Chun and Potter (1995) "A two-stage model for multiple target detection in rapid serial visual presentation." Journal of Experimental Psychology: Human Perception & Performance, 21(1), 109-127, American Psychological Association. T2 End of Stream is adapted with permission from the “No Mask” data in fig. 2(B) of Giesbrecht and Di Lollo (1998) "Beyond the attentional blink: visual masking by object substitution." Journal of Experimental Psychology: Human Perception & Performance, 24(5), 1454-1466, American Psychological Association**.**

Figure 6: Serial position curves of pure T1 performance, pure T2 performance and percentage of T1 – T2 order inversions (swaps) for letters-in-digits task. This is adapted with permission from the data in Table 1 and the “inversion among digits” bars of Figure 8 in Chun and Potter (1995) "A two-stage model for multiple target detection in rapid serial visual presentation." Journal of Experimental Psychology: Human Perception & Performance, 21(1), 109-127, American Psychological Association.

Figure 7: (a) A basic node with a self-excitatory link; (b) a basic node with self-excitation and an inhibitory interneuron; and (c) a basic (non-self sustaining) node with a self-excitatory off node.

Figure 8: Activation dynamics of a gate-trace circuit, which is shown in three consecutive states: (A) a stable (resting) ready state; (B) a dynamic state in response to extrinsic excitation; and (C) a stable active maintenance state.

Figure 9: Structure of Neural-ST^2. Connection patterns between layers are only depicted for some nodes, but apply for the entire layer.

Figure 10: Input, Masking, Item and Task Filtered layer activation traces for a target when it is masked (bottom panel) and unmasked (top panel). Notice that masking layer traces have similar maximum amplitudes in both cases, but the unmasked trace is considerably longer than the masked trace. Broadly speaking, this trace length difference at the Masking layer is turned into amplitude differences at the Item and Task Filtered layers. In particular, the TFL and Item traces are shorter in the unmasked case, because tokenization completes more rapidly and thus, off nodes cut in more quickly.

Figure 11: The type layers of Neural-ST^2. In all cases, apart from the task demand projections, the connectivity depicted for individual nodes extrapolates to all nodes in the layer. We show the constituent components of the Task Filtered and Item layers, i.e. self-loops and off node circuits. Thus, this figure expands out the gray filled nodes in Figure 9.

Figure 12: The token system. Connectivity shown for an individual node is repeated for each node in the given layer. As depicted, all binding pool units are unallocated.

Figure 13: Token dynamics. Bias strength is indicated by point size of the arrowhead denoting the bias.

Figure 14: The binding pool in action.

Figure 15 a-d: The model's performance (a, c) compared to human data (b, d). T2 performance (a, b) represents the accuracy in reporting T2 on trials in which T1 was reported. In c and d, the lines at the top of the graph show T1 accuracy, while the lines at the bottom denote the percent chance for the reported order of T1 and T2 to be inverted. Horizontal axes represent lag, while vertical axes denote accuracy. In the T1+1 blank condition there is no lag-1 case, since that slot is blank. Human data is adapted with permission from Chun and Potter (1995) "A two-stage model for multiple target detection in rapid serial visual presentation." Journal of Experimental Psychology: Human Perception & Performance, 21(1), 109-127, American Psychological Association. In addition, the T2 end of stream data is adapted with permission from Giesbrecht and Di Lollo (1998) "Beyond the attentional blink: visual masking by object substitution." Journal of Experimental Psychology: Human Perception & Performance, 24(5), 1454-1466, American Psychological Association.

Figure 16: Model serial position curves for T1+2 and T2+2 blank compared with the basic and T1+1 blank conditions. The basic blink and T2+2 blank conditions are identical. Heightened T1+2 blank accuracy at lags 1 and 3 arises because at lag-1 the T2 is free from backward masking, while at lag-3 it is free from forward masking.

Figure 17: Activation profiles for 1st stage layers for seen vs missed T2s presented at lag 3 (i.e. during the blink). Profiles are averaged across all T1 times T2 values. The blips in item and task filtered layer activations are caused by the blaster firing (the first time of which is for the T1, before T2 has been presented). It is the small difference in bottom-up trace strength (evident in the Input profiles) that generates the different profiles (and encoding outcomes) at later layers.

Figure 18: Model performance when the T1+1 distracter primes the T2. We also include the model's basic blink and T1+1 blank simulations. We do not include lag 1 or 2 for the T1+1 prime condition. Lag 1 is not a real data point, since the prime occupies this position, while Chua et al (2001) did not consider lag 2, since the proximity of the prime to T2 is a confound.

Figure 19: Data from Neural-ST^2 (on left) and humans (on right), showing blink curves when items are presented (in the usual manner) at an SOA (Stimulus Onset Asynchrony) of 100 ms and when the presentation rate is doubled to an SOA of 50 ms.

Figure 20: Human serial position curves for the letter pairs task. Raw (i.e. non-conditional) T1 and T2 accuracy are shown, along with a measure of overall information extraction, calculated as the mean of T1 and T2 accuracy.

Figure 21: The Gated Auto-associator model.

Figure 22: Processing pathways in the Global Workspace Model. This diagram is adapted from Figure 1B of Dehaene et al (2003) "A neuronal network model linking subjective reports and objective physiological data during conscious perception." Proceedings of the National Academy of Sciences of the USA, 100(14), 8520-8525, Copyright 2003 with permission from National Academy of Sciences, USA.

Figure 23: Battye's Conflict Monitoring model. Grey self-loops indicate a layer with lateral inhibition.

Figure 24: The CODAM model used to simulate the AB. The attentional control signal from the IMC (Inverse Model Controller) to the Object Map is modulated by the Input to Object Map projection. The Monitor to IMC projection is selective, i.e. it excites target items, while inhibiting all other IMC items. This is adapted from

Figure 1 of Fragopanagos et al (2005) "A Neurodynamic Model of the Attentional Blink." Cognitve Brain Research, 24(3), 568-586. Copyright 2005 with permission from Elsevier.

Figure 25: This depiction of the LC model is adapted with permission from figure 2 of Nieuwenhuis et al. (2006) "The Role of the Locus Coeruleus in Mediating the Attentional Blink: A Neurocomputational Theory." Journal of Experimental Psychology: General (in Press), American Psychological Association. Note, inhibitory and crosstalk connections between T1 and D are not shown for simplicity of presentation and point size of arrows indicates weight strength.

Figure 26: (A) Connection types. (B) Connectivity into and out of the binding gates; connectivity shown for one neuron is repeated across that layer.

Figure 27: Full blaster implementation. Output connection thresholds emanating from the Blaster Output (ON) node are denoted 1, 2 and 3, with 1 less than 2 less than 3. Note, connections (a), (b) and (c) are as per Figure 9.

[i] Two such examples are repetition blindness, which obtains when two instances of the same item arise in close temporal proximity (Kanwisher, 1987) and lag 1 sparing, which shows that when items are presented in close temporal proximity, order is not easily retrievable (see section 3.2 and the Lag 1 Sparing subsection of section 3.3).

[ii] For example, spatial overlap of input stimuli, as, e.g., arises in AB experiments, can generate interference at the earliest levels of stage 1; see the Early Visual Processing subsection of section 4.3.

[iii] The capacity of WM is a lower bound on the number of duplicates, since it classically measures the number of *distinct* items that can be stored, however, the number of repetitions of *the same* item, is likely to be a good deal bigger.

[iv] Although, masking in AB experiments is weaker than that arising in classic psychophysics masking experiments (Keysers & Perrett, 2002). In particular, in AB tasks items are presented for around 100ms, which is a good deal longer than in classic masking studies.

[v] That is, even if both targets are bound to the other token, then this ambiguity is resolved by the certainty of the binding to the unambiguously bound token.

[vi] Importantly, this inhibition is only active during a binding allocation and is not active while an existing allocation is being maintained. This effect is obtained because, as clarified in section 4.6, each binding unit is in fact a neural circuit, one part of which is only active during an allocation period; it is this part (the gate node) of each binding unit that projects inhibition to the blaster.

[vii] Although, there is evidence that the unavailability of the TAE may not be absolute. In particular, placing a cue before the T2 can attenuate the blink, suggesting that the cue fires the blaster even while T1 tokenization is ongoing (Nieuwenstein, Chun, van der Lubbe, & Hooge, 2005).

[viii] To pull the 50 ms model data above floor, we have increased the input strength compared to the 100 ms SOA conditions. This is not surprising, since it is consistent with the existence of a non-linear relationship between bottom-up trace strength and length of item presentation / SOA.

[ix] Indeed, the experimental findings in Nieuwenhuis et al. (2006) give added weight to our confirmation of prediction 1.

[x] Although, it should be emphasized that this is a weakness that Nieuwenhuis et al. (2006) acknowledge, and informally argue could be resolved in a revision of their current implementation.

[xi] In fact, it should be noted that none of the prominent informal theories of the blink explicitly consider serial order. So, the interference theory should not be overly criticised in this respect.

[xii] In this respect, it is also interesting to note that the sort of crossover effects that Jolicoeur (1998) report when comparing speeded and unspeeded T1 tasks, have also been identified with ST^2 when the ease of the T1 task is varied (Wyble & Bowman, 2005).

[xiii] These findings resonate with studies showing that discriminating the duration of an event in a (singly perceived) RSVP stream does not generate a blink (Sheppard, Duncan, Shapiro, & Hillstrom, 2002). Presumably, gap judgements do not initiate the generation of new object representations.

[xiv] Although note, Di Lollo et al (2005) and Olivers et al (2005) failed to obtain separate glimpses when they considered sequences of three or more targets in an RSVP task; see section 7.2.

[xv] Note this projection is depicted as inhibitory in Figure 9. But, in fact, it is realized as an excitatory projection onto the Blaster Input off node, which overall yields an inhibitory effect of the binding pool onto the Blaster Input on node.

[xvi] psychophysicstoolbox.org.

[xvii] These letters were chosen by scanning a corpus of over 150 assorted English novels including contemporary as well as classical authors and identifying the set of four letter combinations that were the least likely to appear asymmetrically, either in order, or in pairing. While asymmetries may have existed in other languages or via popular acronyms, targets were always randomly chosen from this set so there could not have been a systematic bias.

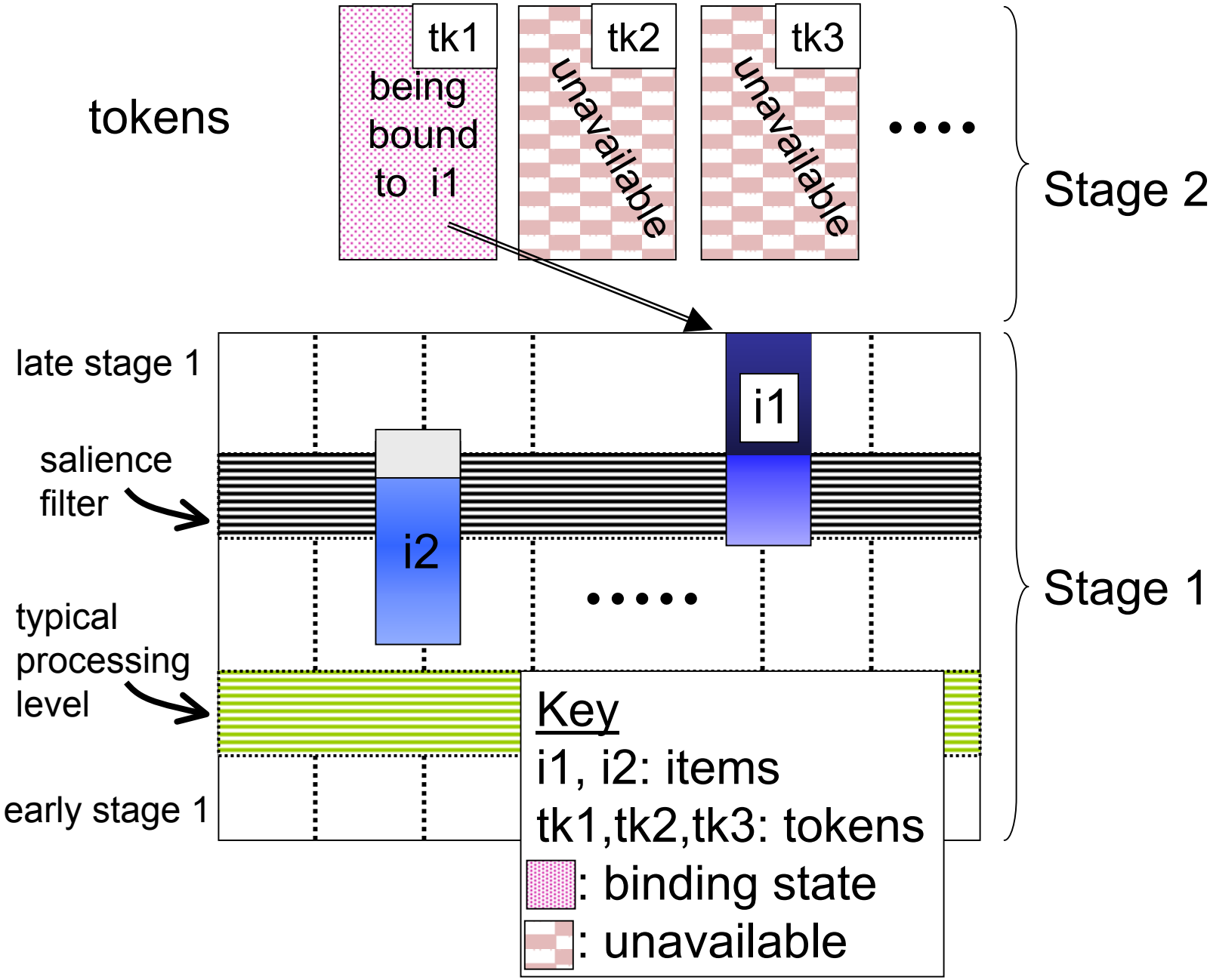

tokens
tk1
being
bound
to i1
tk2
unavailable
tk3
unavailable
Stage 2
late stage 1
salience
filter
typical
processing
level
early stage 1
i1
i2
Stage 1
Key
i1, i2: items
tk1,tk2,tk3: tokens
: binding state
: unavailable

FIGURE 1

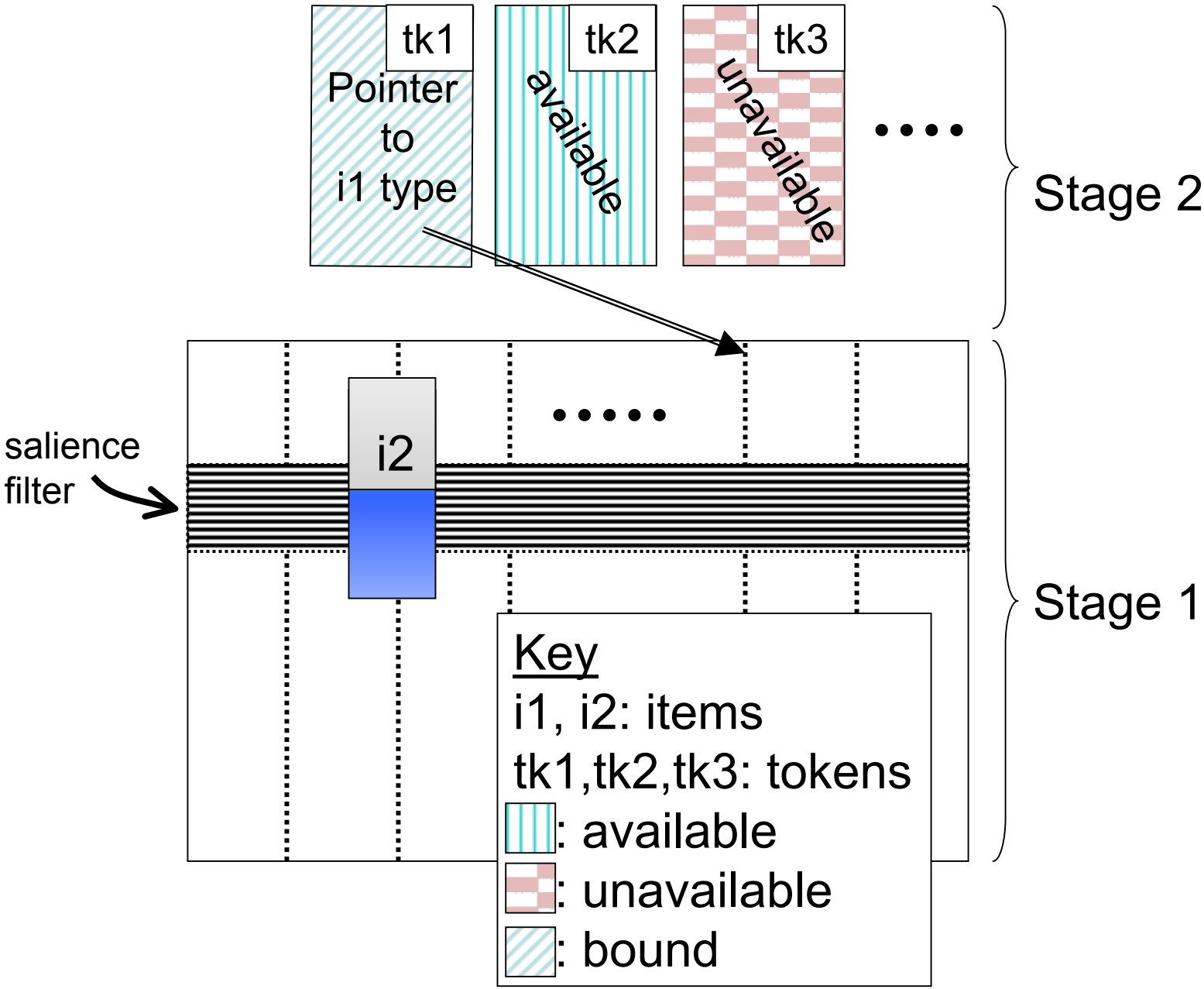

tk1
Pointer
to
i1 type
tk2
available
tk3
unavailable
Stage 2
salience
filter
i2
Stage 1
Key
i1, i2: items
tk1,tk2,tk3: tokens
: available
: unavailable
: bound

FIGURE 2

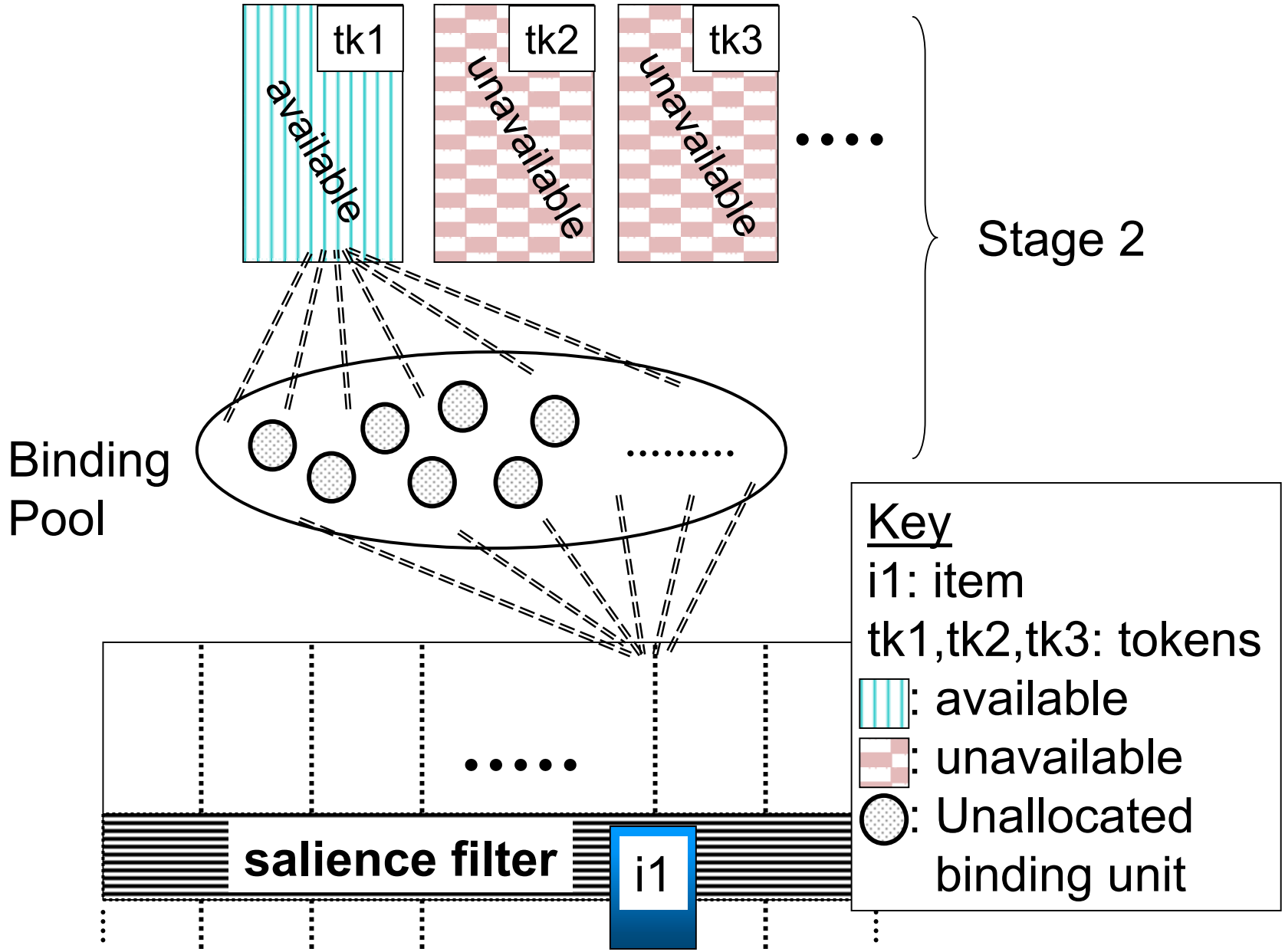

tk1
tk2
tk3
available
unavailable
unavailable
Stage 2
Binding
Pool
salience filter
i1
Key
i1: item
tk1,tk2,tk3: tokens
: available
: unavailable
: Unallocated
binding unit

FIGURE 3

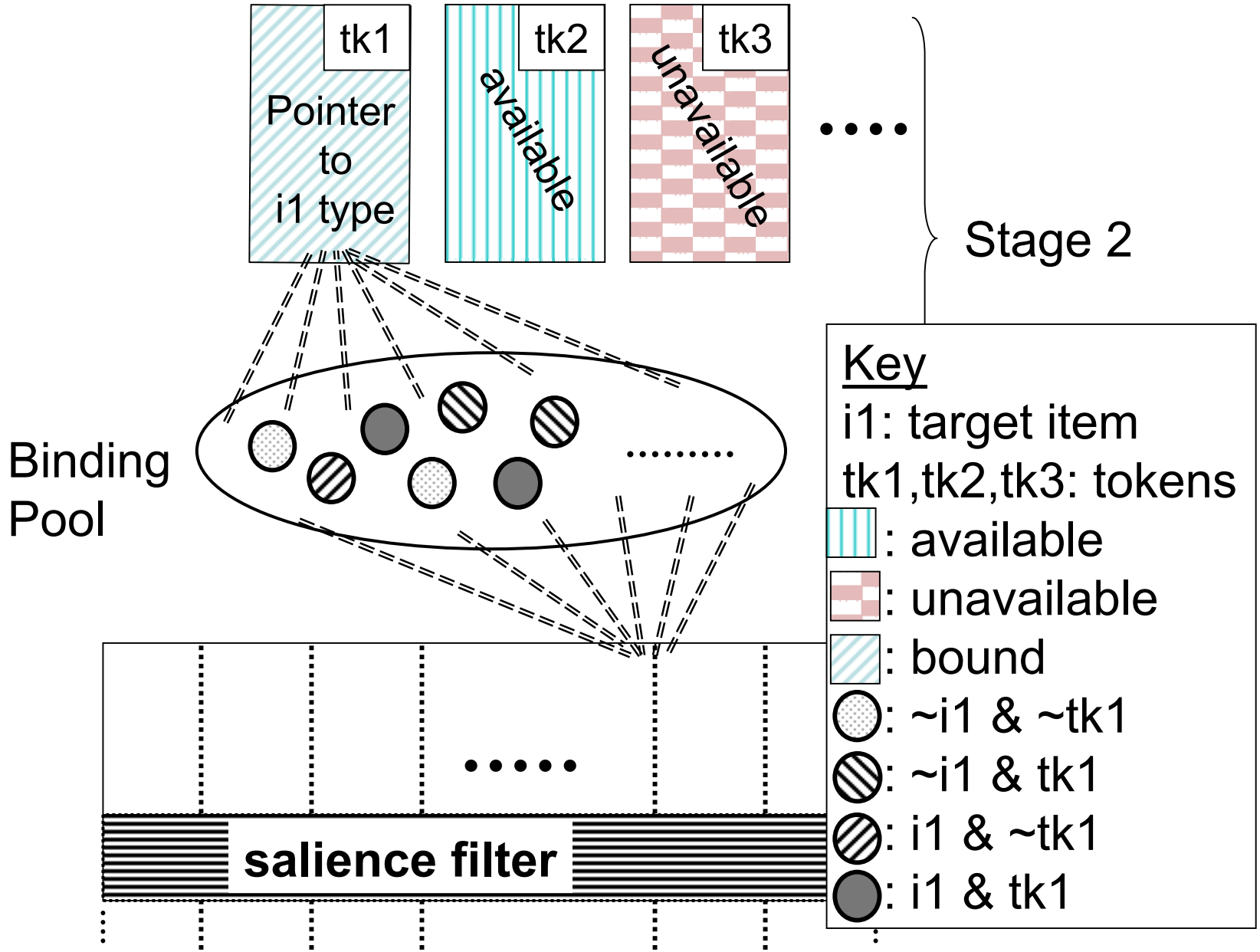

tk1
Pointer
to
i1 type
tk2
available
tk3
unavailable
Stage 2
Binding
Pool
salience filter
Key
i1: target item
tk1,tk2,tk3: tokens
: available
: unavailable
: bound
: ~i1 & ~tk1
: ~i1 & tk1
: i1 & ~tk1
: i1 & tk1

FIGURE 4

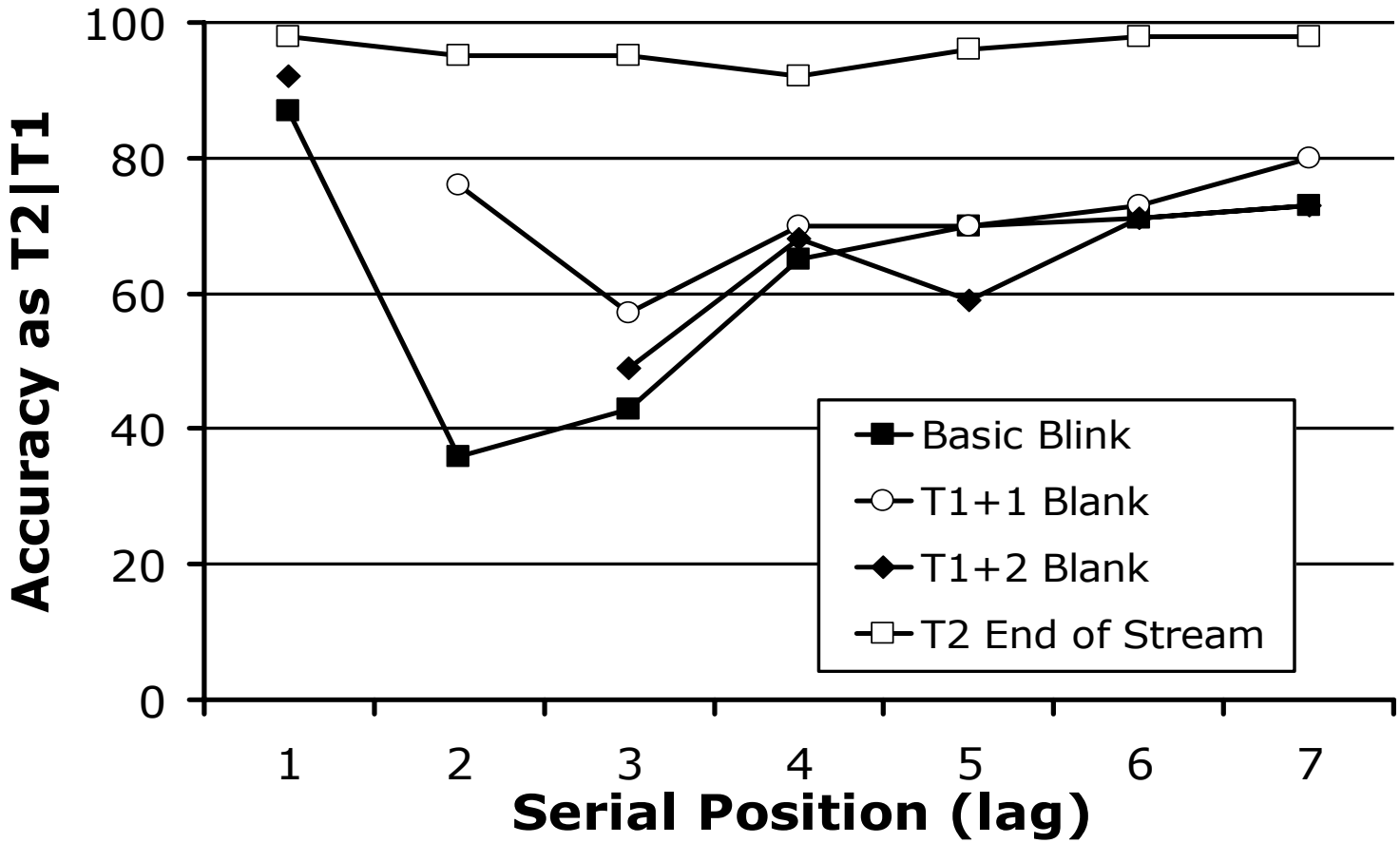

Accuracy as T2|T1
100
80
60
40
20
0
1
2
3
4
5
6
7
Serial Position (lag)
Basic Blink
T1+1 Blank
T1+2 Blank
T2 End of Stream

FIGURE 5

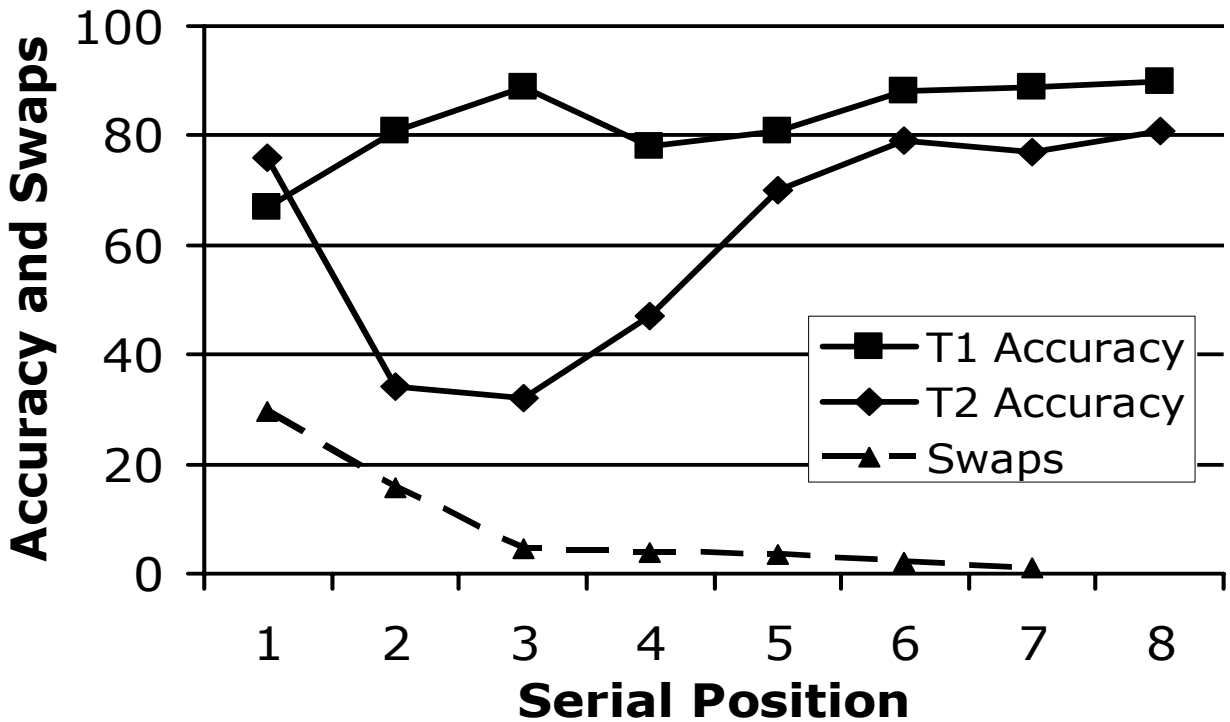

Accuracy and Swaps
100
80
60
40
20
0
1
2
3
4
5
6
7
8
Serial Position
T1 Accuracy
T2 Accuracy
Swaps

FIGURE 6

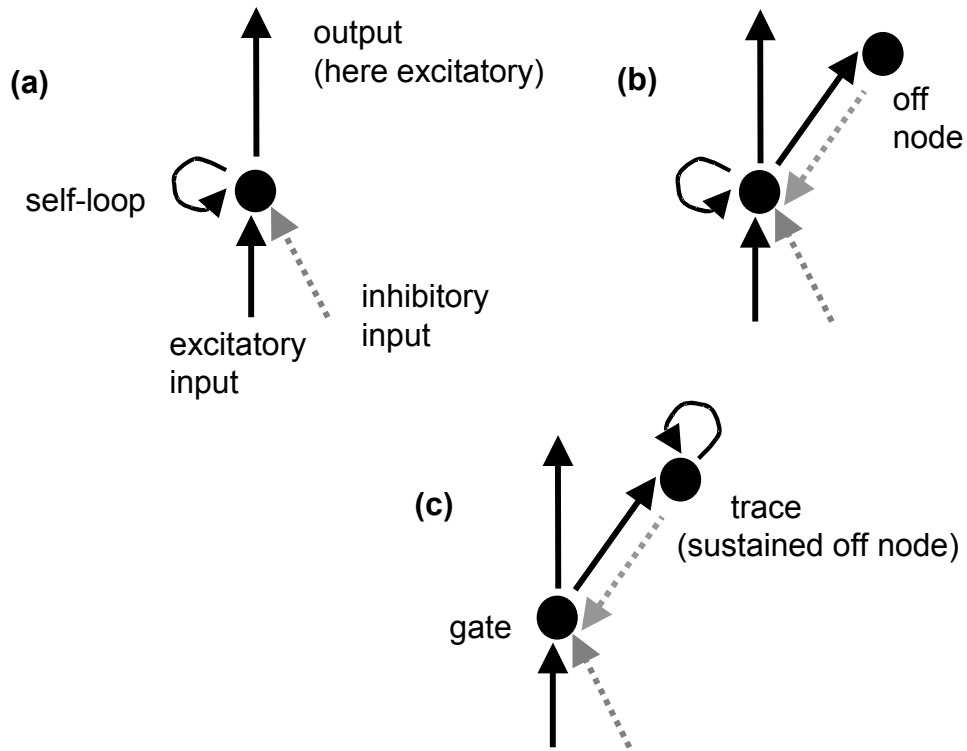

(a)
output
(here excitatory)
self-loop
inhibitory
input
excitatory
input
(b)
off
node
(c)
trace
(sustained off node)
gate

FIGURE 7

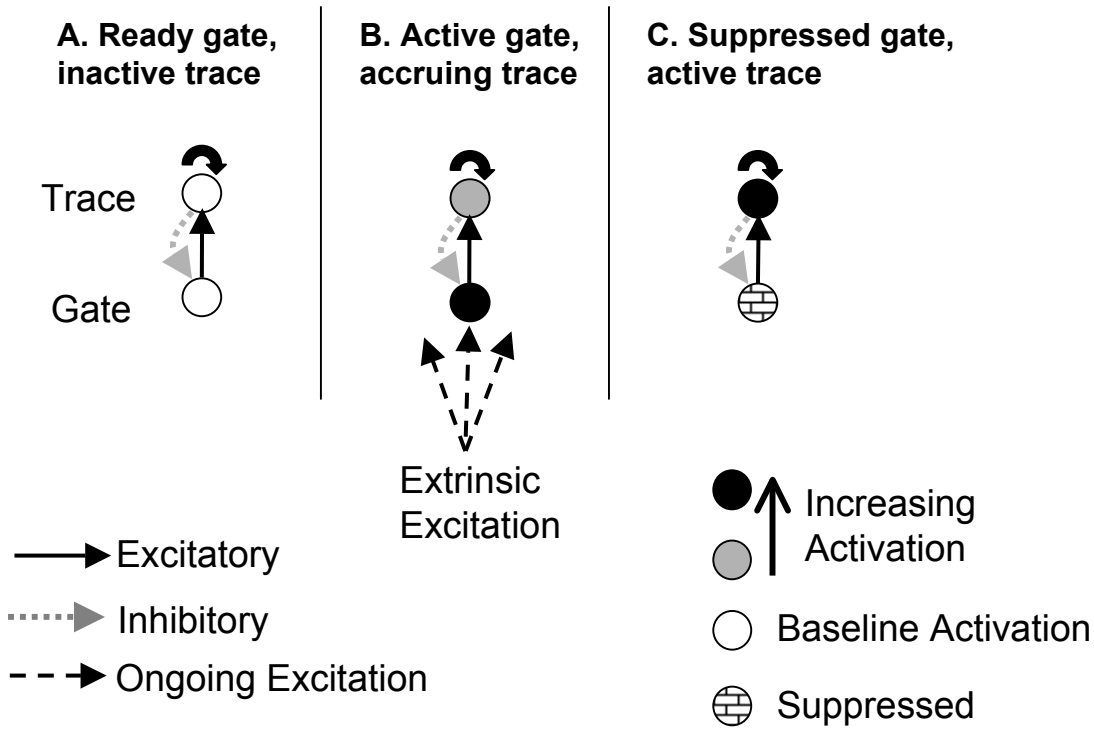

A. Ready gate, inactive trace
B. Active gate, accruing trace
C. Suppressed gate, active trace
Trace
Gate
Extrinsic Excitation
Excitatory
Inhibitory
Ongoing Excitation
Increasing Activation
Baseline Activation
Suppressed

FIGURE 8

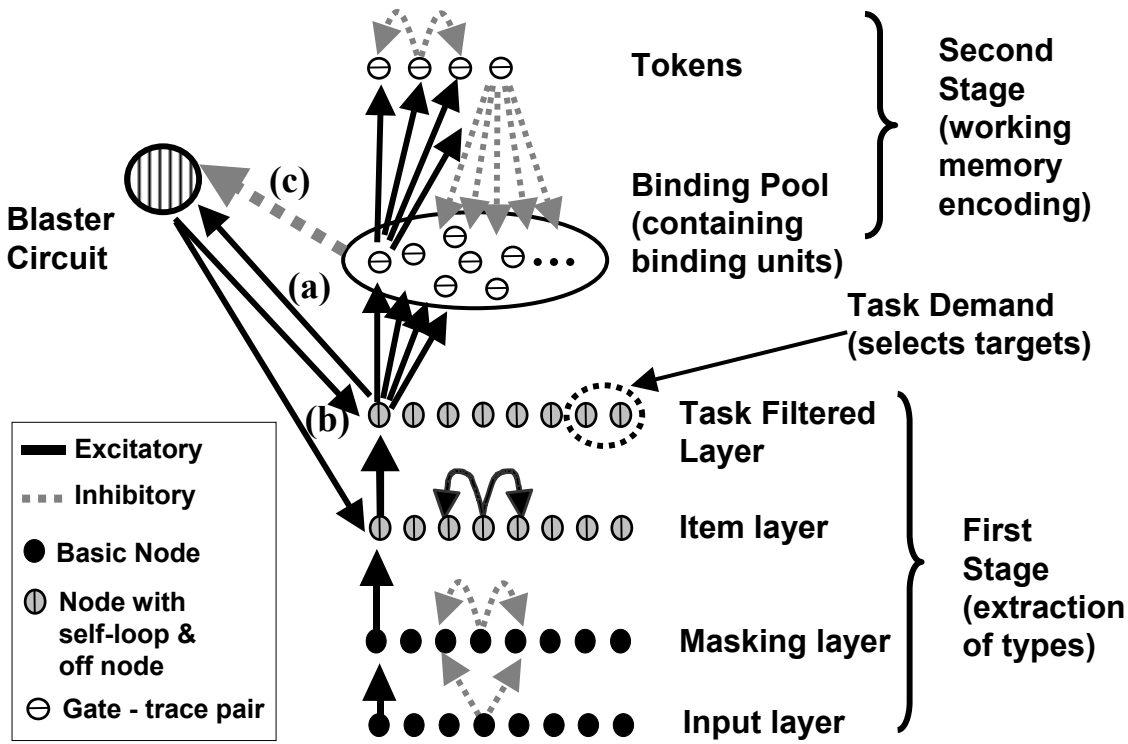

Tokens
Second Stage (working memory encoding)
Binding Pool (containing binding units)
Blaster Circuit
(c)
(a)
(b)
Task Demand (selects targets)
Task Filtered Layer
Item layer
Masking layer
Input layer
First Stage (extraction of types)
Excitatory
Inhibitory
Basic Node
Node with self-loop & off node
Gate - trace pair

FIGURE 9

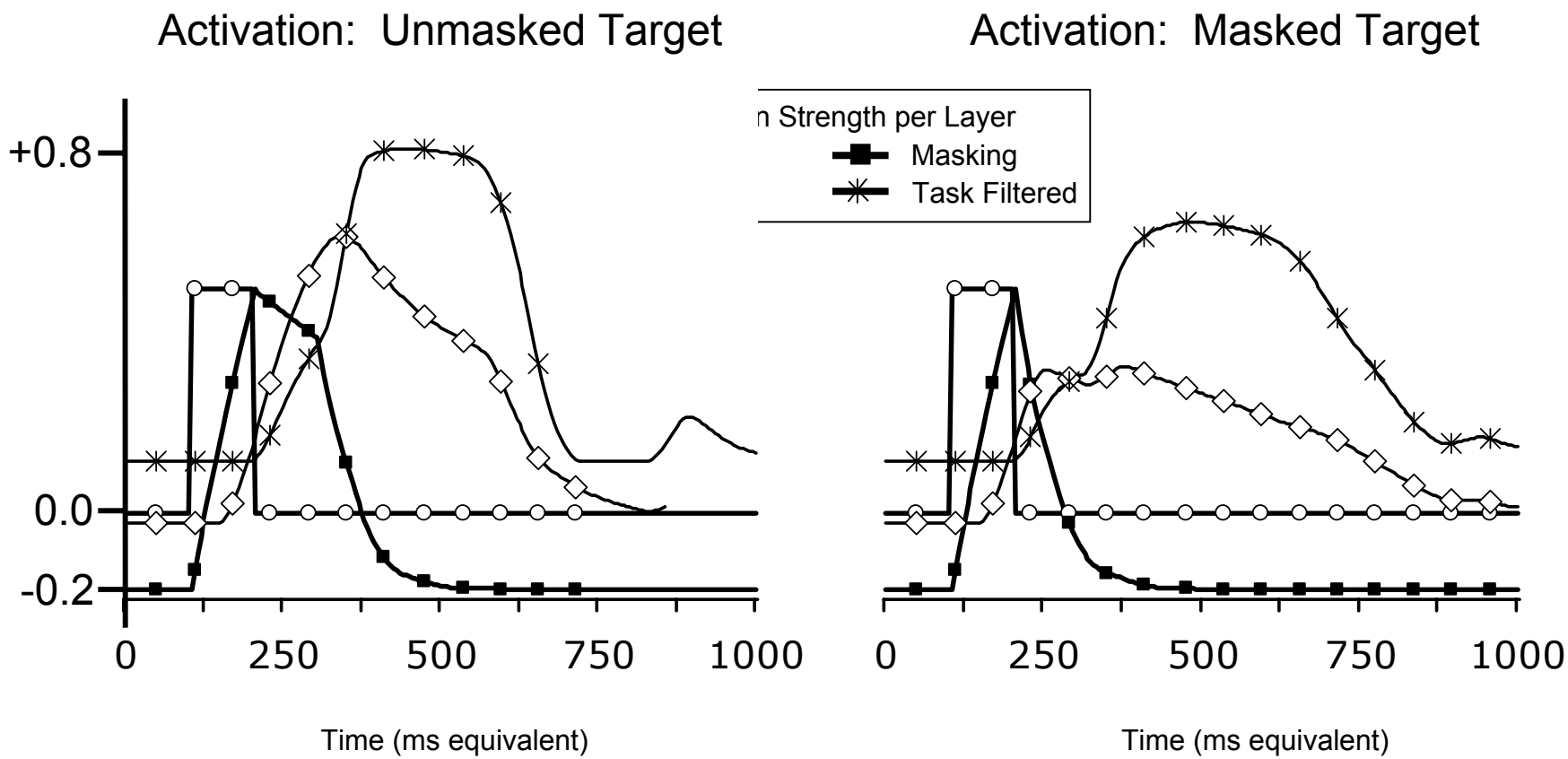

Activation: Unmasked Target
Activation: Masked Target
n Strength per Layer
Masking
Task Filtered
+0.8
0.0
-0.2
0
250
500
750
1000
Time (ms equivalent)
Time (ms equivalent)

FIGURE 10

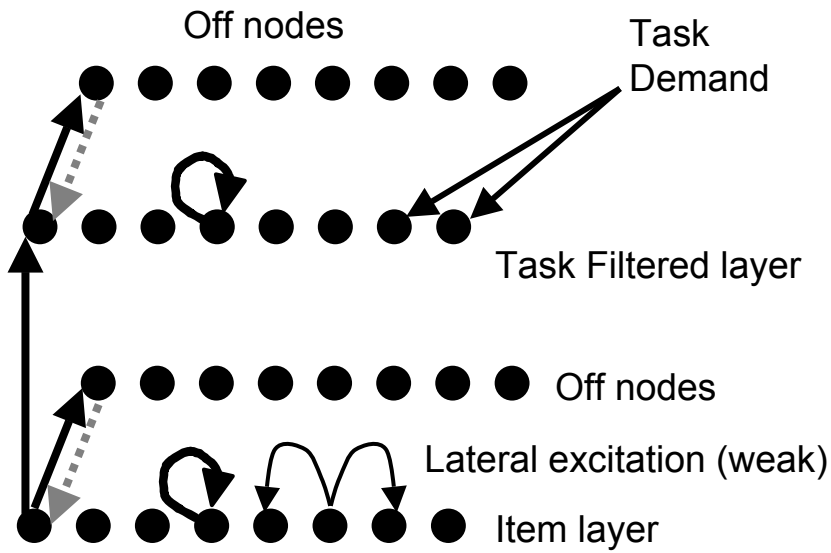

Off nodes
Task
Demand
Task Filtered layer
Off nodes
Lateral excitation (weak)
Item layer

FIGURE 11

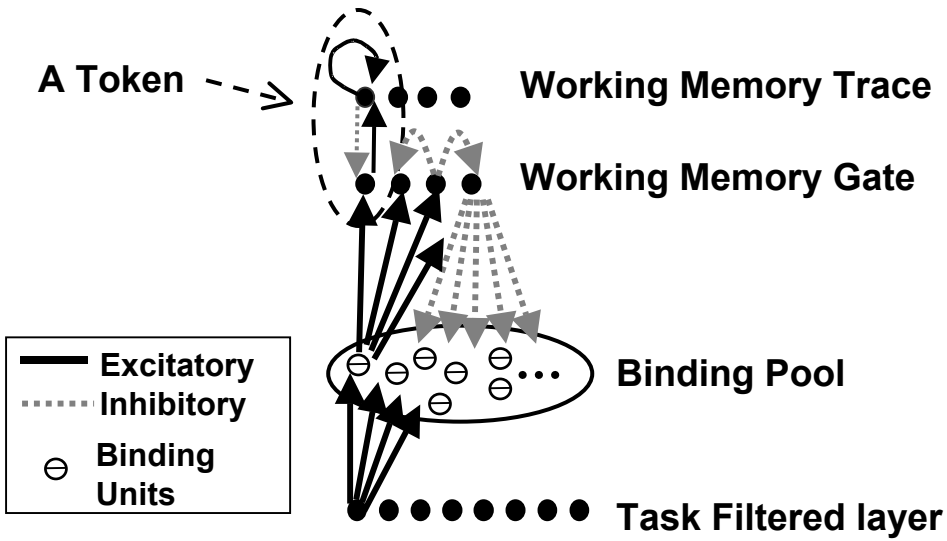

A Token
Working Memory Trace
Working Memory Gate
Excitatory
Inhibitory
Binding Units
Binding Pool
Task Filtered layer

FIGURE 12

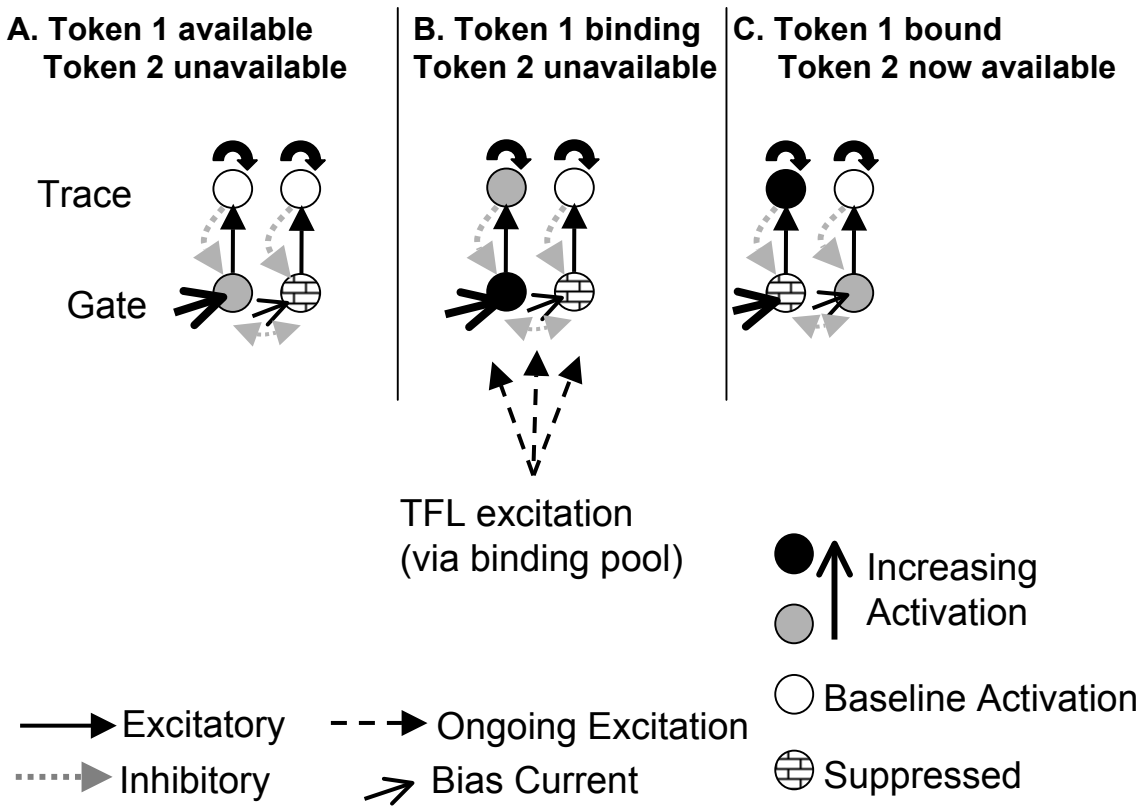

A. Token 1 available
Token 2 unavailable
B. Token 1 binding
Token 2 unavailable
C. Token 1 bound
Token 2 now available
Trace
Gate
TFL excitation
(via binding pool)
Increasing
Activation
Baseline Activation
Suppressed
Excitatory
Inhibitory
Ongoing Excitation
Bias Current

FIGURE 13

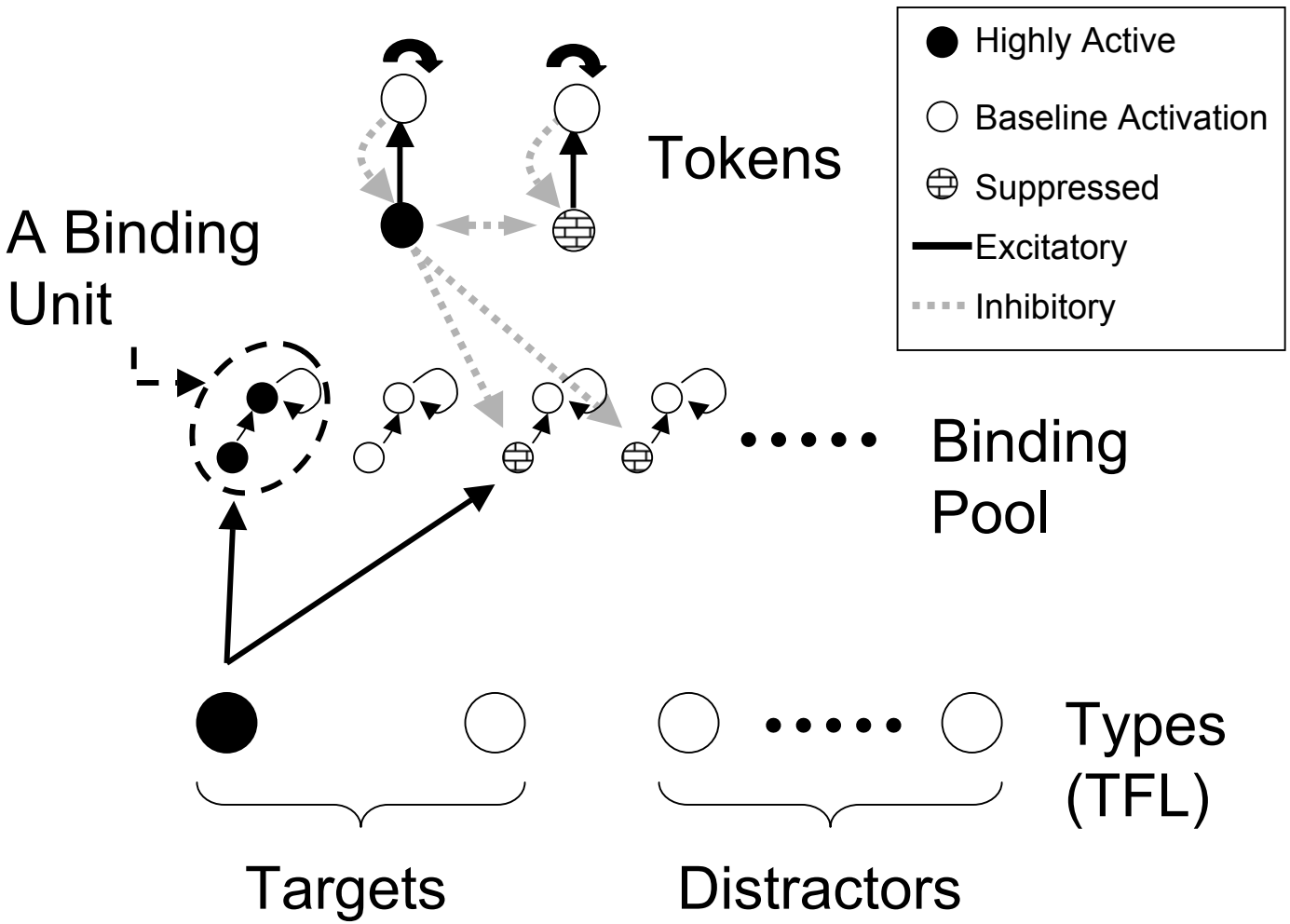

Highly Active
Baseline Activation
Suppressed
Excitatory
Inhibitory
Tokens
A Binding Unit
Binding Pool
Types (TFL)
Targets
Distractors

FIGURE 14

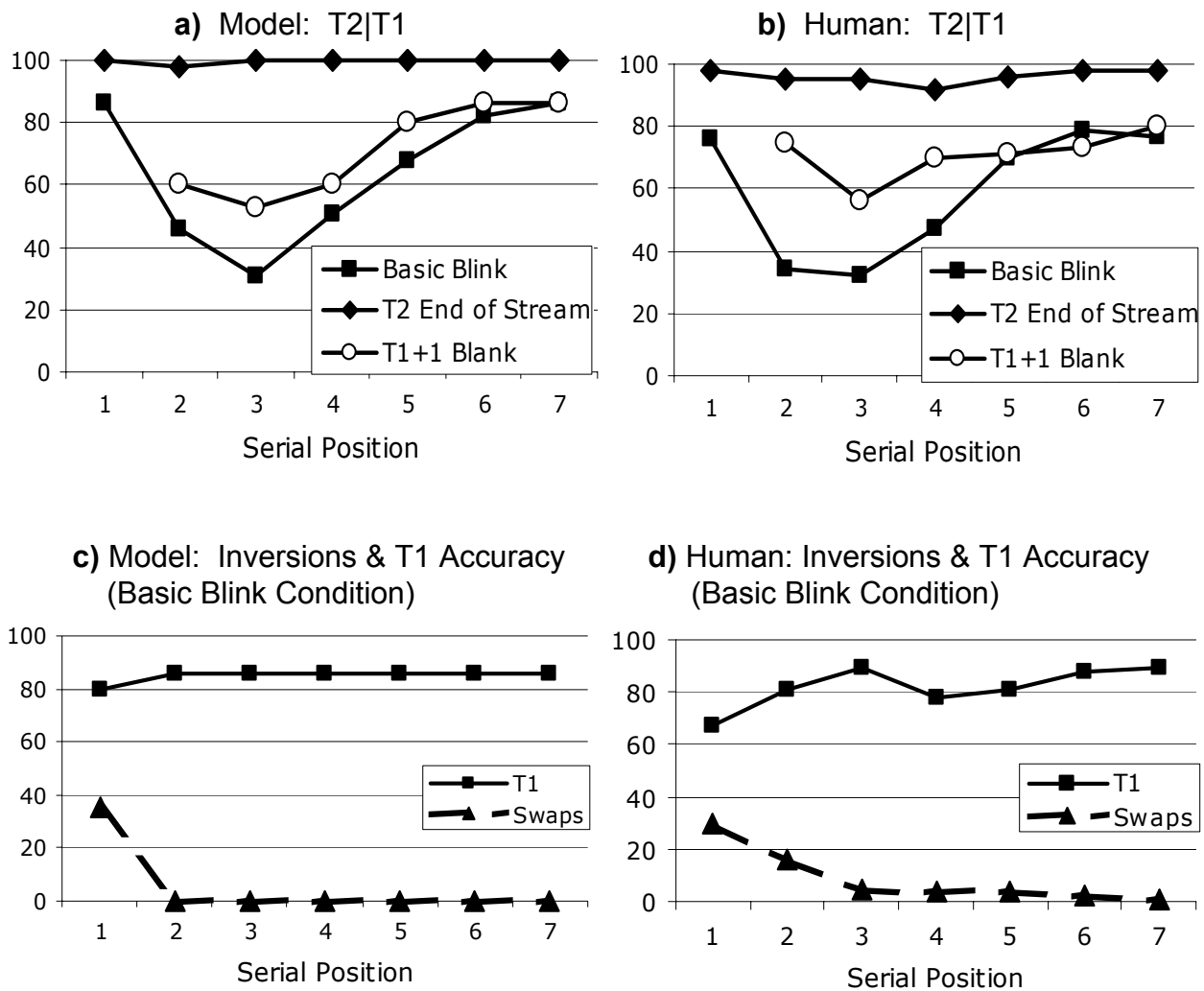

a) Model: T2|T1
b) Human: T2|T1
c) Model: Inversions & T1 Accuracy (Basic Blink Condition)
d) Human: Inversions & T1 Accuracy (Basic Blink Condition)
Basic Blink
T2 End of Stream
T1+1 Blank
T1
Swaps
Serial Position
0
20
40
60
80
100
1
2
3
4
5
6
7

FIGURE 15

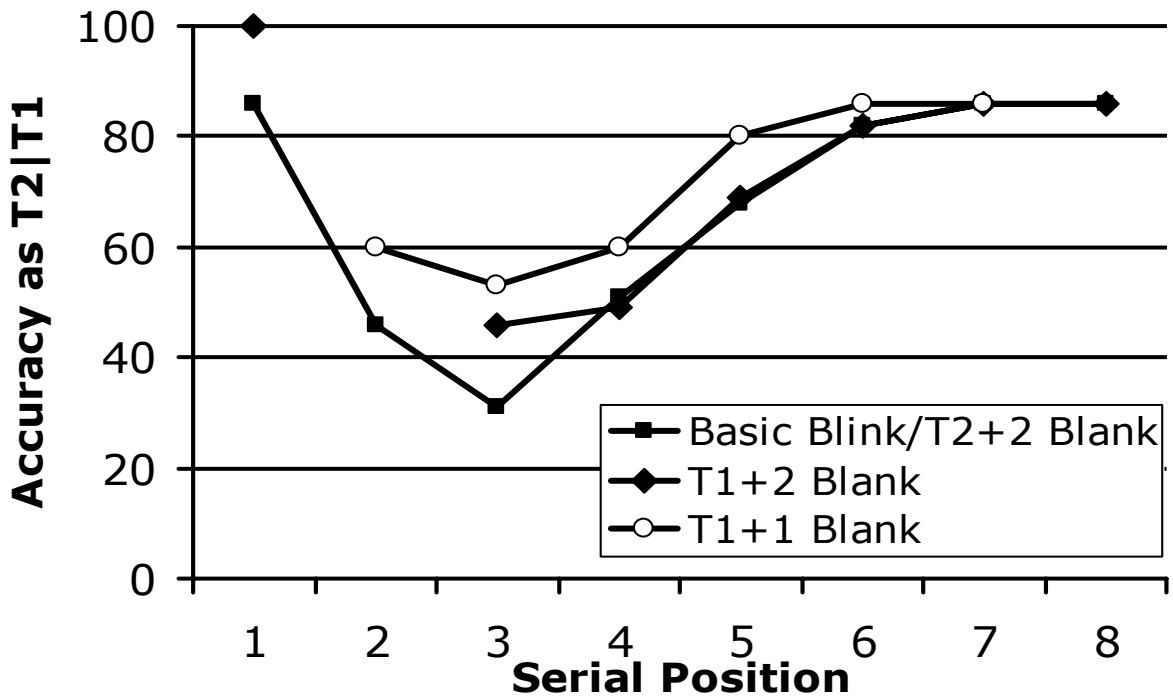
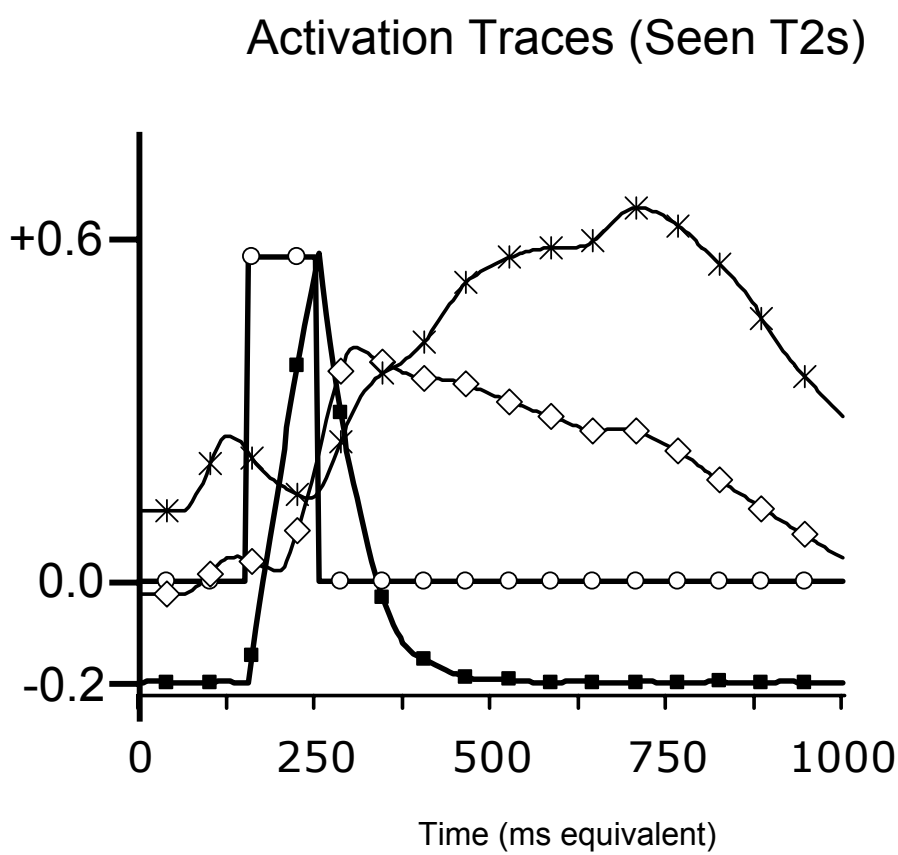

100
80
60
40
20
0
Accuracy as T2|T1
1
2
3
4
5
6
7
8
Serial Position
Basic Blink/T2+2 Blank
T1+2 Blank
T1+1 Blank

FIGURE 16


Activation Traces (Seen T2s)
+0.6
0.0
-0.2
0
250
500
750
1000
Time (ms equivalent)

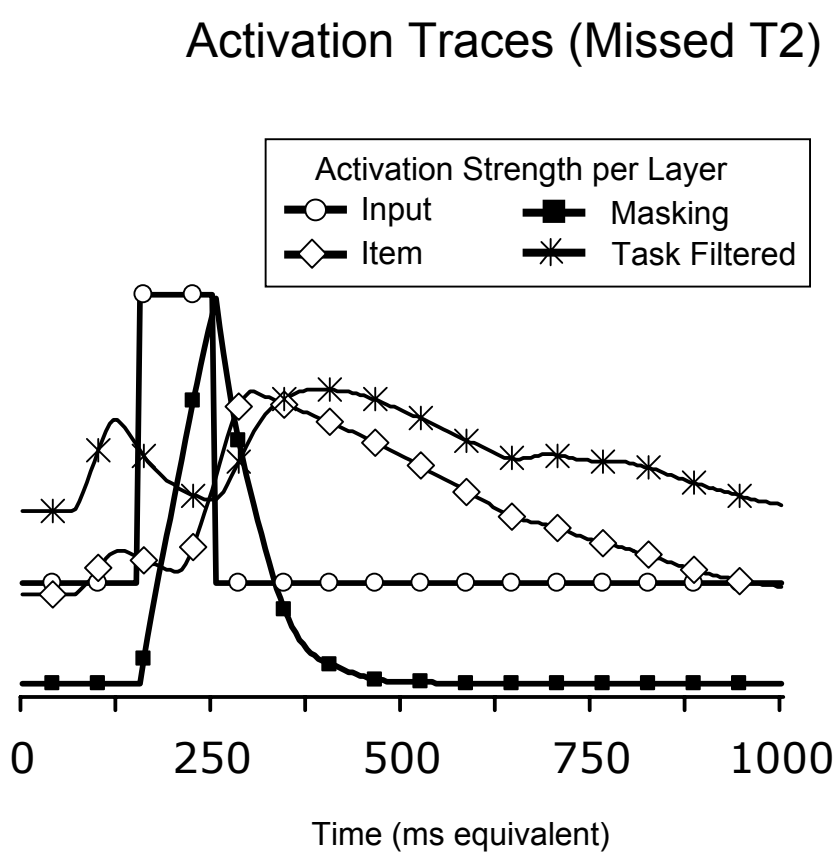

Activation Traces (Missed T2)
Activation Strength per Layer
Input
Masking
Item
Task Filtered
0
250
500
750
1000
Time (ms equivalent)

FIGURE 17

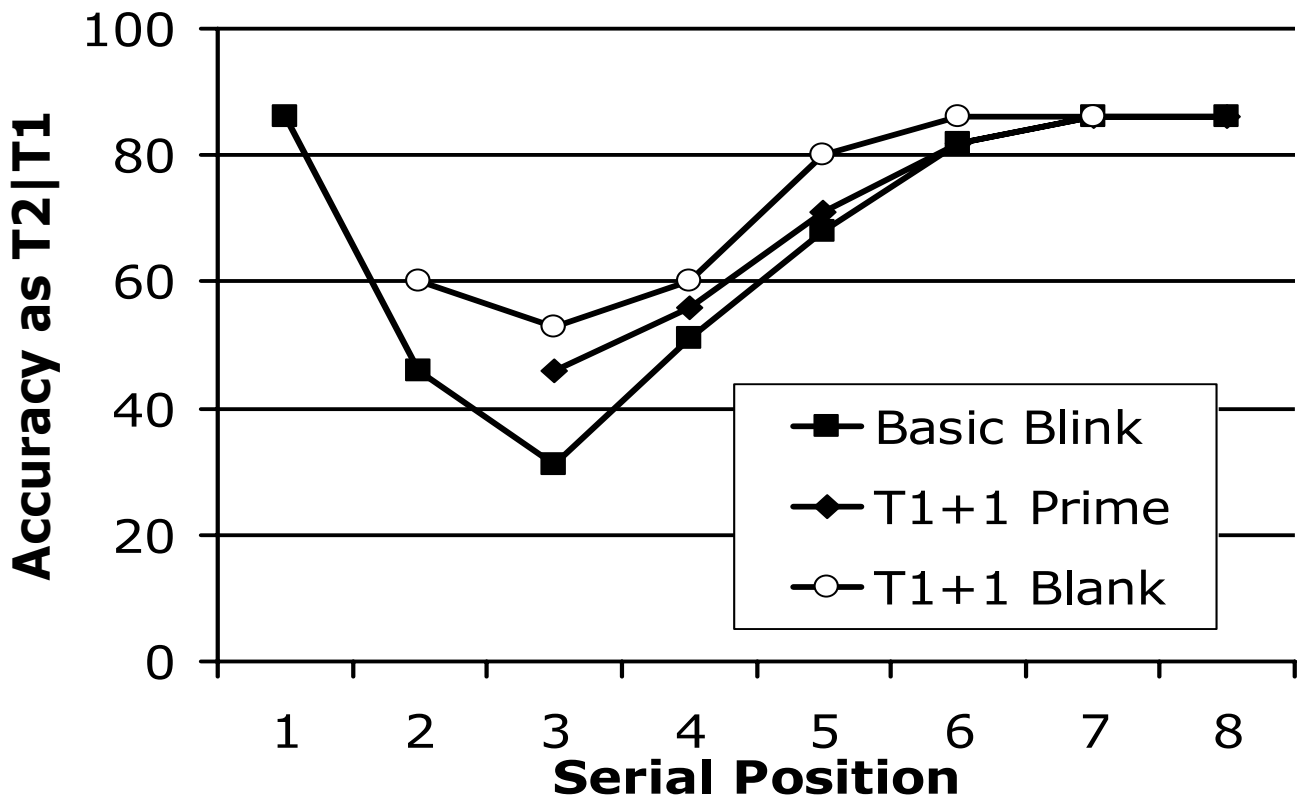

100
80
60
40
20
0
Accuracy as T2|T1
1
2
3
4
5
6
7
8
Serial Position
Basic Blink
T1+1 Prime
T1+1 Blank

FIGURE 18

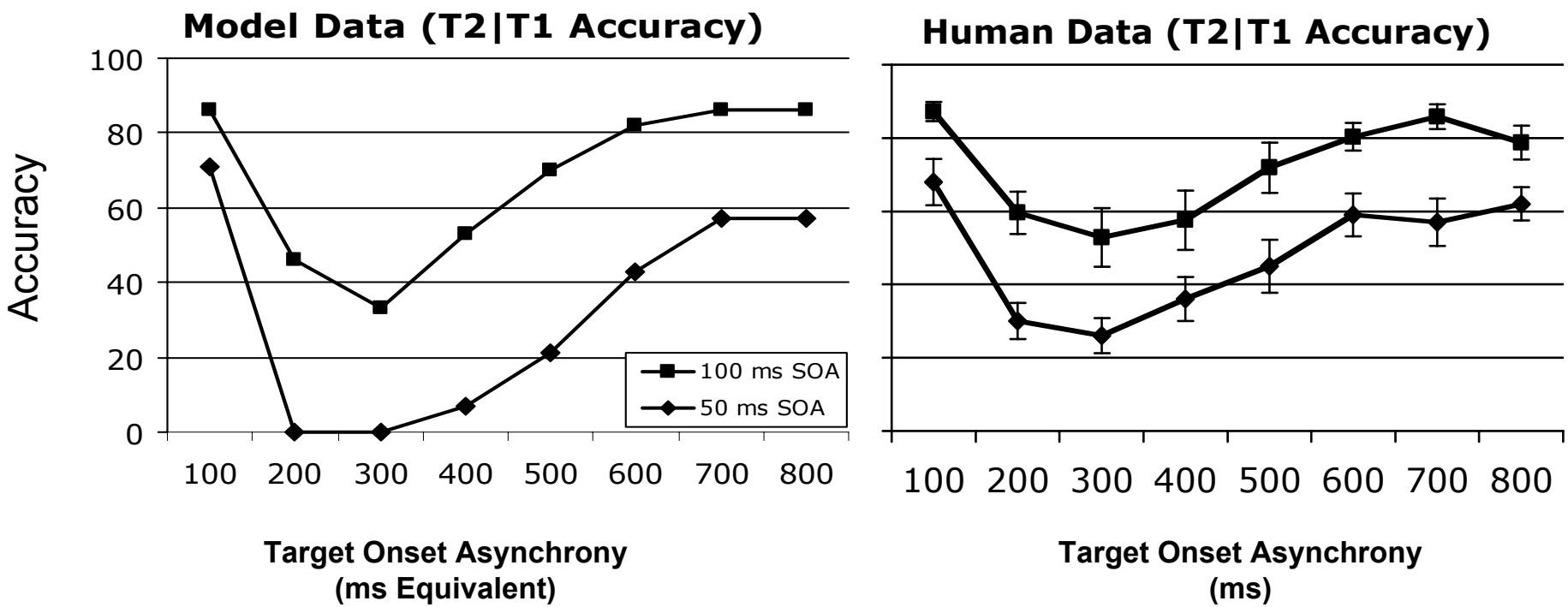

Model Data (T2|T1 Accuracy)
Accuracy
100
80
60
40
20
0
100 ms SOA
50 ms SOA
100 200 300 400 500 600 700 800
Target Onset Asynchrony
(ms Equivalent)
Human Data (T2|T1 Accuracy)
100 200 300 400 500 600 700 800
Target Onset Asynchrony
(ms)

FIGURE 19

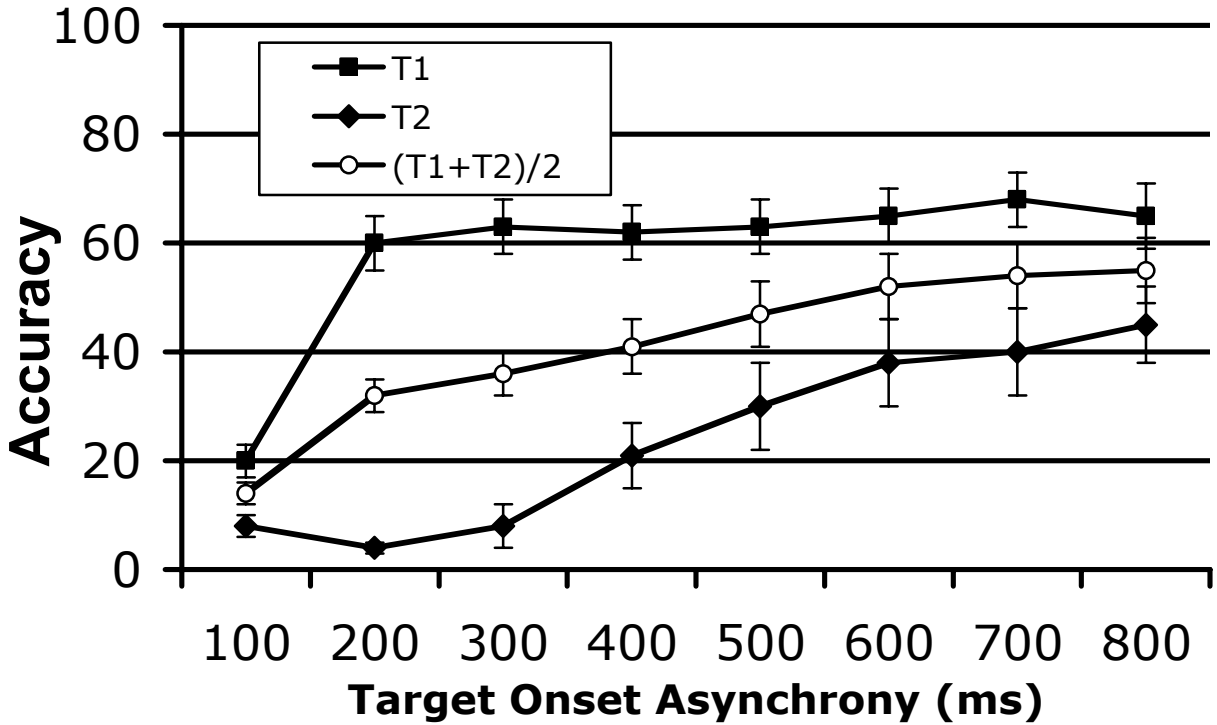

100
80
60
40
20
0
Accuracy
T1
T2
(T1+T2)/2
100
200
300
400
500
600
700
800
Target Onset Asynchrony (ms)

FIGURE 20

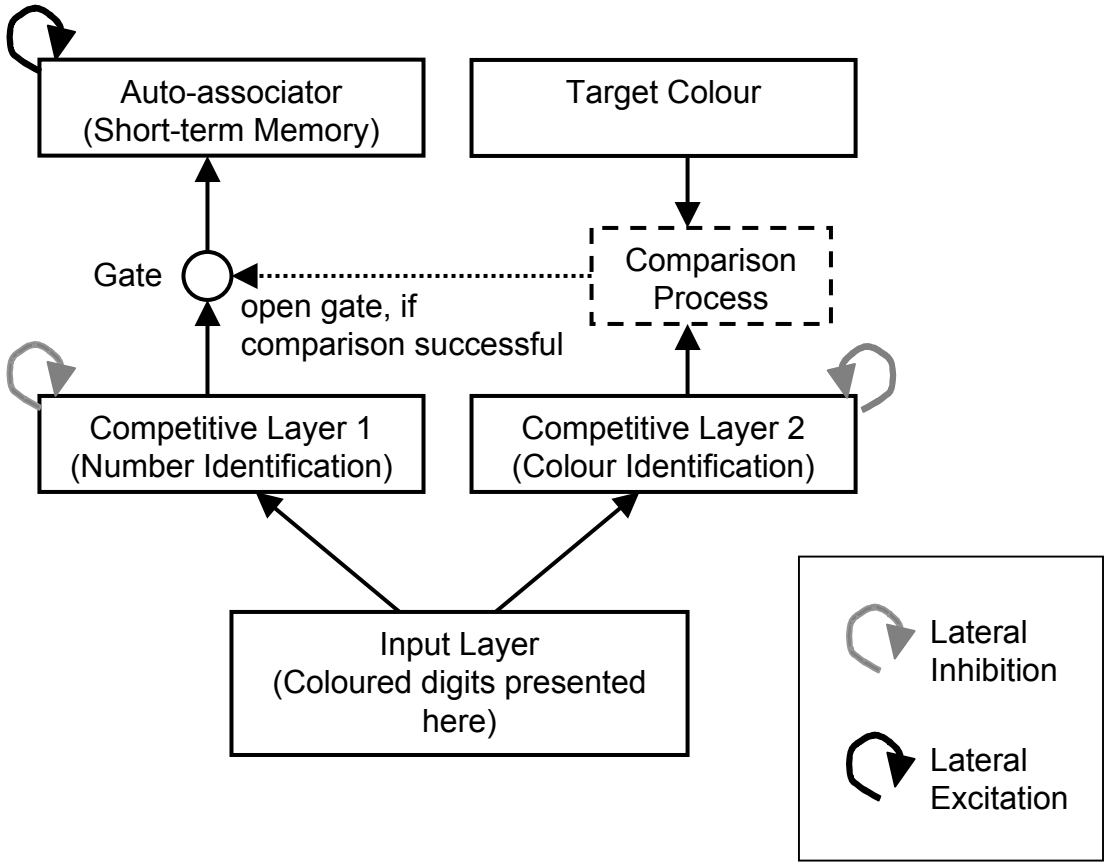

Auto-associator
(Short-term Memory)
Target Colour
Gate
open gate, if
comparison successful
Comparison
Process
Competitive Layer 1
(Number Identification)
Competitive Layer 2
(Colour Identification)
Input Layer
(Coloured digits presented
here)
Lateral
Inhibition
Lateral
Excitation

FIGURE 21

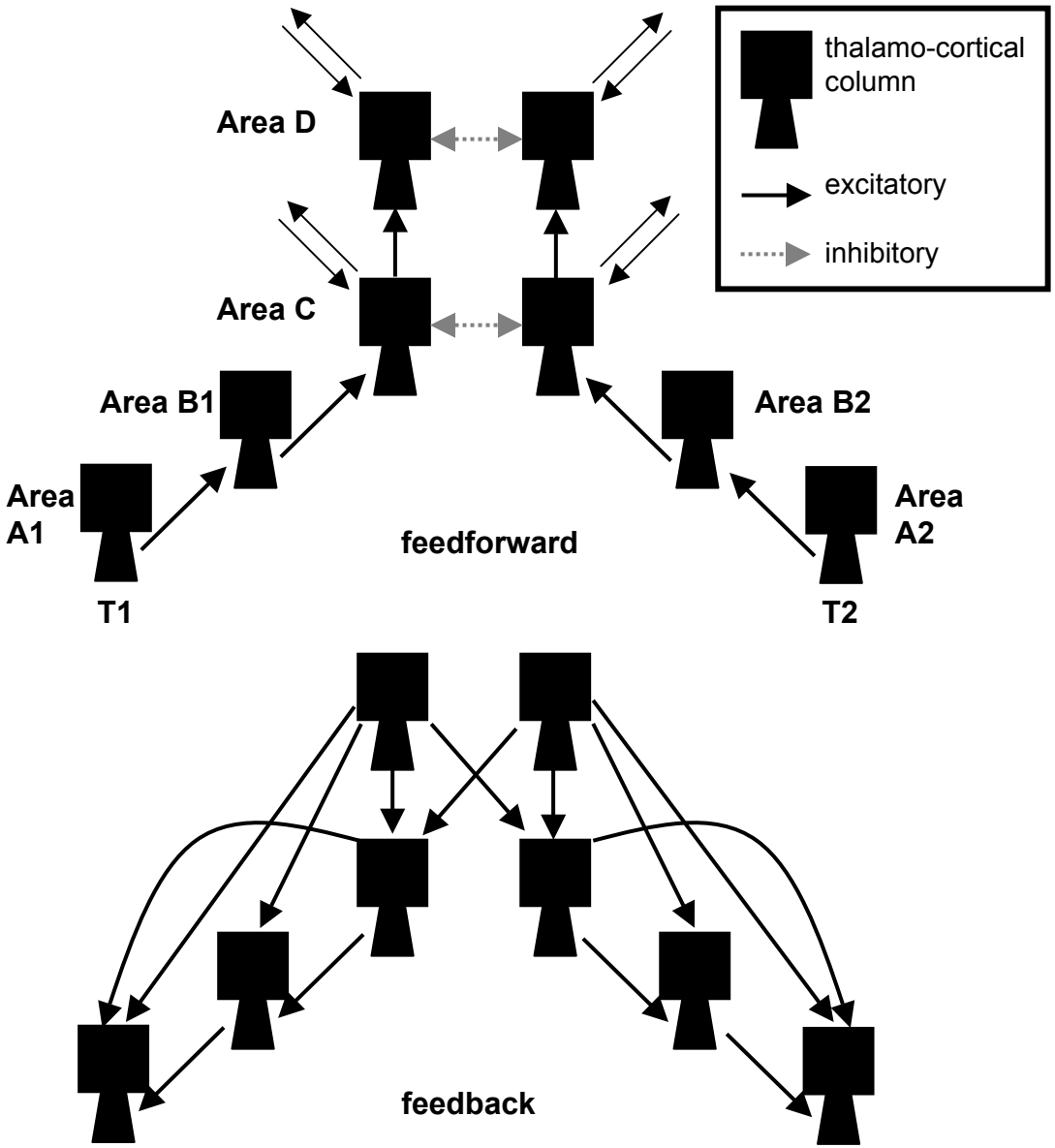

thalamo-cortical column
excitatory
inhibitory
Area D
Area C
Area B1
Area B2
Area A1
Area A2
T1
T2
feedforward
feedback

FIGURE 22

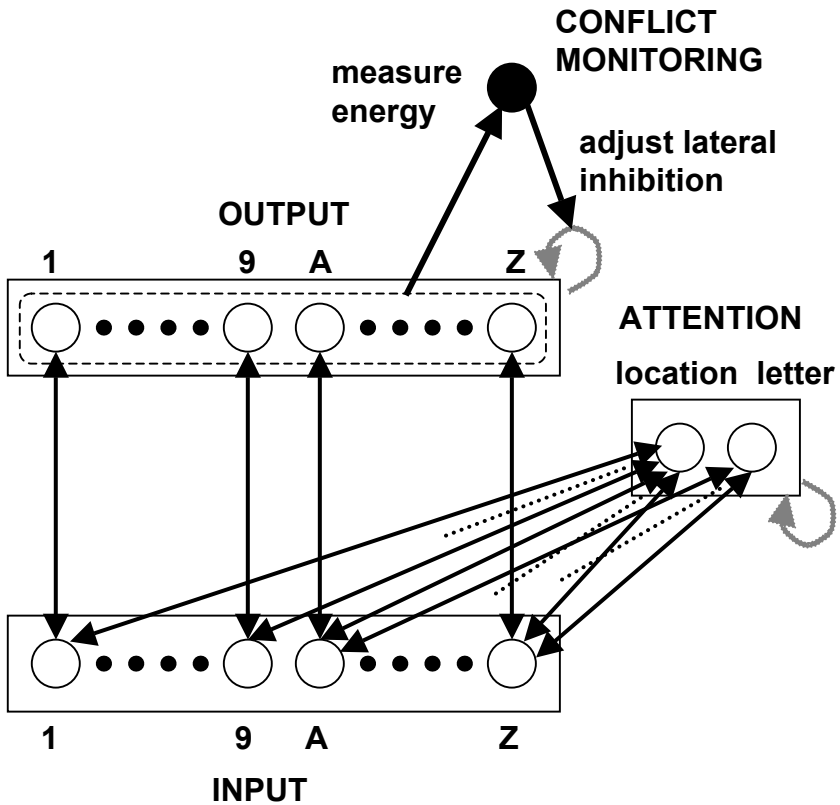

CONFLICT
MONITORING
measure
energy
adjust lateral
inhibition
OUTPUT
1
9
A
Z
ATTENTION
location
letter
1
9
A
Z
INPUT

FIGURE 23

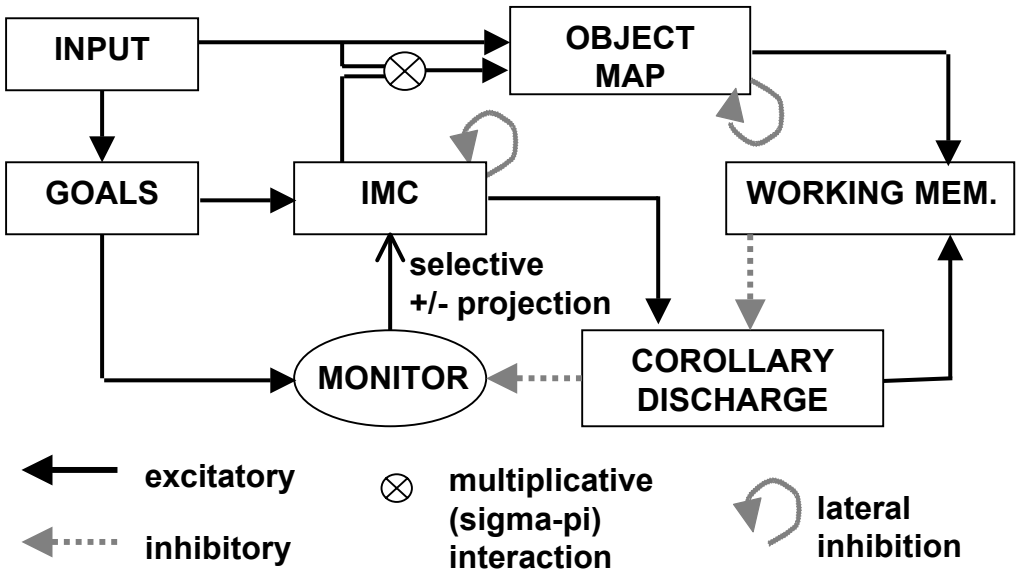

INPUT
OBJECT MAP
GOALS
IMC
WORKING MEM.
selective +/- projection
MONITOR
COROLLARY DISCHARGE
excitatory
inhibitory
multiplicative (sigma-pi) interaction
lateral inhibition

FIGURE 24

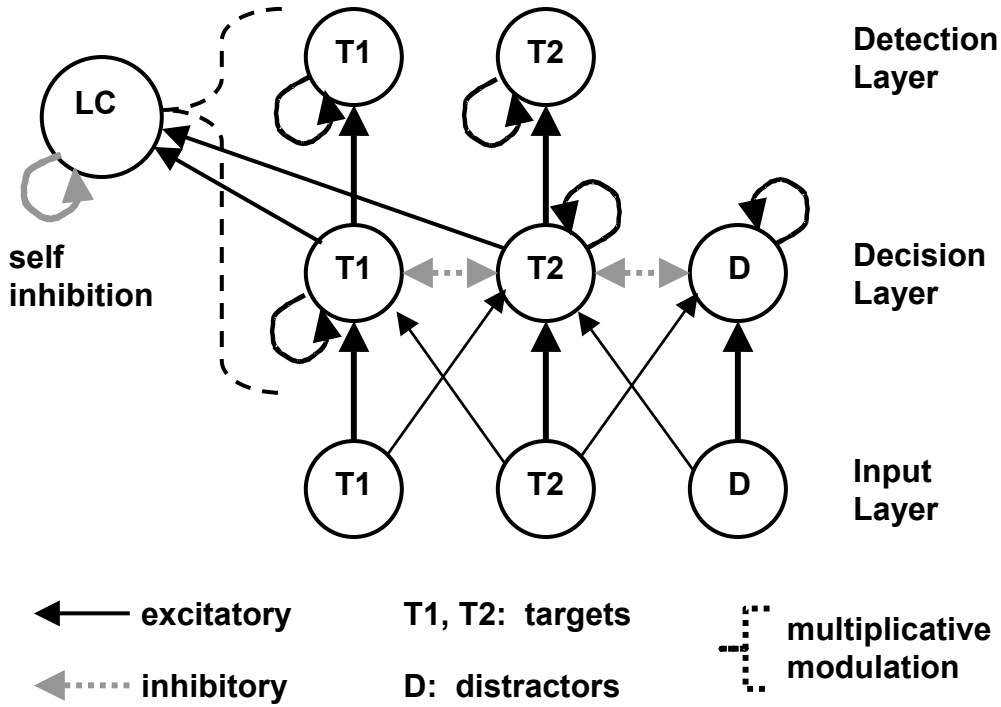

T1
T2
LC
Detection
Layer
self
inhibition
T1
T2
D
Decision
Layer
T1
T2
D
Input
Layer
excitatory
inhibitory
T1, T2: targets
D: distractors
multiplicative
modulation

FIGURE 25

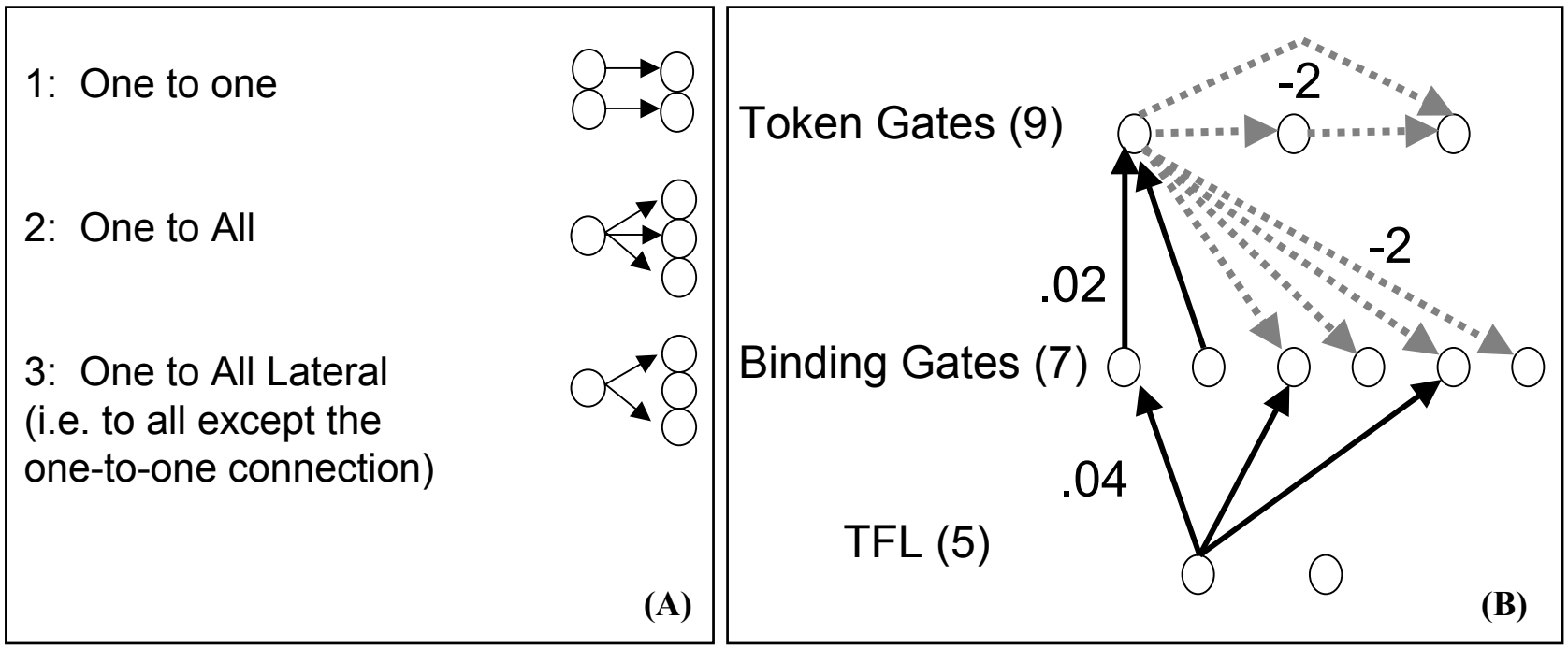

1: One to one
2: One to All
3: One to All Lateral
(i.e. to all except the
one-to-one connection)
(A)
Token Gates (9)
-2
-2
.02
Binding Gates (7)
.04
TFL (5)
(B)

FIGURE 26

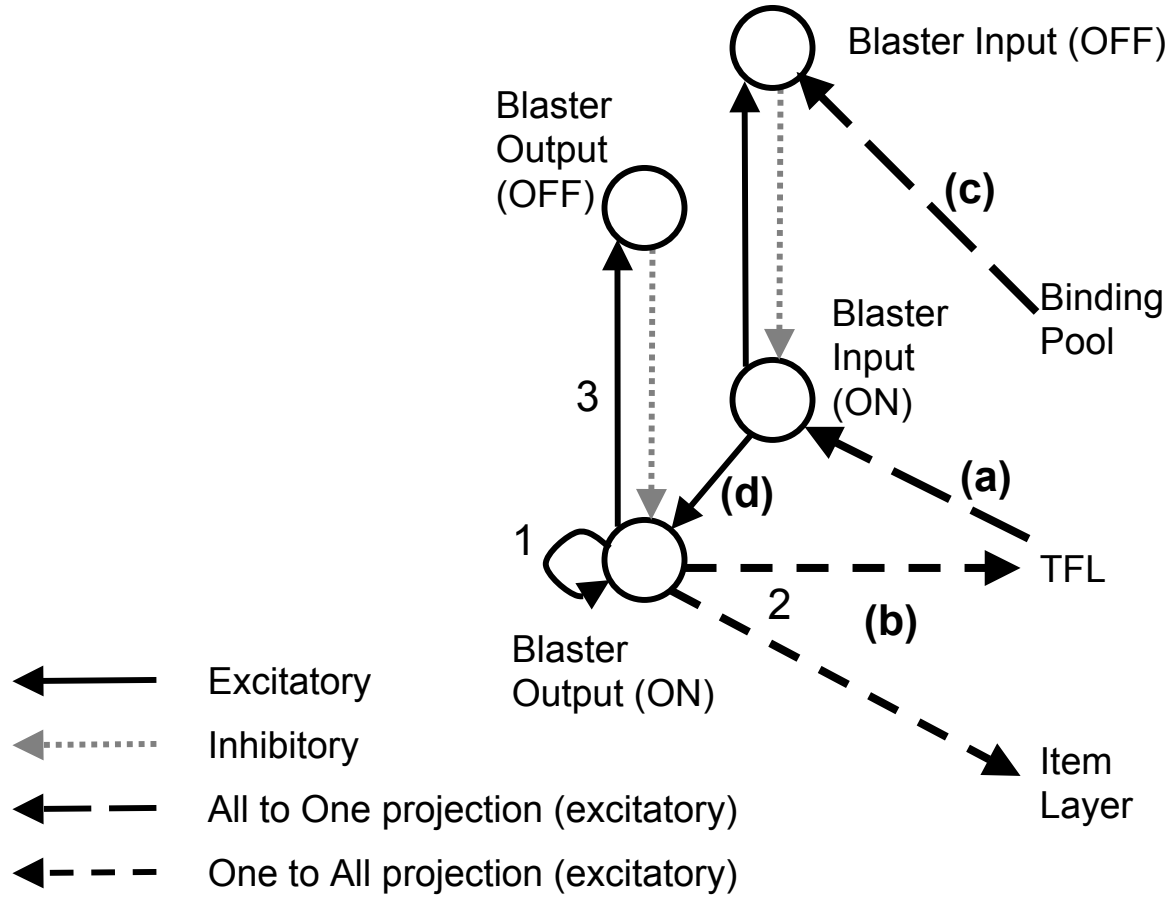

Blaster Input (OFF)
Blaster
Output
(OFF)
(c)
Binding
Pool
Blaster
Input
(ON)
3
(a)
(d)
1
TFL
2
(b)
Blaster
Output (ON)
Item
Layer
Excitatory
Inhibitory
All to One projection (excitatory)
One to All projection (excitatory)

FIGURE 27