



Cognitive Science 36 (2012) 62–101

Copyright © 2011 Cognitive Science Society, Inc. All rights reserved.

ISSN: 0364-0213 print / 1551-6709 online

DOI: 10.1111/j.1551-6709.2011.01213.x

RACE/A: An Architectural Account of the Interactions Between Learning, Task Control, and Retrieval Dynamics

Leendert van Maanen,^{a,b} Hedderik van Rijn,^b Niels Taatgen^c

^a*Cognitive Science Center Amsterdam, University of Amsterdam and Department of Artificial Intelligence, University of Groningen*

^b*Departments of Psychology and Department of Artificial Intelligence, University of Groningen*

^c*Department of Artificial Intelligence, University of Groningen*

Received 7 June 2010; received in revised form 3 February 2011; accepted 16 May 2011

Abstract

This article discusses how sequential sampling models can be integrated in a cognitive architecture. The new theory Retrieval by Accumulating Evidence in an Architecture (RACE/A) combines the level of detail typically provided by sequential sampling models with the level of task complexity typically provided by cognitive architectures. We will use RACE/A to model data from two variants of a picture–word interference task in a psychological refractory period design. These models will demonstrate how RACE/A enables interactions between sequential sampling and long-term declarative learning, and between sequential sampling and task control. In a traditional sequential sampling model, the onset of the process within the task is unclear, as is the number of sampling processes. RACE/A provides a theoretical basis for estimating the onset of sequential sampling processes during task execution and allows for easy modeling of multiple sequential sampling processes within a task.

Keywords: Cognitive architecture; ACT-R; Accumulator model; Sequential sampling process; Drift diffusion model; PRP/dual-task; Picture–word interference

1. Introduction

As theories about cognitive processing become more complex, the need for formal modeling of cognition increases. Because of this, many recent results in cognitive psychology are formulated by formal modeling of the hypothesized processes. Two of the most influential research paradigms in the field of formal modeling of cognition are the sequential sampling framework (e.g., drift diffusion model, Ratcliff, 1978; accumulator model, Vickers,

Correspondence should be sent to Leendert van Maanen, Cognitive Science Center Amsterdam, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands. E-mail: l.vanmaanen@uva.nl

1970; and related connectionist models, e.g., Usher & McClelland, 2001) and cognitive architectures (e.g., ACT-R, Anderson, 2007; EPIC, Meyer & Kieras, 1997a; Soar, Newell, 1990). However, the classes of problems addressed by these two paradigms differ widely. Sequential sampling modelers have traditionally focused on tasks in which participants are required to select a response based on a particular feature of the stimulus (or sometimes a combination of features). These models often provide very detailed accounts of one particular task. However, these tasks are typically quite simple because sequential sampling models cannot easily account for complex interactions between multiple sampling processes. In contrast, the cognitive architecture framework has a tradition in modeling behavior that requires the interaction of multiple aspects of cognition. Cognitive architectures focus more on unification and generalization and models of very complex behavior. Because of the focus on generalization, cognitive architectures sometimes ignore the level of detail required to model low-level decision-making behavior. In this work we want to integrate these research lines to be able to address simple decision-making behavior in the context of complex tasks.

The most prominent success of the sequential sampling framework is the ability to account for both response speed and accuracy within a single theoretical framework (Ratcliff & Smith, 2004). In particular, sequential sampling models provide an elegant explanation for the speed–accuracy tradeoff (Luce, 1986; Wickelgren, 1977) often observed in decision-making behavior. In addition, sequential sampling models often account for the shape of response time distributions (e.g., Ratcliff, 1978), rather than just the mean response time and the associated variance. A third important accomplishment of sequential sampling is the ability to account for response time distributions for both correct responses and incorrect responses. Often people make particular errors in decision making, and sequential sampling provides explanations for the mechanisms that result in fast or slow errors.

Sequential sampling models have provided much insight in the behavioral correlates of decision making. However, computational models that are derived from these theoretical accounts often only model the decision-making process itself. As such, these models fail to appreciate that retrieving declarative knowledge (needed for the decision) does not stand alone but is always part of the execution of a particular task. This aspect becomes especially important when the complexity of the task under scrutiny is increased. For example, a sequential sampling model of a perceptual decision task may be relatively straightforward (e.g., Forstmann et al., 2008; Ratcliff & McKoon, 2008), but performing multiple sequential perceptual decisions as part of operating a complex system (e.g., a flight management system, Taatgen, Huss, Dickison, & Anderson, 2008) may change performance in ways that are hard to account for using only standard sequential sampling concepts.

In contrast, cognitive architectures have had considerable success explaining cognition in an integrative way. This means that theories developed within a cognitive architecture are supported by theories of other aspects of cognition. For instance, for an explanation of visual search tasks, in which participants are asked to search an array of stimuli for a previously presented target stimulus, it is important to include a theory of decision processes to account for a stopping rule of the visual search, and of declarative memory to explain how

participants retrieve the target from memory. Therefore, it makes sense to study different cognitive phenomena within one framework, so that one theory (for instance, about visual search) remains consistent with others (for instance, about decision making or declarative memory).

The integration of multiple aspects of cognition becomes especially apparent when studying the concurrent execution of multiple tasks (Salvucci & Taatgen, 2008). Executing a secondary task interferes in many cases with performance on a primary task, and cognitive architectural models have provided explanations of the specific bottlenecks in the combined task execution (e.g., Borst, Taatgen, & Van Rijn, 2010; Meyer & Kieras, 1997a). In addition, cognitive architectures often incorporate theories on the changes the cognitive system undergoes as a result of learning (e.g., Anderson & Schooler, 1991; Laird, Rosenbloom, & Newell, 1986; Taatgen & Lee, 2003). As a result, changing behavior in complex tasks can be modeled as well (e.g., Anderson, Taatgen, & Byrne, 2005; Taatgen & Anderson, 2002; Van Rij, van Rijn, & Hendriks, 2010; Van Rijn, Van Someren, & van der Maas, 2003).

Sequential sampling models have provided much insight in the behavioral correlates of decision making. Cognitive architectures on the other hand provide a theory of task execution (Newell, 1990), but the explanation provided by these models is not always at the level of detail of sequential sampling models. The theory put forward in this article reconciles both approaches. In the remainder of this article, we will present an integrated theory of a sequential sampling model and a cognitive architecture. As a starting point we will introduce the cognitive architecture ACT-R (Anderson, 2007). Secondly, we will discuss those aspects of the sequential sampling framework that are relevant for the integrated theory. Third, we will present the integrated theory itself, called Retrieval by Accumulating Evidence in an Architecture (RACE/A). After RACE/A has been introduced, we will demonstrate its importance in two experiments that are exemplary for the kind of problems that RACE/A is intended to address. In Experiment 1, participants were asked to perform a picture–word interference (PWI) task in which certain stimuli are repeated. In addition and at the same time, participants were given a secondary task (this experimental design constitutes a psychological refractory period [PRP] design). The PRP design enables us to temporally separate multiple stages in the execution of the PWI task. This temporal separation has a differentiating effect on the sequential sampling processes involved in PWI. The cognitive architecture provides a way of scheduling these processes that is not included in the default sequential sampling framework. At the same time this type of behavior cannot be modeled without a sequential sampling component, because such a model would lack a theoretical account of the competitive processes involved in PWI. Experiment 1 demonstrates how RACE/A accounts for the interaction between short-term memory processes and long-term memory processes.

In Experiment 2, participants are again required to perform a PWI task in a PRP design. However, this time an extra condition is included that influences participants' behavior in the other conditions. RACE/A provides an explanation of the response time data in terms of the scheduling of sequential sampling processes and the way they influence each other. We will again discuss why this pattern is hard to account for using only a traditional cognitive architecture, or only a traditional sequential sampling model.

1.1. ACT-R

ACT-R is a hybrid cognitive architecture in which behavior on a task can be described by a sequence of production rule executions. The rules specify which actions may be executed given certain conditions. If the conditional side of a production rule matches against the current information state, then the actions associated with that rule are executed. The current information state is represented by a set of buffers, each containing one piece of information. Which information is present at a certain point in time in the buffers is determined by specialized modules that each process one type of information. For instance, the visual module handles visual perception, and the motor module executes motor commands. The declarative module is used for storing and retrieving declarative memory information, the speech module handles the speech output, the aural module handles auditory perception, and the goal and imaginal are modules for keeping track of (sub)goals and intentions. The modules can be regarded as theories on that particular aspect of cognition, and the production rule system connects these theories to account for overall behavior (Gray, 2007). For example, the declarative module implements a theory of declarative memory that is to a large extent based on Anderson's earlier work on memory representations (Anderson & Milson, 1989; Anderson & Schooler, 1991) and Instance Theory (Logan, 1988).

A typical ACT-R model run consists of the sequential execution of production rules. Each production rule triggers a set of actions, which may change the buffers' status, which in turn may trigger another production rule. Because each action requires a certain amount of time, the model latency is an aggregation of the timing of all actions (cf. Sternberg, 1969). This is illustrated by Fig. 1A. Each line in the figure represents a different type of information processing module, with the bottom line representing the production rule module. The visual module processes the stimulus (the elephant), followed by a production rule that starts a memory retrieval. If a fact is retrieved (in this case, the word "Elephant"), a next production rule triggers a vocal response. Note, however, that, although not discussed in this example, modules may operate in parallel, so the latency is not simply the sum of the timing of all actions. For example, the visual module may be busy processing new information, while at the same time the declarative module may retrieve factual information. Where the modules may operate in parallel, each individual module can only execute one action at a time. Therefore, the timing is also influenced by how the model schedules all the subprocesses involved in executing the task (Byrne & Anderson, 2001). RACE/A copies the way ACT-R schedules task execution. That is, a model in RACE/A also consists of a set of production rules that are triggered when a particular state is met. In addition, RACE/A copies the processing theories of most of the individual modules. The exceptions to this are the declarative memory module and the visual module. These modules are extended with a new retrieval mechanism based on sequential sampling theory. To fully appreciate the changes in the retrieval process of ACT-R that we propose, we will first discuss how memory retrieval is currently modeled in ACT-R.

The current ACT-R mechanism of retrieval is as follows. Each fact (called a chunk) is represented by an activation value representing the log odds that a chunk will be needed in

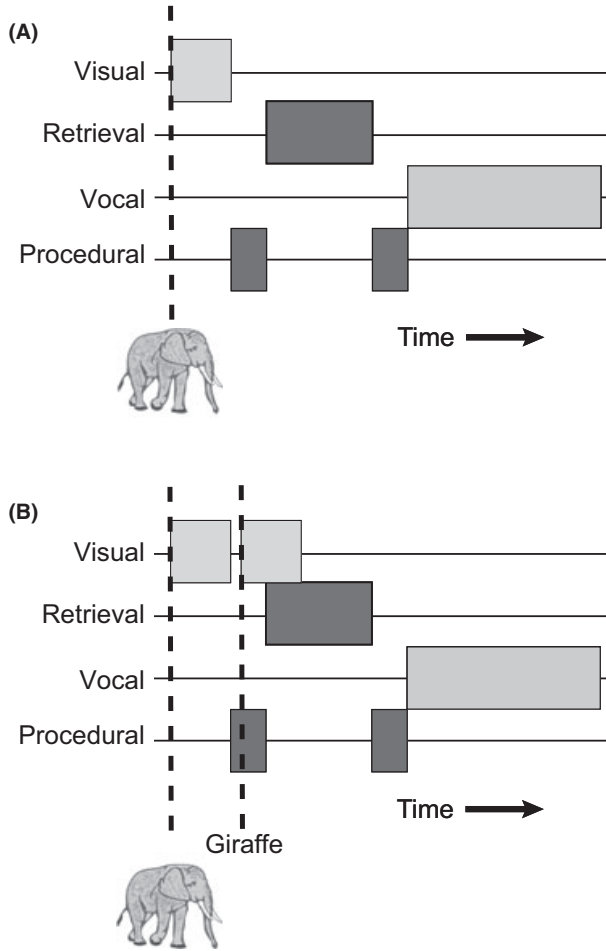


Fig. 1. (A) Schematic representation of a simplified ACT-R model of perception and subsequent naming of an elephant. Each horizontal line represents processing in a dedicated module. The vertical dashed line represents the stimulus onset. (B) In a standard ACT-R model, if the word “Giraffe” is presented after the elephant picture, it cannot influence the retrieval process anymore.

the near future. The log odds is partly determined by a component describing the history of usage of a chunk called the *base-level activation* (B_i , Anderson & Schooler, 1991):

$$B_i(t) = \ln \left(\sum_{j=1}^n (t - t_j)^{-d} \right). \tag{1}$$

The base-level activation equation reflects the theory that declarative memory is optimally adapted to the environment. That is, chunks that are most active are the ones that are most likely needed, given the demands of the environment. By incorporating both the

frequency with which particular information is used and the recency of these occurrences, the base-level activation predicts learning effects as well as forgetting (e.g., the power laws of learning and forgetting; Anderson, Fincham, & Douglass, 1999). In Eq. 1, t_j represents the time of the j th presentation of a memory chunk and d is the parameter that controls decay, which in most ACT-R models is fixed at 0.5 (Anderson et al., 2004).

In the standard conception of ACT-R, the total activation is the sum of the base-level activation, moment-to-moment noise ($\varepsilon(t)$ in Eq. 2), and another component describing the influence of the current context (*spreading activation*). The spreading activation component in ACT-R—represented in Eq. 2 by the summation—is composed of the associative values (S_{ji}) of other chunks (chunks j in Eq. 2) to chunk i , weighed by W_j , representing the importance of the associated chunks:

$$A_i(t) = B_i(t) + \sum_j W_j S_{ji} + \varepsilon(t). \quad (2)$$

Currently, the assumption behind ACT-R's spreading activation mechanism is that chunks that are temporarily available to central cognition (i.e., chunks that are present in the buffers) increase the probability that related chunks will be needed. The association that exists between two chunks (S_{ji}), reflects the pattern of co-occurrences of the two chunks in previous cognitive processing (Anderson & Milson, 1989). For instance, in the presence of a green stimulus in the visual buffer, the probability of retrieval of chunks that are related to green—such as a chunk representing grass or a chunk representing the concept of Ireland—increases. This is because grass and green as well as Ireland and green often co-occur.

If a production rule fires that requests the retrieval of a chunk with a particular set of properties, ACT-R selects the most active chunk that matches the request and computes a retrieval latency. This default ACT-R retrieval process can be thought of as a one-shot process. Currently, ACT-R does not have an explanation of how the retrieval behavior would change if new and relevant information would become available *during* the retrieval process. This is illustrated by Fig. 1B. If information becomes available after the retrieval process has been initiated, a default ACT-R model would first complete that retrieval process before considering any new information. However, from studies in which the presentation duration of stimuli was manipulated, it is known that even for very short time intervals (<100 ms) the current visual information state influences cognitive processing (e.g., Glaser & Dünghoff, 1984; Marcel, 1983). For example, Glaser and Dünghoff found that presenting different kinds of words 50 ms after a picture reliably changes response times, which is a clear indication that the retrieval time has been influenced. In addition, ACT-R does not account for competitive effects that arise when multiple chunks match the retrieval request. Instead, the model would just return the most active chunk (but see Van Rijn & Anderson, 2003; Schneider & Anderson, 2011 for alternative approaches).

The RACE/A theory extends ACT-R by substituting the one-shot retrieval mechanism by a more dynamic component that models the retrieval process. Thus, the static retrieval model in ACT-R is substituted by one in which the activation is continuously updated.

In addition to changes to the declarative memory module of ACT-R, RACE/A also differs from standard ACT-R in the way visual information processing is modeled. The current visual information processing theory in ACT-R assumes that conceptual information related to a visual stimulus is available in the visual buffer 85 ms after stimulus presentation (Anderson, 2007). The exact mechanisms behind this process are yet unaccounted for (although a number of mechanisms have been proposed, e.g., Fleetwood & Byrne, 2006; Salvucci, 2001; Tamborello & Byrne, 2007). While RACE/A does not aspire to present a definitive account of visual processing, it assumes that an evidence accumulation mechanism has some access to (partial) information even before the 85 ms threshold has passed. This information can therefore be used to model the process of retrieving stimulus-related information (Treisman & Gelade, 1980).

1.2. Sequential sampling framework

Underlying many sequential sampling models is the idea that response options are neurally represented by pools of neurons, the firing rates of which are inherently noisy. Therefore, to make a decision (i.e., choose one of the options), multiple samples of these firing neurons must be accumulated to retrieve a stable representation of the decision (for reviews, see Smith & Ratcliff, 2004; Wang, 2008). This is often represented by racing stochastic processes, in which each process accumulates information with a particular rate until one of the values crosses a preset decision criterion value (see Fig. 2). Sequential sampling models are often used to describe competitive processes in which two or more response options are in competition. Typically, the first response option to reach the criterion value is the option that is chosen, and the time needed to reach the criterion is the retrieval time.

An important question in the study of decision making is whether the decision criterion consists of an absolute value (e.g., Vickers, 1970) or reflects a difference between

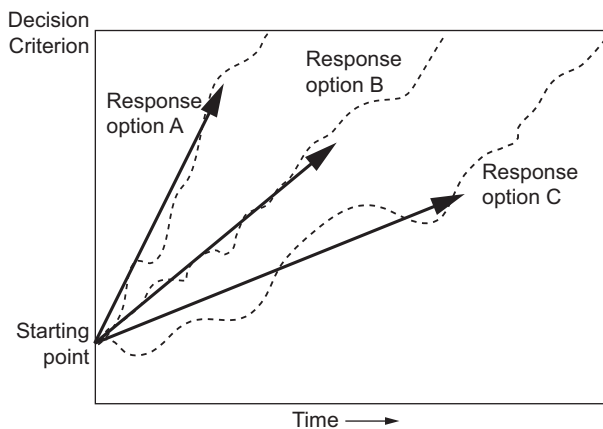


Fig. 2. Schematic representation of a typical accumulator model. The arrows indicate the mean drift rate for each alternative. In this case, response option A is the first to reach the decision criterion, which is here represented by an absolute value.

accumulators (e.g., Ratcliff, 1978). In models with an absolute decision criterion, all accumulators race until one of them crosses a preset threshold value. In models with a relative decision criterion, the decision criterion takes the value of the highest accumulator relative to the value of all other accumulators. RACE/A will adopt a relative decision criterion because such an implementation is most in line with the notion of activation in ACT-R, which reflects a relative log odds of the need of a particular chunk.

A second important property of any sequential sampling model is the start point of accumulation. Often this is assumed to be equal for all accumulators (as in Fig. 2), reflecting no prior preference for any of the response alternatives, but it could also be biased toward one of the alternatives (Morton, 1969; Ratcliff & McKoon, 1997). The start point of accumulation is often considered as a free parameter when fitting data (e.g., Wagenmakers, Ratcliff, Gomez, & McKoon, 2008) or assumed to be equal for all accumulators (e.g., Wagenmakers, Van der Maas, & Grasman, 2007). In RACE/A the start point of accumulation is based on the history of usage of the accumulating response options. For this, RACE/A applies the activation values of chunks, taking into account how a priori likely it is that one of the chunks is required for further processing (Anderson & Schooler, 1991).

A third parameter that is relevant in most sequential sampling models is the mean drift rate. The mean drift rate represents the average speed with which a particular process accumulates. Drift rate differences are often taken to represent differences between the stimuli. For example, responses to frequent items are generally faster than responses to infrequent items (e.g., in lexical decision tasks, Wagenmakers et al., 2008), which can be represented by a higher drift rate for frequent items. In RACE/A, the drift will also be determined by properties of the stimulus. In addition, as we will discuss in more detail later, the drift also depends on associations with other chunks.

2. Retrieval by accumulating evidence in an architecture

In this section, we will develop the dynamics of the RACE/A theory. We will first present the equations that govern the activation dynamics and show how they relate to both ACT-R concepts and sequential sampling concepts. Next, we will show how RACE/A can account for reaction time effects.

The accumulation process can be characterized by two equations that determine the long-term dynamics and the short-term dynamics of the activation. The long-term dynamics of the activation are expressed by the default base-level activation equation from ACT-R (B_i , Eq. 1). The short-term dynamics are mediated by spreading activation from other chunks and the presence or absence of perceptual stimuli (C_i , described below). Thus, the total activation in RACE/A is

$$A_i(t) = B_i(t) + C_i(t). \quad (3)$$

This indicates that the total activation of a chunk is the sum of the log odds of a chunk (reflected by B_i) and the accumulated evidence for that chunk (reflected by C_i). Note that

this new computation of activation (Eq. 3) includes a base-level activation component (B_i) and a context component (C_i), analogous to the original ACT-R activation equation (Eq. 2).

The accumulation process continues until a certain decision criterion is reached. The duration of the accumulation process is defined as the time from a request for a retrieval of information (scheduled by a production rule) until the decision criterion is reached.

Similar to most sequential sampling models, the short-term activation dynamics can be represented by a starting point, a drift, and a decision criterion, which we discuss below.

2.1. Starting point

The starting point of the accumulation reflects the prior probability that a chunk is needed (Morton, 1969; Ratcliff & McKoon, 1997). This is represented by the current activation level of a chunk at the onset of retrieval. This can involve the usage history of a chunk, reflected in the base-level activation, as well as short-term activation. Chunks with a high base-level activation start the accumulation process at a higher starting point and are thus more likely to be retrieved from memory. In addition, a retrieval that is scheduled shortly after another retrieval is still influenced by the activation that accumulated during the previous retrieval. The amount of residual activation depends on a decay parameter discussed in the next section. This approach has a history in sequential sampling models (Morton, 1969; Ratcliff & McKoon, 1997) in which sometimes prior probabilities are considered as starting points of the accumulation process. The approach also has a strong link to activation modeling in traditional cognitive architectures. ACT-R, for example, incorporates a rational analysis of memory (Anderson & Schooler, 1991), meaning that activation in ACT-R reflects a higher probability of retrieval of frequently used chunks (as well as recently used chunks).

2.2. Drift

Drift in RACE/A is a function of stimuli in the environment, as well as the currently active internal chunks, such as declarative facts. All facts and stimuli, which will be collectively referred to as *sources of activation*, continuously spread excitatory activation toward associated chunks. This means that chunks with more sources of activation (more evidence) or with sources with more activation (“stronger” evidence) accumulate faster than chunks with less sources of activation or sources with less activation. In the absence of evidence for a particular chunk, the short-term activation will decay.

We incorporated decay, spreading activation, external activation from stimuli, and a noise term in Eq. 4:¹

$$dC_i = [-\alpha C_i + \beta \sum_j S_{ji} A_j + \gamma \sum_k S_{ki} E_k + \varepsilon_i] dt. \quad (4)$$

Eq. 4 captures the dynamics of short-term activation of one chunk (chunk i) over time.

The first term in Eq. 4 represents the decay of short-term activation. The speed of decay is controlled by a parameter α . In the absence of spreading activation, the short-term

activation of a chunk i will decay to 0. Because the total activation of a chunk is the sum of the short-term activation and the base-level activation (Eq. 3), the activation will decay to its base level.

The second term in Eq. 4 represents spreading activation from other chunks. Spreading activation is the sum of the activation of other chunks (A_j , which is the sum of B_j and C_j), weighted by the associations that exist with chunk i (S_{ji}). The spreading activation term is scaled by a factor β .

The third term in Eq. 4 represents the influence of external factors (i.e., stimuli, E_k). Of importance is that a stimulus will only activate a subset of chunks (i) that are related to the stimulus (S_{ki}). The external activation term is scaled by a factor γ . Because it is often hard to disentangle the influences of these different parameters, in practice γ will be estimated as 1 and S_{ki} will be estimated as either 1 or 0, indicating the presence or absence of a relation between chunk and stimuli. All stimulus-related variance will thus be accounted for by estimating E_k .

Finally, noise over the activation value of a chunk is expressed by a noise sample ε_i that is drawn at each time step from a logistic distribution. The logistic distribution was chosen here because it approximates the normal distribution but is computationally simpler (cf. Anderson & Lebiere, 1998, p. 64).

To summarize, the decay parameter α together with scaling factors β and γ determine the drift of the chunks in the system. The chunk that receives the most spreading activation (scaled by β and γ) from sources of activation will (in the absence of noise) be the first to reach the decision criterion.

Eq. 4 resembles—and is in fact derived from—the leaky competitive accumulator model (LCAM; Usher & McClelland, 2001). In LCAM the activation of a unit (a chunk) depends on an input term, a decay parameter (“net leakage”), and lateral inhibition between the competing units. Eq. 4 also contains a term for external input ($\gamma \sum S_{ki} E_k$) and a decay term (α). The crucial difference between the two models is that RACE/A assumes an excitatory spreading activation term ($\beta \sum S_{ji} A_j$), whereas LCAM assumes an inhibitory term. This difference reflects our different choice of decision criterion. Where Usher and McClelland (2001) used an absolute criterion to terminate the accumulation process, RACE/A uses a relative criterion. This warrants different interactions between the units, as discussed below.

2.3. Decision criterion

We chose a relative decision criterion for RACE/A. This choice is motivated by considerations of Bayesian optimality, in line with standard ACT-R. We assume that a memory retrieval is optimal if the expected retrieval time is minimized for a particular error rate (Brown, Steyvers, & Wagenmakers, 2009). Thus, the system continuously checks if there is enough evidence for one of the retrieval options that the probability that that particular option is correct is on or above the error rate. Optimality in this sense means continuously computing the posterior probability for each option until one of these exceeds a criterion value (Dragalin, Tartakovsky, & Veeravalli, 1999).² As in all sequential sampling models, the time needed to exceed the criterion value constitutes the retrieval time (e.g., Ratcliff, 1978).

Because activation values in RACE/A (and ACT-R) represent the log odds that a particular chunk is needed or relevant, the posterior probability for each chunk can be retrieved by exponentiating the activation values:

$$\frac{e^{A_i}}{\sum_j e^{A_j}} \geq \theta. \quad (5)$$

Besides being optimal, this formulation has a number of benefits over an absolute decision criterion. A first benefit is that the posterior probability decreases if the associative value between chunks increases. This means that competition resulting in slower reaction times can be modeled without resorting to an explicit inhibitory mechanism. This is in line with the rational analysis framework incorporated in ACT-R: Chunks that often co-occur have a strong positive association (Anderson & Milson, 1989) and not a negative one. Therefore, items that often co-occur but have a slowing-down effect on each other can be accounted for (Van Maanen and Van Rijn, 2007a, 2007b). A second benefit is that this formulation of a decision mechanism has an explicit neural implementation. Basal ganglia circuits that are involved in neural decision making have been shown to compute activation values for a certain decision criterion in a similar way as RACE/A (Bogacz & Gurney, 2007; Zhang & Bogacz, 2010). In particular, Bogacz and colleagues hypothesize that basal ganglia implement optimal decision making by constantly monitoring the ratio of activations until the activation of one of the alternatives crosses a preset criterion value. The mathematical model they propose for this process implements Eq. 5. Therefore, RACE/A can be considered neurally plausible.

2.4. Activation dynamics

Before discussing memory retrievals using RACE/A in the broader scope of complex task execution, it is important to understand the activation dynamics of this model. In this section, we present the results of a set of simple simulations of how RACE/A behaves under varying circumstances. In these simulations, two unrelated chunks are competing for retrieval from declarative memory. Because there are only two unrelated chunks in the system, Eq. 4 can be expressed as

$$dC_i = [-\alpha C_i + S_{ji}E_j]dt. \quad (6)$$

In Eq. 6, j represents the stimulus associated with chunk i , and E_j the strength of that stimulus. For tractability, we assume, here, that the base-level activation is a constant value. Because in this simulation the chunks are unrelated, the $S_{ji}A_j$ term from Eq. 4 is zero, which means that the activation growth in the model does not depend on spreading activation, which has therefore been excluded from Eq. 6. For explanatory purposes, we set the noise component to zero.

Fig. 3A (left) presents how activation develops over time for the default case, in which both chunks have the same starting point, and one stimulus is present that

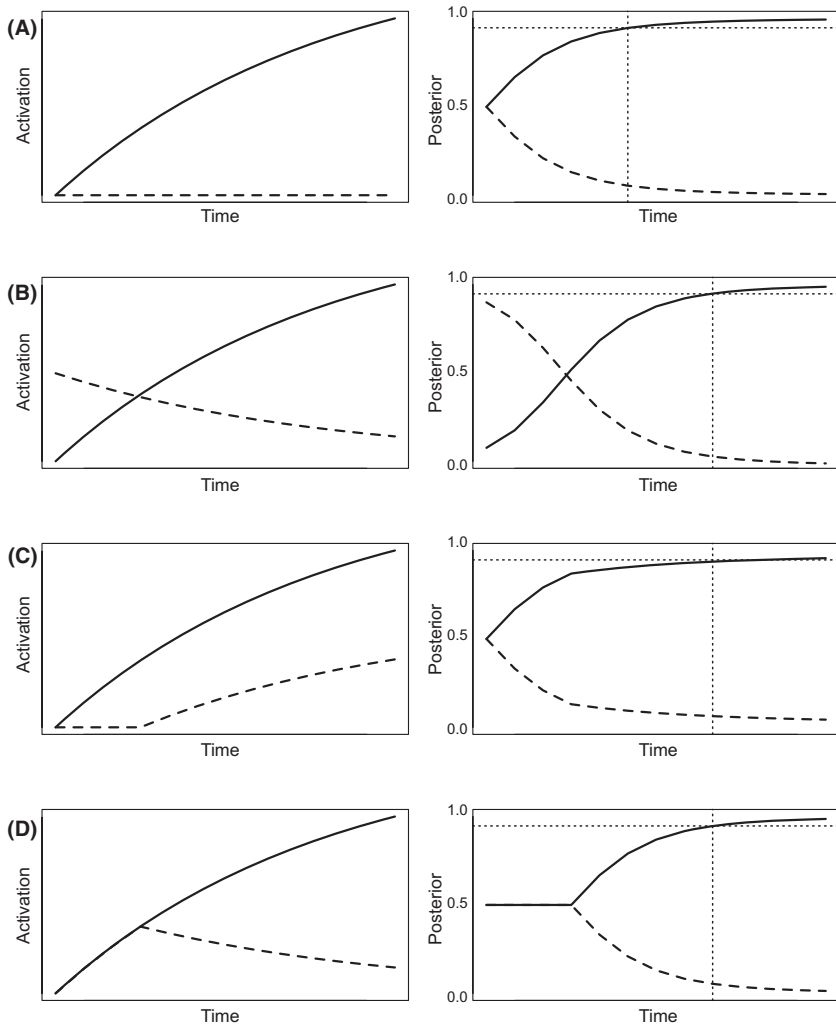


Fig. 3. RACE/A dynamics for four typical retrieval processes without spreading activation. Left: activation of two chunks. Right: posterior probabilities for two chunks. The vertical dotted lines indicate the retrieval times. See text for a description of panels A through D. RACE/A, Retrieval by Accumulating Evidence in an Architecture.

activates one chunk. That chunk accumulates activation, whereas the other remains at the value that is determined by the base-level activation. Fig. 3A (right) presents the probability for both chunks, with the dotted horizontal line indicating when the decision criterion (θ) was met. Because there are two equally probable chunks in this competition, the probability at the start of the process is 0.5 for both chunks, but it quickly changes in favor of the accumulating chunk, until it crosses the decision criterion, and a choice is made.

Fig. 3B presents the activations and probabilities for two chunks, one of which receives activation from a stimulus, but the other chunk (the non-target chunk) has a higher starting point. This is for instance the case if the non-target chunk has a higher activation because it has been recently retrieved. The activation of the non-target chunk will decay to its base-level activation, while the activation of the target chunk increases. Under these conditions, retrieval of the target chunk takes longer because the initial posterior probability for the activated chunk is smaller due to the higher initial activation of the non-target chunk.

In Fig. 3C, a situation is depicted in which one chunk is activated at a later moment in time than the other. This has been hypothesized to represent PWI experiments in which the presentation of the two stimulus dimensions (picture and word) are temporally separated (Van Maanen & Van Rijn, 2007a). Even though the retrieval process of the first chunk has already been initiated, activation of the second chunk at a later moment in time influences the posterior probability and thus increases the retrieval time of the first chunk.

The last simulation addresses a condition in which two chunks are activated, but one of the chunks is only activated for a short duration (Fig. 3D). This is similar to masked priming experiments in which one stimulus is only available for a short duration and thus only has a short interval in which to activate related concepts, while another stimulus remains present for a longer interval (Van Maanen & Van Rijn, 2007b). Initially, both chunks accumulate, because both receive activation from a stimulus. After one of the stimuli disappears, the activation of one of the chunks decays, but the other chunk accumulates further. Again, the retrieval time is increased as compared to the default situation in Fig. 3A, because the activation of the decaying chunk still influences the posterior probability. The simulations depicted by Fig. 3C,D are difficult to explain with the traditional approach to memory in cognitive architectures, because they simulate conditions in which the sources of activation change *during* the retrieval process, thereby altering the accumulation of evidence and the retrieval time. Because of the static nature of most cognitive architectures, any information that becomes available during the retrieval process would not be able to influence the process. Therefore, these kind of short-term effects on cognitive processing discussed here are hard to include in a traditional cognitive architectural model.

To summarize, the posterior probability for a particular chunk is negatively accelerated by the activation of other chunks that compete with that chunk for retrieval from declarative memory. This competitive process results in an increase in retrieval time.

In all simulations presented in Fig. 3 the chunks were not associated. If the chunks are associated, the retrieval times are longer (Fig. 4). Under these conditions, both chunks accumulate due to spreading activation from the other chunk (cf. Fig. 4A with Fig. 3A). This “co-activation” of the chunks results in slower accrual of the posterior probabilities, and therefore a longer retrieval (Fig. 5). Many effects have been associated with a differentiation of associative strength, including the semantic gradient effect found in PWI studies (e.g., W. R. Glaser & Döngelhoff, 1984; Klein, 1964; Rayner & Springer, 1986; Smith & Magee, 1980) and the fan effect (Anderson, 1974; Anderson & Reder, 1999). If two concepts are associated, for instance, because they belong to the same semantic category, they will excite each other, increasing decision criterion and thus the retrieval time and the response latency.

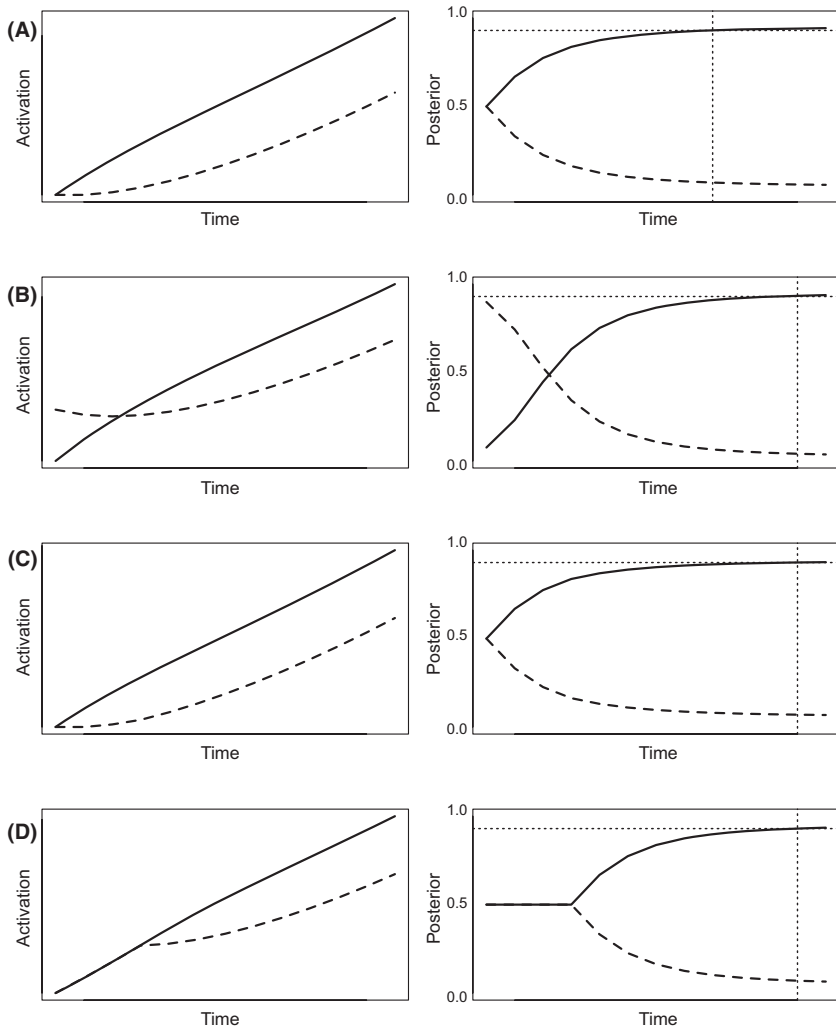


Fig. 4. RACE/A dynamics for four typical retrieval processes with spreading activation between the chunks. Left: activation of two chunks. Right: posterior probabilities for two chunks. The vertical dotted lines indicate that a chunk has been retrieved. Note that the scales in this figure differ from Fig. 3. RACE/A, Retrieval by Accumulating Evidence in an Architecture.

2.5. Architectural aspects of RACE/A

The previous sections illustrated the model's behavior in various circumstances, focusing on the short-term dynamics of RACE/A. We will now turn to the role of the cognitive architecture in RACE/A. Recall that in ACT-R the timing of all the subprocesses depends on the (serial) execution of production rules and the (parallel) execution of module operations. As a result, the particular specification of production rules determines *when* a particular

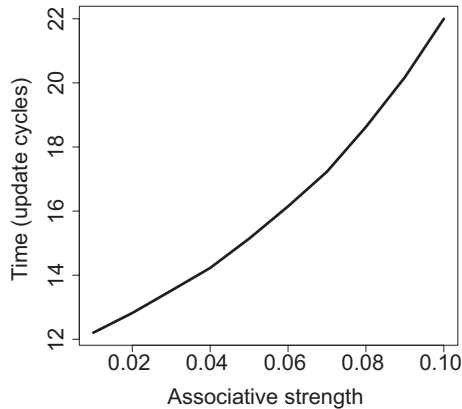


Fig. 5. Retrieval times as a function of associative strength in RACE/A. Time is represented as update cycles. RACE/A, Retrieval by Accumulating Evidence in an Architecture.

subprocess commences. In RACE/A, this means that the production rules determine when an accumulation process starts. The exact start time of an accumulator is important because the accumulation process depends on the information that is present during the accumulation, which may differ depending on the start time.

In addition, the timing of the accumulation process also has an impact on the starting point value of accumulation. First, as indicated by Eq. 3, the accumulation of activation of a chunk starts at the base-level activation of that chunk (expressed by Eq. 1). Decay of base-level activation insures that as time progresses the base-level activation will be lower. Therefore, if an accumulation process is started later in time, the starting point of the accumulation process will be lower. Conversely, if by delaying the accumulation process an additional encounter of a chunk has increased the base-level activation, the starting point of accumulation for that chunk will be higher. It should be noted that sequential sampling models could be devised to account for these parameter changes (e.g., Dutilh, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2009 describe how stimulus repetitions affect sequential sampling model parameters). However, RACE/A provides a theory-driven way of estimating these values, as computation of the base-level activation values is firmly grounded in memory research (e.g., Anderson & Schooler, 1991).

A second way in which the starting point of accumulation can be altered is by assuming multiple accumulation processes in sequence. If two processes are scheduled to occur in close temporal proximity, and some (or all) of the chunks in the second process also accumulated activation in the first process, then for those repeated chunks the starting point of accumulation will be increased when the second process starts. This differs from the context in which the base-level activation was increased due to an additional encounter, because the repeated chunks are not necessarily retrieved in the first accumulation process. Because of decay of the accumulated activation, the interval between the accumulation processes matters for the starting point estimate as well.

The cognitive architectural aspects of RACE/A provide a theory on learning of declarative facts (implemented with the base-level activation) and a theory of task control (implemented with production rules). The interaction of these with a model of short-term activation dynamics provided by the sequential sampling aspects of RACE/A will enable us to explain complex behavior at a more detailed level. This will be tested in the following sections. In Experiment 1 participants perform a complex task in which stimuli are repeated. This is subsequently modeled using RACE/A. In particular, we make use of the short-term activation dynamics inspired by the sequential sampling framework in combination with the learning mechanisms inherited from ACT-R. To illustrate how the particular task setup changes behavior, and how this can be modeled with RACE/A, we performed Experiment 2. Experiment 2 is equal to Experiment 1 except for the inclusion of an additional condition that alters the behavior associated with the other conditions. The assumption underlying the model of Experiment 2 is that participants deploy a slightly different task strategy than in Experiment 1 that explains the different behavior in this task.

3. Experiment 1: Repetition priming in a dual task

In Experiment 1, participants are required to perform a PWI task in a PRP design. The experiment is designed to illustrate two aspects of RACE/A. The first aspect is the interaction between short-term memory processes and long-term memory processes. Following a rational analysis of memory, RACE/A assumes that information that is encountered multiple times will be increasingly likely to be needed in the future. This is reflected in an adapted base-level activation value for those chunks (Eq. 1) representing an increased starting point of the hypothesized accumulation processes underlying PWI (Van Maanen & Van Rijn, 2007a, 2010; Van Maanen, Van Rijn, & Borst, 2009). In Experiment 1 pictures are repeated multiple times. RACE/A predicts that the repetitions increase the base-level activation of the picture-related chunks, leading to increased starting points of the picture-related chunks and thus to shorter accumulation processes.

The second aspect of RACE/A addressed by Experiment 1 relates to the interference effects in a PRP task. In a PRP design, participants are asked to perform two tasks sequentially. The first task is often relatively simple, whereas the second task is the task of interest. The interval between the stimulus onsets of the two tasks is manipulated (stimulus onset asynchrony [SOA]). A typical finding, known as the PRP effect (Telford, 1931), is a negative correlation between SOA and response latency on the second task. Responses to the first task are typically unaffected by varying the SOA. The PRP effect has been explained by the assumption that both tasks share a cognitive resource that can only be used by one task at a time (e.g., McCann & Johnston, 1992; Pashler, 1994; Welford, 1967, 1980; but see Meyer & Kieras, 1997a). Thus, the second task is delayed at short SOAs because the first task still requires a critical resource, as illustrated in Fig. 6. As the interval between the tasks increases, the delay becomes smaller, resulting in a faster response on the second task.

Accounting for complex multi-task designs is one of the strong points of many cognitive architectures. For example, the EPIC architecture has been extensively applied to study PRP

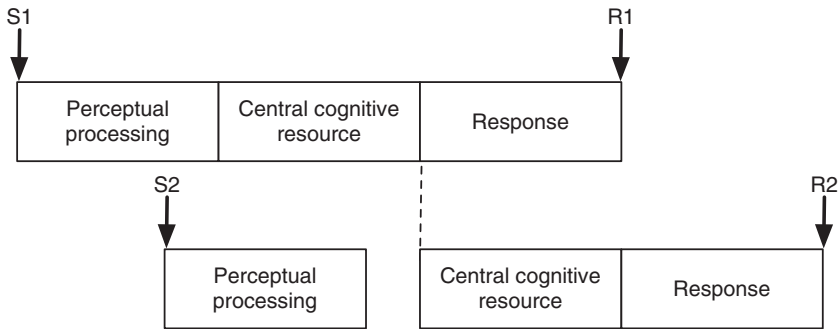


Fig. 6. Gantt diagram of the PRP effect. The top bar indicates processing in the first task. The bottom bar indicates processing in the second task. S1, stimulus of task 1; S2, stimulus of task 2; R1, response on task 1; R2, response on task 2. PRP, psychological refractory period.

effects (e.g., Meyer & Kieras, 1997a, 1997b) and ACT-R has been applied to multi-task interference as well (e.g., Altmann & Gray, 2008; Borst et al., 2010; Salvucci & Taatgen, 2008; Taatgen, Van Rijn, & Anderson, 2007). Therefore, in order to account for the PRP aspects of Experiment 1, it makes sense to use a cognitive architecture approach.

Applying the PRP design to PWI (in which PWI is the second task) Dell'Acqua, Job, Peressotti, and Pascali (2007) have shown that the effect size of PWI decreases with decreasing SOA. In line with the typical PRP rationale, Dell'Acqua et al. argued that these results indicate that the locus of interference in PWI is located before the singular resource that both tasks share. The reasoning behind this is that an interval between the presentation of the stimuli of the first and second task generates a delay in execution of the second task (referred to as ‘cognitive slack time’). This delay is caused because the shared resource is not free for the second task when it is needed, and therefore task execution is delayed (Fig. 6). Dell'Acqua et al. argue that because the effect locus of PWI is before the shared resource, participants use the cognitive slack time to resolve the interference. If the interval increases, the delay of the second task decreases. The interference cannot be resolved in the slack time anymore and therefore becomes apparent in the reaction times.

Thus, Experiment 1 includes aspects that are typically modeled within the sequential sampling framework (interference) as well as aspects that are often modeled using cognitive architectures (PRP). However, we contend that in order to provide a formal theory of the combination of these aspects, a combination of modeling techniques is required.

3.1. Methods

3.1.1. Participants

Twenty-three students of the University of Groningen (mean age 22.7, 14 male, 9 female) took part in this experiment for course credit. All were native speakers of Dutch and had normal hearing and normal or corrected-to-normal vision.

3.1.2. Stimuli

Forty-nine images were taken from an image set (PD/DPSS, Dell'Acqua, Lotto, & Job, 2000). The images that were selected for inclusion in this study had a naming agreement of 95%. Of each image, two PWI stimuli were created that consisted of the image, with a word written in the center of the image. The words were selected as follows: For the related condition, category members of the image descriptors were selected. The words for the unrelated condition were then selected from a lexical database (CELEX; Baayen, Piepenbrock, & Van Rijn, 1993) and matched to the related distractors with respect to word length (plus or minus 1 letter) and word frequency ($\pm 10\%$).

The tones for the primary task consisted of a 300, 600, and 1,200 Hz tone, similar to the experiment conducted by Dell'Acqua et al. (2007).

3.1.3. Design

For each participant, an experimental list was created in which each image was combined with each relatedness condition (related and unrelated) and every SOA (100, 350, and 800 ms). This resulted in 294 trials per participant, in which each picture was repeated six times. The lists were pseudo-randomized in such a way that (a) the same condition (relatedness or SOA) did not occur more than twice in a row, and (b) the same tone did not occur more than twice in a row.

3.1.4. Procedure

Each trial started with the presentation of a fixation cross for 1,500 ms followed by the tone-classification tone for 150 ms. After the SOA, the PWI stimulus was presented. The word and the picture that formed the PWI stimulus were presented simultaneously. The participants were instructed to always respond to the tone first, and then to the PWI stimulus. The tone had to be classified as low, medium, or high pitched by pressing the *b*, *n*, or *m* keys, respectively, with the index, middle, and ring fingers of the right hand. For the PWI task, the participants were instructed to name the picture. Responses were recorded using a voice response box. If participants failed to answer in the correct sequence, a screen reminding them of the correct procedure was presented.

The participants were tested individually. First, each participant practiced the tone-classification task in isolation (15 trials). Second, participants were familiarized with the PWI procedure by responding to a set of non-experimental PWI stimuli in single-task setting. Both PWI conditions were practiced for nine trials, for a total of 18 trials. Before each of these practice blocks, the speed of responding was stressed as the important factor. Third, the participant was presented with a practice block of the dual task (again nine trials per condition). After this practice block, the actual experiment started. Two filler trials preceded the experimental block and were excluded from analyses.

3.2. Results

One participant was excluded from the analyses because of an excessive error rate (52% erroneous trials). Trials on which participants responded to the picture–word stimulus before

responding to the tone were excluded from further analysis (2.4% of the trials). In addition, trials in which the participant's response on either the PWI stimulus or the tone was more than three standard deviations from the participant's mean per relatedness times SOA combination were excluded (2.0% and 2.0%, respectively). Because these exclusion criteria partly overlap, this resulted in exclusion of 5.2% of the trials. Following Dell'Acqua et al. (2007), no other trials were excluded.

The overall pattern in the data is presented in Fig. 7A. A linear-mixed effects model (Bates, 2005) was fit to the data to find the relative contribution of the factors to the response latency in the PWI task.³ Relatedness and SOA level, and the number of repetitions were included as fixed effects, together with the interactions between these factors. Participant and Picture were included as random effects, to account for intersubject and interitem variability. Markov-chain Monte Carlo (MCMC) sampling of the mixed-effects model (10,000 simulations; Baayen, Davidson, & Bates, 2008) showed that SOA, Repetition, and a two-way interaction between SOA and Relatedness significantly contributed to the variance in the data (Table 1). Also, the three-way interaction between SOA, Relatedness, and Repetition reached significance.

Further analysis of the effects per SOA level revealed that this three-way interaction is caused by the decreasing interference over repetitions at a SOA of 800 ms. This interaction

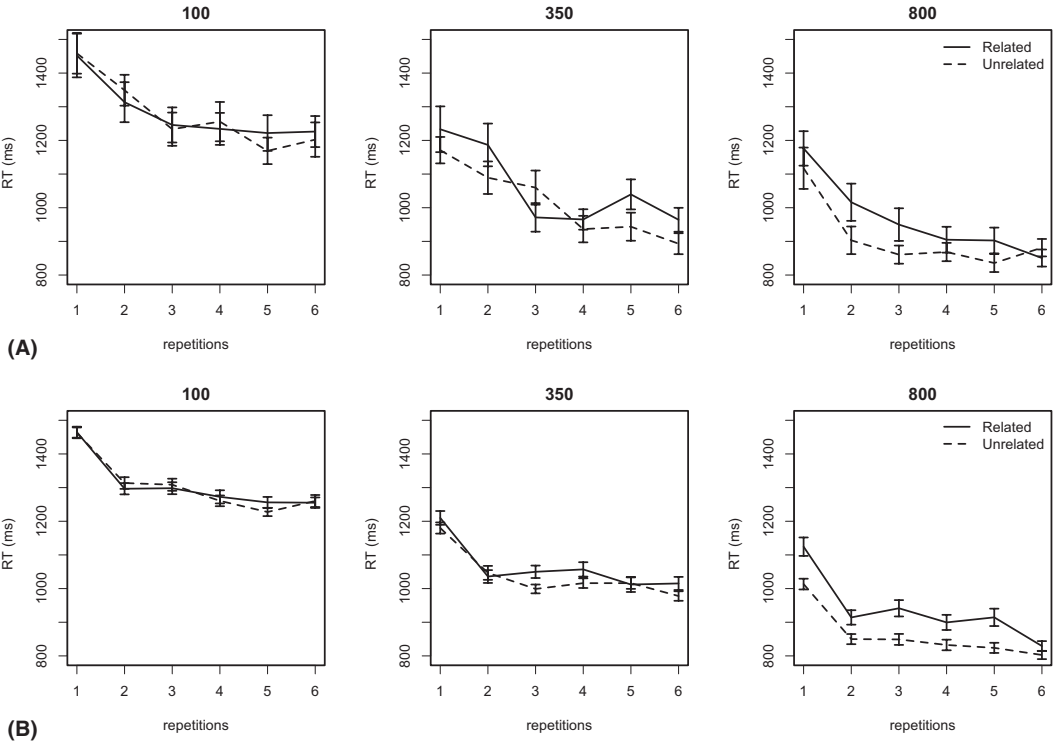


Fig. 7. Latency data (A) and model (B) from Experiment 1. Each panel represents the latency on each SOA level for all repetitions of the picture. SOA, stimulus onset asynchrony.

Table 1

Mean coefficient estimates and confidence intervals for the linear-mixed effects model of Experiment 1

Factor	Estimate	95% CI Lower	95% CI Upper	<i>p</i> -Value
(Intercept)	1452.74	1366.45	1541.04	<.001*
Relatedness	38.93	-42.71	121.94	.36
SOA	-0.41	-0.53	-0.29	<.001*
Repetition	-46.89	-62.49	-30.99	<.001*
Relatedness × SOA	-0.21	-0.38	-0.04	.014*
Relatedness × Repetition	-14.35	-35.64	8.32	.20
SOA × Repetition	-0.01	-0.04	0.02	.61
Relatedness × SOA × Repetition	0.04	0.00	0.09	.040*
Random effects				
Picture	86.04	66.88	106.68	
Participant	140.31	99.09	185.11	

Notes. * = $p < 0.05$

SOA, stimulus onset asynchrony.

was not significant at other SOAs (Table 2). To summarize, this experiment replicates the typical PRP effect, in which response latencies on the second task are negatively correlated with SOA, the typical effect of interference when the distractor word is related but not identical to the to-be-named picture, and, additionally, a relatively straightforward effect of repetition priming. The interaction indicates that at the 800 ms SOA, the repetition effect was smaller for trials with related distractors as compared to the 100 ms SOA trials.

3.3. Discussion

The difference in the response time latencies reflects the difference between the related and the unrelated condition: the interference effect. Since in the related condition picture and word are semantically related (but not in the unrelated condition), the interference may be an effect of competition between semantically related concepts. In both conditions, the word and picture activate a conceptual representation, but in the related condition, it is harder to decide on the correct conceptual representation of the picture (e.g., Glaser & Dünghoff, 1984; Van Maanen & Van Rijn, 2007a).

The size of the interference effect decreased with decreasing SOA. This observation suggests that the competition between the word and the picture is resolved early in the mental processing stream. If the interference effect under single-task conditions is located early in the process (for instance, during the visual processing of the picture), then this effect will be absorbed in the cognitive slack that is created by the PRP paradigm. Therefore, the presence of cognitive slack decreases the overt interference effect (Dell'Acqua et al., 2007; Fig. 6). On longer SOAs the slack time is decreased. Consequently, any additional processing that might have been resolved in the slack time is now observable. Therefore, the PWI effect reappears at longer SOAs. This effect replicates the findings of Dell'Acqua et al. (2007).

To study the details of the interactions between the effects identified earlier (repetition, interference, PRP), a RACE/A model is required. All three panels of Fig. 7A show latency

Table 2

Mean coefficient estimates and confidence intervals for the linear-mixed effects model of Experiment 1, separated per SOA level

Factor	Estimate	95% CI Lower	95% CI Upper	<i>p</i> -Value
SOA = 100 ms				
(Intercept)	1430.52	1326.77	1529.38	<.001*
Relatedness	24.11	-61.40	117.53	.59
Repetition	-40.00	-57.92	-22.53	<.001*
Relatedness × Repetition	-9.26	-33.38	14.97	.44
Random effects				
Picture	92.79	64.85	122.79	
Participant	175.22	125.50	233.94	
SOA = 350 ms				
(Intercept)	1257.57	1171.96	1342.83	<.001*
Relatedness	-51.87	-134.64	30.50	.22
Repetition	-55.98	-70.68	-41.54	<.001*
Relatedness × Repetition	2.83	-18.12	24.35	.79
Random effects				
Picture	88.79	62.43	117.70	
Participant	141.27	100.59	190.06	
SOA = 800 ms				
(Intercept)	1148.65	1070.51	1220.43	<.001*
Relatedness	-132.09	-213.41	-57.91	.002*
Repetition	-53.54	-66.68	-40.29	<.001*
Relatedness × Repetition	22.04	3.57	41.21	.02*
Random effects				
Picture	93.15	68.96	118.49	
Participant	110.90	74.96	149.36	

Note. SOA, stimulus onset asynchrony.

curves very similar to the power law of learning (Newell & Rosenbloom, 1981), indicating that participants become faster with more repetitions. This is in line with the idea that at each repetition, the conceptual representation of the picture is retrieved from memory, strengthening its memory trace, and making it easier to retrieve the concept at the next presentation. The effect of repetition on interference can be modeled by incorporating the long-term declarative memory mechanisms of RACE/A. In addition, the interference effects can be modeled using the short-term activation dynamics in RACE/A. The PRP effect can be accounted for by a particular specification of production rules.

3.4. Model of Experiment 1

The RACE/A model of Experiment 1 consists of two separate models: one model for the PWI task and one model for the tone-classification task.⁴ The model of the tone-classification task consists of two production rules: Rule 1 executes when a tone is detected and stores information about the tone in the aural buffer (using the standard ACT-R theory on auditory processing). Rule 2 maps the tone information onto a motor response and executes that response.

The model of the PWI task assumes that the task consists of three stages: a perceptual stage, a decision stage, and a response stage (e.g., Dell'Acqua et al., 2007; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Van Maanen et al., 2009). When the PWI stimulus appears on the screen, a production rule fires to retrieve conceptual information (referred to as a concept chunk) related to the stimulus, and stores this in the visual buffer. This process is governed by the short-term dynamics of RACE/A. The concept chunks can be regarded as representations of semantic properties of a certain concept. Even though this stage comprises more than pure perceptual processing, we will refer to this part of the task as the perceptual stage. In the decision stage, the model tries to retrieve a lemma chunk that relates to the concept chunk that was stored in the visual buffer. Lemma chunks can be regarded as sets of orthographic and syntactic properties of a word. When an appropriate lemma chunk has been retrieved, the model continues with the response stage, in which it first retrieves a morphophonological chunk containing information on articulatory properties of the response, and then initiates a motor program to execute a vocal response.

Two stimuli are presented simultaneously to the model, a picture and a word. When a stimulus is presented, it spreads activation to related chunks in declarative memory (following Eq. 4). To represent more prior practice with processing words than with processing pictures (e.g., Cohen, Dunbar, & McClelland, 1990; MacLeod & Dunbar, 1988), spreading activation from the word stimuli is higher than from pictures. This means that E_{word} in Eq. 4 is set at a higher value than E_{picture} (see Table 3 for the parameter values).

Text stimuli engages directly on the lemma chunks, while for pictures, conceptual information needs to be retrieved before the lemma associated with an image can be retrieved. This reflects the finding that words can be pronounced without a conceptual level (Glaser & Glaser, 1989; La Heij, Happel, & Mulder, 1990). By contrast, activation of the depicted concept is necessary for naming an image.

Because lemmas spread activation to the conceptual representations that relate to them (Fig. 8), the presentation of a distractor word causes interference at the conceptual level,

Table 3

Model parameters. Parameters of Model 2 are only displayed when they differ from those in Model 1. Default ACT-R parameters (other than those related to activation) are not included

Parameter	Model 1	Model 2	Description
α	0.65		Short-term decay
β	0.7		Spreading activation scale
γ	0.7		External activation scale
θ	0.975		Decision criterion
d	0.5		Long-term decay
ε	0.43		Moment-to-moment noise
E_{word}	1.375	3.5	External activation from words
E_{picture}	0.55	1.4	External activation from pictures
	2500 ms		Retrieval duration threshold mean
	440 ms		Retrieval duration threshold scale
	45%		Skip concept-lemma check

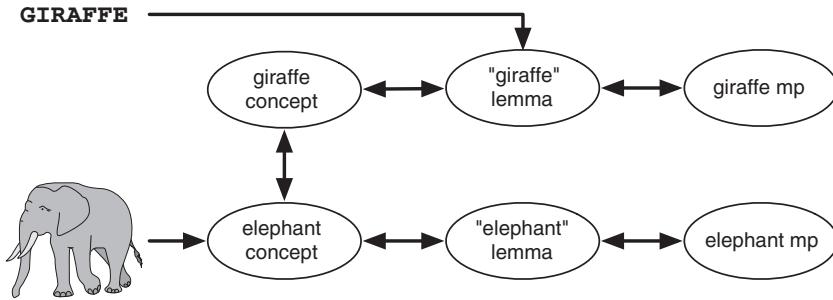


Fig. 8. Connections between different chunk types in the PWI model for an example stimulus. Note that in the experiments the word and picture dimensions of the stimulus overlap. mp, morphophonological representation; PWI, picture–word interference.

decreasing the concept's posterior probability. Likewise, spreading activation from the concept chunks to the lemma chunks causes interference at the lemma level. This reflects the principle in RACE/A that competition is not localized in a particular stage but may be distributed over all stages in which declarative memory is engaged. The competition in the model is distributed over the different subsets of chunks, but in different magnitudes.

The different probabilities determine the latency difference between the related and unrelated PWI conditions. In the related condition, the concepts of the target and the distractor spread activation to each other, decreasing the target's probability. This mutual excitation is not present in the unrelated condition, resulting in less competition, and thus a faster retrieval. These dynamics are also apparent from the simple simulations presented in Figs. 3a and 4a, which reflect the competition in the unrelated and related condition, respectively.

Once a concept has been retrieved, the model initiates a response. First it retrieves a lemma representation that encodes the syntactic information associated with the desired response; then, it retrieves a motor program to articulate the desired response. After retrieval of the lemma, the model checks if the retrieved lemma matches the information from the initial retrieval (i.e., the conceptual information). If not, the model retries to retrieve the correct lemma, excluding the just retrieved information from the retrieval set. However, this concept-lemma check is not always performed as in more typical picture-naming behavior, when no distractor is present; thus, it is highly unlikely that an incorrect lemma is retrieved. In our model, we estimated that in 45% of all trials no concept-lemma check is performed. Note that not performing a check does not always result in an incorrect trial, as the model may have retrieved the intended information. Typically, errors due to omitting the concept-lemma check are relatively fast. The second source of error reflects slow errors. If the duration of a memory retrieval crosses a certain threshold (the retrieval duration threshold), a retrieval error is signaled, and the model reports an incorrect trial. This type of error may reflect response time contaminants (Ratcliff & Tuerlinckx, 2002).

The retrieval duration threshold was sampled at each trial from a logistic distribution with mean 2,500 ms and scale 440 ms. The retrieval error represents the second source of error for the model.

Because of the task instruction to withhold the vocal PWI response until the tone classification is made, the model includes a control state that insures that the retrieval of the response lemma (as part of the PWI task) does not start until motor preparation for the button press (as part of the tone-classification task) has been initiated. As soon as the tone-classification response has been initiated, the PWI mechanisms can access declarative memory again, so the PWI task can continue before the actual button press is made.

Each time a chunk has been retrieved from memory, its base-level activation is updated because it has gained an additional reference. This means that the starting point in the accumulation process differs during the next retrieval for which that chunk is a competitor. Thus, at each repetition of a picture, the starting point for the associated conceptual representation is higher than on the previous repetition. A higher starting point results in a faster retrieval of the relevant concept, and therefore a faster response. Moreover, a fast retrieval results in little time for the other—distractor-related—chunks to accumulate activation, resulting in a smaller interference effect. This accounts for the repetition priming effect observed in Experiment 1.

3.4.1. Model results and discussion

To fit the model to the data, we first created a model using the parameter values that are considered the default in the ACT-R architecture. Next, we manually adjusted the additional RACE/A parameters until we obtained a model fit that was close to the data.

Similar to the experiment, we excluded outliers that were more than three standard deviations of the mean of each SOA-Relatedness combination. In the model, these outliers represent cases in which the model is unable to come to a decision due to multiple chunks accumulating at the same rate. When the model reaches a deadline, the accumulation is halted and the model retries to retrieve the desired chunk, resulting in prolonged-response latencies.

Fig. 7B shows that the model accounts for the repetition effect (RMSE = 56 ms). As the number of repetitions increases, the responses become faster. In the model, strengthening the memory trace of conceptual information as well as information at response levels causes the repetition priming effect. Previous studies have provided evidence for the existence of this dual process. For example, bilinguals show repetition priming for concepts that are presented in one language on the first presentation and in another language on the next presentation (e.g., Francis, Augustini, & Saenz, 2003; Francis & Sáenz, 2007). Thus, a repetition benefit was present in the absence of a repetition of responses. This is evidence that conceptual information is being reinforced on the first presentation, enabling a faster response on the second presentation. On the other hand, certain lexical decision studies demonstrate that non-word response latencies are also decreased by repetition, indicating that repetition priming also has a speed-up effect if no conceptual representation is present, as is generally believed to be the case for non-words in lexical decision experiments (e.g., Wagenmakers, Zeelenberg, Steyvers, Shiffrin, & Raaijmakers, 2004; Zeelenberg, Wagenmakers, & Shiffrin, 2004).

In addition to a model fit for the correct trials, the model also provides error rates (Fig. 9). The model assumes two sources of error to account for these responses. If the duration of a memory retrieval crosses a certain threshold, a retrieval error is signaled, and the model reports an incorrect trial. We assume that these trials reflect response time contaminants (Ratcliff & Tuerlinckx, 2002) in which participants make a guess or generally resort

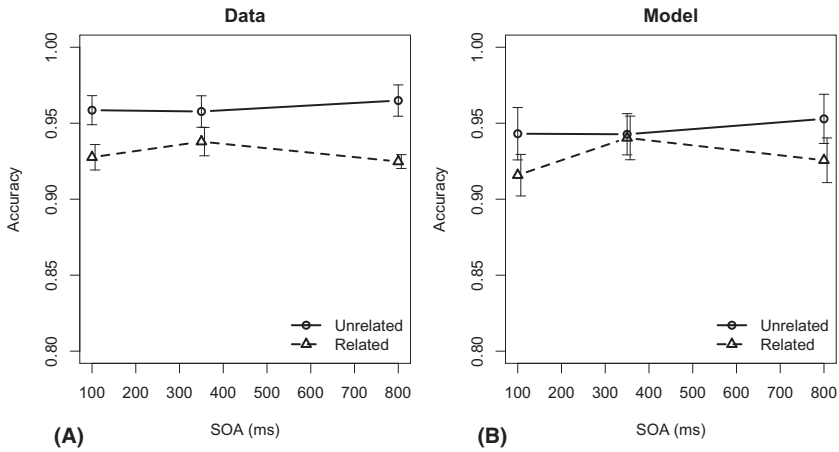


Fig. 9. Accuracy data (A) and model (B) of Experiment 1, aggregated over stimulus repetitions. SOA, stimulus onset asynchrony.

to a different strategy. On average, this leads to a latency of 2,091 ms and is committed on 1.1% of the trials (Type I errors in Fig. 10). The second source of error is generally fast (899 ms on average) and occurs on 4.8% of the trials, caused by an incorrect lemma retrieval in those trials in which no concept-lemma check is performed (Type II errors). The combination of these two types of error predicts similar error rates and error latencies as observed in the data (Figs. 7 and 8).

An important question in the context of this article is whether both the cognitive architecture and the sequential sampling model are necessary to account for the data. The model of this experiment strongly suggests that this question has to be answered affirmatively with

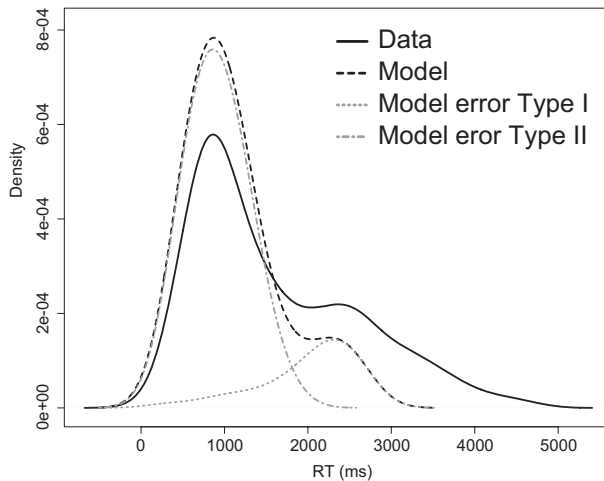


Fig. 10. Latency distribution for incorrect responses for Experiment 1 and the model. The gray lines indicate the distributions of the two error types predicted by the model.

respect to the cognitive architecture. Although the error pattern observed in Experiment 1 can be easily accounted for assuming a single process, it is necessary to schedule multiple sequential sampling processes in a particular scheme to model the mean response time data. This is because the start point of accumulation is determined by the interval between a previous sampling process and the onset of the current process. While admittedly such a scheduling mechanism may be added to any sequential sampling theory, the theoretical implications of chaining a series of accumulation processes would constitute a cognitive architecture in its own right. Therefore, instead of coming up with new architectures, it is worthwhile to test whether well-established cognitive architectures can perform this function.

In addition, the decrease of the interference effect observed in the data at the longest SOA is accounted for by strengthening the activation of declarative chunks at each encounter. This leads to higher start points at successive retrieval processes, a decrease in retrieval time, and less interference. The activation values are computed with the rational analysis of memory incorporated by the ACT-R cognitive architecture. Again, similar mechanisms may be added to sequential sampling theories if desired, but the current approach combines the strengths of a long-standing tradition of a rational analysis of memory incorporated in a cognitive architecture with precise sequential sampling models.

It is conceivable that a model can be developed that accounts for the data using a completely different explanation. For example, an architectural model of the Stroop task has been proposed that hypothesizes that Stroop effects occur because of an unsuccessful retrieval of the color information and a subsequent retry (Altmann & Davidson, 2001). However, by incorporating competitive mechanisms in our model we align with earlier research into the nature of interference processes (e.g., Cohen et al., 1990; Roelofs, 1992).

Experiment 1 and Model 1 illustrate how RACE/A naturally interacts with an important intrinsic property of the cognitive architecture ACT-R. The theory of long-term learning and forgetting that was already present in the architecture was extended with a theory on short-term dynamics of memory retrieval. This way, the interaction between (long-term) repetition priming and (short-term) interference could be explained. Also, the model makes use of the timing of the different subprocesses that are controlled by the production rule system. Thus, the model makes use of the possibility to formulate a theory on task execution in terms of the cognitive architecture.

However, one may be able to account for the same pattern in the data with a sequential sampling model in which separate starting point parameters are estimated for each repetition. RACE/A, however, provides a principled estimate of the starting points, whereas a pure sequential sampling model would leave them as free parameters.

4. Experiment 2: Cognitive control in a dual task

An important assumption in the model of Experiment 1 is that participants do not attend the word while naming the picture. That is, although the word interferes with the retrieval of the picture-related concepts and lemmas, the word is not fully processed. In the model this is implemented by a task strategy that only attempts to retrieve the concept related to the

picture (but at the same time, the word spreads activation to related chunks). The amount of spreading activation from the word stimulus is controlled by a parameter E_{word} that determines the drift rate of the associated lemma (together with spreading activation from the associated concept).

Experiment 2 was designed to influence the amount of attention that was given to the word. This way, we manipulated the amount of spreading activation from the word stimulus. We achieved this by adding an extra condition in which the word and the picture designate the same conceptual information. This condition will be referred to as the congruent condition. It has been hypothesized that including congruent trials in Stroop-like interference tasks increases the latency difference between related- and unrelated-target distractor pairs due to a decreased suppression of the automatic reading response (Logan & Zbrodoff, 1979). This effect has been attributed to the amount of processing of the irrelevant stimulus dimension (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Van Maanen & Van Rijn, 2010). In terms of the model, this would be reflected by a manipulation of the E_{word} parameter setting. Because—as in the model of Experiment 1—the model will have multiple dependent accumulating processes, changing the spreading activation from the word not only influences the magnitude of the interference but also the way the accumulating processes affect each other.

4.1. Methods

4.1.1. Participants

Twenty-two students of the University of Groningen (mean age 22.2, 14 male, 8 female) took part in this experiment for course credit. All were native speakers of Dutch and had normal hearing and normal or corrected-to-normal vision. The participants that took part in Experiment 1 were excluded from participation in Experiment 2.

4.1.2. Stimuli

The stimuli were the same as in Experiment 1, except that for the congruent condition, a third set of PWI stimuli was created. The distractor words for these stimuli were the Dutch image descriptors.

4.1.3. Design

The design was the same as in Experiment 1.

4.1.4. Procedure

The procedure was the same as in Experiment 1. Due to the length of the experimental block (441 trials and three filler trials), the participants were allowed three breaks, after 25%, 50%, and 75% of the trials.

4.2. Results

Again, we excluded trials according to the following criteria: Responses that were more than three standard deviations from a participants' mean were excluded (2.1% on the PWI

stimulus, and 2.3% on the tone, respectively). Trials in which the responses were in the incorrect order were also excluded (5.3%). Overall, 7.7% of the trials were excluded. The data of Experiment 2 are presented in Fig. 11A. We fitted a linear-mixed effects model with Relatedness, SOA, and Repetition as fixed effects and Participant and Picture as random effects. MCMC simulations on the factors revealed that there were main effects of SOA level, Relatedness, and Repetition (Table 4). In addition, all interactions were significant except the SOA times Repetition interaction.

In contrast to Experiment 1, the interaction between SOA and Relatedness does not seem to be due to the cognitive slack time induced by the PRP effect. Post hoc *t* tests show that response times differ between Relatedness conditions on both the shortest and the longest SOA (all $p < .003$). Only at the middle SOA of 350 ms we did not find a difference between

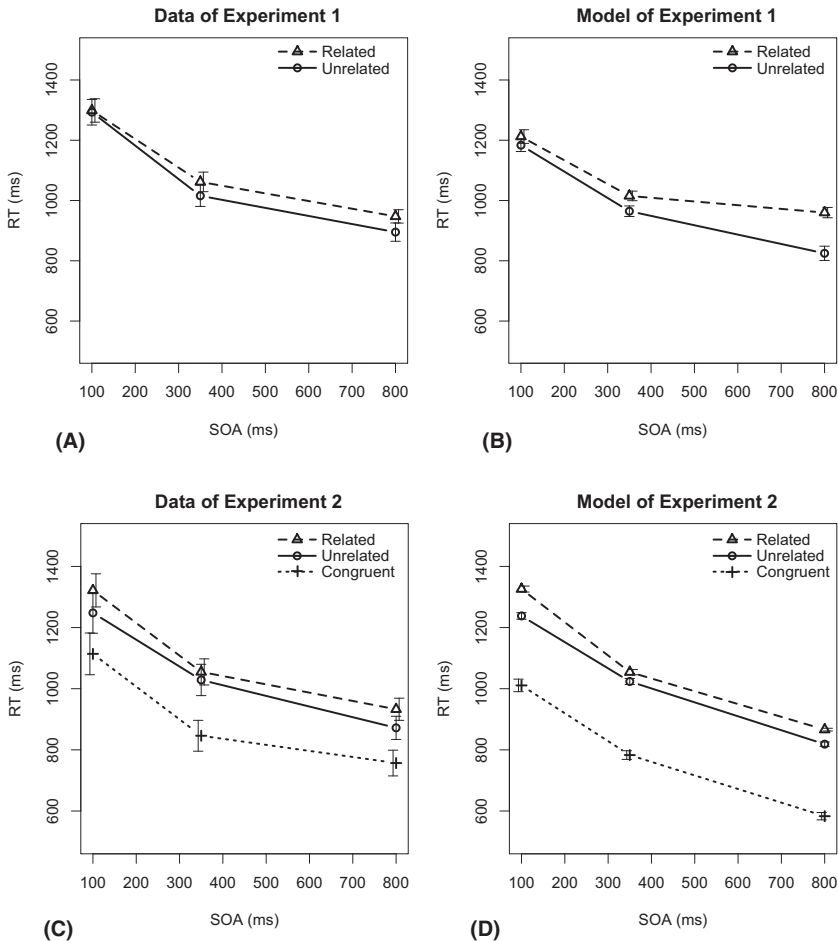


Fig. 11. Data and models for Experiments 1 and 2. (A) Data of Experiment 1, collapsed over stimulus repetitions for comparison. (B) Model of Experiment 1, collapsed over stimulus repetitions. (C) Data of Experiment 2. (D) Model of Experiment 2. SOA, stimulus onset asynchrony.

Table 4

Mean coefficient estimates and confidence intervals for the linear-mixed effects model of Experiment 2

Factor	Estimate	95% CI Lower	95% CI Upper	<i>p</i> -Value
(Intercept)	1208.80	1112.61	1313.96	<.001*
Related	409.41	332.20	486.98	<.001*
Unrelated	237.53	157.96	317.23	<.001*
SOA	-0.46	-0.56	-0.36	<.001*
Repetition	-21.85	-33.81	-10.79	<.001*
Related × SOA	-0.31	-0.47	-0.16	<.001*
Unrelated × SOA	-0.17	-0.34	-0.02	.032*
Related × Repetition	-43.84	-59.15	-27.89	<.001*
Unrelated × Repetition	-16.64	-32.20	-1.32	.034*
SOA × Repetition	-0.01	-0.03	0.02	.58
Related × SOA × Repetition	0.06	0.03	0.09	<.001*
Unrelated × SOA × Repetition	0.04	0.00	0.06	.022*
Random effects				
Picture	64.31	48.34	78.85	
Participant	195.28	147.24	247.42	

Note. SOA, stimulus onset asynchrony.

Relatedness conditions (the Related and Unrelated condition did not differ). This is in contrast with Experiment 1, in which the Related and Unrelated condition did not differ at the shortest SOA (post hoc *t* test: $p = .8$) but did differ at the longest SOA (post hoc *t* test: $p = .003$).

4.3. Discussion

The typical PRP speed-up is present in the data of Experiment 2, suggesting a correct operationalization of the PRP design and indicating the presence of cognitive slack time. Therefore, the observation that in the presence of congruent trials the mean response latencies per relatedness condition differ at the shortest SOA (100 ms) suggests that the interference between distractor word and picture was longer than could be absorbed in the maximum cognitive slack time. This is different from the result of Experiment 1 in which the interference could be absorbed in the cognitive slack time. In line with our operationalization of Experiment 2, we hypothesize that participants process the word more often or to a higher level than in Experiment 1. As a result, the competition between conceptual representations will be stronger and hence the interference of the related words will be higher. However, it cannot simply be the case that the interference in the first stage of the process becomes larger. That would mean that for the long SOA level of 800 ms, the interference effect would become even larger, because with an SOA of 800 ms there is no cognitive slack time in which part of the interference can be resolved. Since the data show no interaction between interference and SOA, this simplest explanation does not seem likely. An explanation for this observation might be that the interference becomes more distributed over different stages of the task. This means that the extended processing of the word does

not only cause interference during the initial conceptual processing of the picture but also during response stages. This explanation is in line with cascading models (Levelt et al., 1999; McClelland, 1979) in which activation from different stages spreads to each other, affecting processing later in the task.

In the current experiment, this explanation would mean that due to cascading activation the interference is partly located at later (response) stages. Following the PRP logic, this part of the interference effect would not be resolved in the cognitive slack time because it does not commence until after the central resource (recall Fig. 6). RACE/A offers an implementation of this idea if we assume that the difference between Experiments 1 and 2 can be expressed by the level of processing of the word. We hypothesize that the presence of congruent trials in Experiment 2 leads to less suppression of the automatic reading response and a higher level of processing of the word dimension of the stimulus.

4.4. Model of Experiment 2

As in Experiment 2, the focus of Model 2 is on the effect of presence or absence of congruent trials. For this reason we now ignore the effects of repetition that were relevant in Experiment 1 and Model 1. The model performs the same task as in Experiment 1, but we assumed that participants in Experiment 2 process the word to a higher extent than in Experiment 1. To account for this, we adapted the parameters that control spreading activation from the stimuli (Table 3). All other parameters were kept constant.

If word and picture are congruent, the new values for spreading activation from the stimulus mean that the activation of the lemma associated with the picture is high when the model attempts a lemma retrieval. This retrieval will therefore be faster than in the model of Experiment 1. The model can now directly continue with retrieving the chunks that are required for response execution. This accounts for faster responses in the congruent condition. However, if word and picture are not congruent (i.e., in the related or unrelated condition), the model initially retrieves the incorrect lemma. The reason for this is that the incorrect lemma's activation was increased by the presentation of the word, which spreads more activation than the picture. Thus, the incorrect lemma is much more activated than the correct lemma, which represents the picture. The model will now check if the retrieved lemma matches the information that initiated the retrieval. If not, the model retries to retrieve the correct lemma, this time excluding the previously retrieved lemma from the retrieval set. However, the incorrect lemma still spreads activation—mediated by the concept chunks—to other lemmas, increasing the relative decision criterion. Therefore, retrieval in the related condition is slower than in the unrelated condition. We hypothesized that the presence of congruent trials in Experiment 2 decreased the suppression of the automatic reading response relative to Experiment 1. In the model this was represented by a higher value for the E_{word} parameter, reflecting more spreading activation from the word. With this adjustment the model does not show an interaction between SOA and Relatedness. Thus, extra processing of the word may cause more interference, which therefore may not be fully absorbed in the cognitive slack time, resulting in an overt latency difference at short SOAs.

4.4.1. Model results and discussion

Figs. 11–13 present the fit of the model to Experiment 2. Similar to the empirical data, the model shows an increased interference effect at a SOA of 100 ms as compared to Experiment 1. Also, the model shows the standard PRP effect (Fig. 11, RMSE = 56 ms). The model follows the error rates in the data as well (Fig. 12), which do not depend on SOA. Similar to the experimental data, the model shows a bimodal latency distribution of the incorrect trials (Fig. 13). As for Experiment 1, the model error pattern consists of two types of errors. Fast errors in this model are estimated to have a mean latency of 626 ms and are committed in 3.0% of the trials. The slow errors are committed in 1.8% of the trials and have an average latency of 2,100 ms. The overall error rate in this model is slightly lower than the error rate in the model of Experiment 1, which is caused by the underestimation of the errors in the Congruent condition. The model does not make the fast Type II errors in the congruent condition, because that would require the retrieval of a lemma chunk that is neither related to the concept chunk, nor to the word dimension of the stimulus. The incorrect retrieval will then be followed by a failure to check the correctness of the lemma. The likelihood of this combination is very low. In addition, the model also never reaches the deadline in the congruent condition, because all evidence amasses to the same chunk, resulting in fast retrievals.

5. General discussion

In this article we have presented a modeling framework that combines insights from sequential sampling models of reaction time with insights from cognitive architectures. We have demonstrated with two experiments and two models how modeling the specific time course of a task enables us to estimate the specific onsets and offsets of sequential sampling processes.

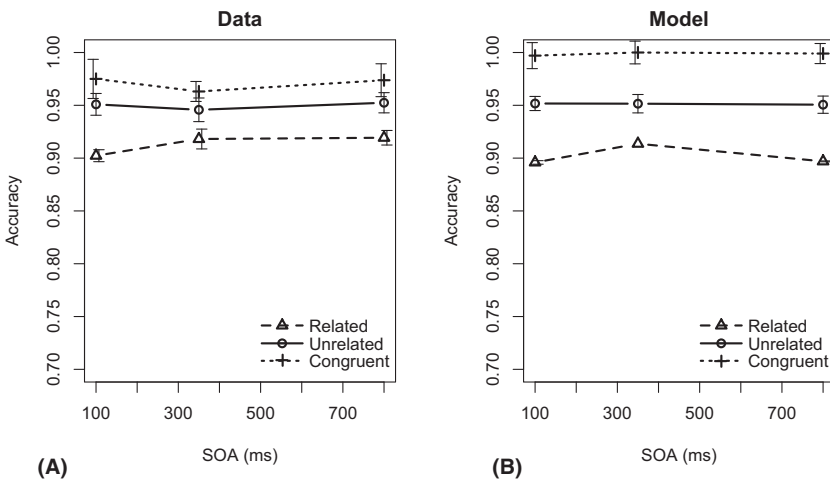


Fig. 12. Accuracy data (A) and model (B) of Experiment 2. SOA, stimulus onset asynchrony.

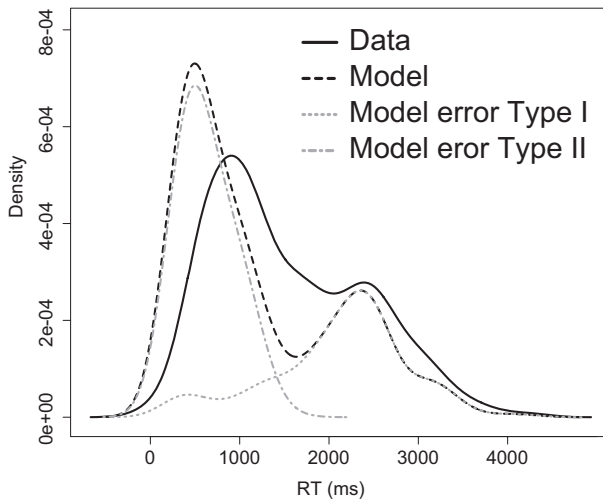


Fig. 13. Latency distribution for incorrect responses for Experiment 2 and the model. The gray lines indicate the distributions of the two error types predicted by the model.

Experiment 1 and the RACE/A model of Experiment 1 showed how short-term dynamics of activation can be combined with long-term activation dynamics to account for the interaction between interference and repetition priming in PWI. In addition, we showed how these mechanisms can be embedded in a production rule system so as to account for PRP effects. For long-term activation dynamics this approach is far from new, but the integration of short-term activation dynamics and a cognitive architecture is novel.

Experiment 2 and the RACE/A model of Experiment 2 show that RACE/A may provide an explanation for interference effects caused by cognitive control mechanisms thought to act on the suppression of an irrelevant stimulus dimension. In addition, the RACE/A models of Experiment 1 and Experiment 2 provide an explanation for the additive interference effects observed in Experiment 2, which are not present in Experiment 1. Less suppression of the word may lead to more interference that can no longer be absorbed in the cognitive slack time.

5.1. Related work

Previous work that integrates sequential sampling models and high-level explanations of cognitive behavior include Brown, Marley, Donkin, and Heathcote (2008), Kent and Lamberts (2005), Lee and colleagues (Lee & Corlett, 2003; Lee & Cummins, 2004; Lee & Dry, 2006), Nosofsky and Palmeri (1997), and Smith and Ratcliff (2009). These approaches are characterized by a sampling model in which the drift rate is determined by psychological concepts, rather than a free parameter as in the classical sequential sampling models. For example, the exemplar-based random walk model (EBRW; Nosofsky & Palmeri, 1997) models classification as a random walk process. Here, the similarity between the target and the exemplars of the category that are stored in memory determines how the model steps

toward the response options (the categories). In a similar vein, Lee and Corlett (2003) propose a sequential sampling model approach to classify sentences by topic. Here, the drift depends on the log odds that a word in a sentence is exemplary of a particular topic and has occurred in previous topical sentences before (cf. Anderson & Schooler, 1991). If the words are highly exemplary for a topic, the drift is also high and the model decides quickly and accurately. If the words are not topic exemplars, the model will reject that topic as the topic of the current sentence. This approach is similar to RACE/A because it also considers how often a particular item (or chunk, in RACE/A terminology) has been encountered. RACE/A, however, also considers that more than one sequential sampling process may be executed during task execution, and it provides a theory on the scheduling of these individual processes.

Another integrated approach is the SAMBA model for absolute identification (Brown et al., 2008). Accumulation in SAMBA is driven by features of the stimulus. The magnitude of the stimulus is compared to the magnitudes of the upper and lower boundaries of the stimulus range, and the resulting relative value determines the drift rate. In addition to the stimulus-dependent drift rate, SAMBA also includes decay of the accumulation after a response, to account for sequential or between-trial effects. We take this idea one step further by not only providing principled estimates for the starting point and drift rate parameters of the accumulation process but also integrating the accumulation process in a full-fledged architecture of cognition. This way we can model interactions of multiple memory retrievals, multiple learning mechanisms, and even multiple tasks.

5.2. Differences with ACT-R

RACE/A differs from default ACT-R in a number of ways. However, these adaptations do not change the central assumptions (Cooper, 2007) that underlie ACT-R. The first obvious deviation from the default architecture is the inclusion of an accumulative process for memory retrievals. This process replaces the, admittedly simpler, algebraic model. The new model increases the explanatory power of the architecture. However, in the absence of competition between chunks RACE/A makes the same predictions as ACT-R. This special case is also addressed by Anderson (2007, appendix 3.1), who compares ACT-R's latency predictions with a sequential sampling model (Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007) similar to RACE/A. Assuming no noise (and no competitors), the retrieval time in RACE/A can be described as a ballistic model (cf. Brown & Heathcote, 2008), in which the retrieval time is a function of the distance between the decision criterion and the posterior probability for the chunk. Fig. 14 presents the predicted retrieval time for various activation values. Especially for low activation values, RACE/A closely follows the ACT-R predictions of retrieval times.

Another important difference is that RACE/A assumes that all chunks in declarative memory spread activation. ACT-R has always implemented a more delimited account of spreading activation in which only those chunks that are currently attended (i.e., chunks that are in the buffers) spread activation. The justification for this change is two-fold. First, experimental evidence from priming studies suggest that priming behaves as a full spreading

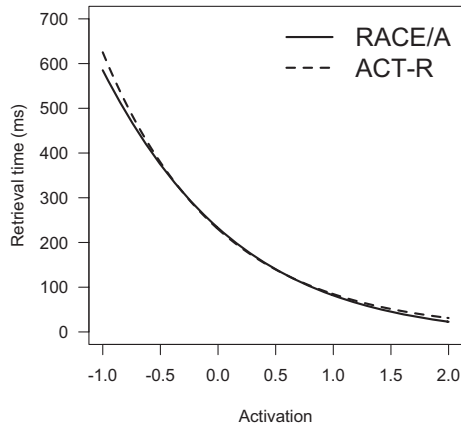


Fig. 14. RACE/A and standard ACT-R latency predictions for various activation levels. For RACE/A, these predictions address the special case when there are no competitors. RACE/A, Retrieval by Accumulating Evidence in an Architecture.

activation mechanism. For example, it has been shown that priming effects persist if multiple unrelated items are presented between prime and target, which is difficult to agree with the ACT-R account (Becker, Moscovitch, Behrmann, & Joordens, 1997; Joordens & Besner, 1992). Second, by assuming a full spreading activation mechanism, we do not have to make the assumption of a dedicated set of chunks that have a privileged status in that they spread activation (and others do not). We feel that the increased emphasis on detailed computation of activation in RACE/A calls for a more detailed account of spreading activation than in default ACT-R.

A third difference between ACT-R and RACE/A is the way in which perceptual encoding is treated. ACT-R traditionally models this by a fixed constant interval for perceptual processes that are required to identify the perceived information (85 ms, Anderson, 2007). RACE/A treats this as a special case of memory retrieval. We assume that features of the stimulus become available during the first stages of perceptual encoding (cf. feature integration theory, Treisman & Gelade, 1980). These features are associated with declarative information that consecutively competes for retrieval. The winning information is then stored in the visual buffer (instead of the retrieval buffer). So, instead of a fixed duration of perceptual encoding, we have implemented a variable encoding time based on features of the stimulus and a selection process.

5.3. Future research

In this article we discussed two settings in which the need for a combined approach such as RACE/A is eminent. However, we did not applied RACE/A to its full potential. One of the successes of the sequential sampling framework is the ability to account for response time distributions (for both correct responses and error responses), instead of only means. For example, the drift rate and the decision criterion of a sequential sampling model have

different effects on the shape of response time distributions. The exact effects of these parameters have been an active research topic in the sequential sampling community (e.g., Dutilh et al., 2009; Ratcliff & Tuerlinckx, 2002). The current work opens up the possibility to study particular response time distributions in complex tasks. One important aspect that will arise during this research is how sequences of accumulator models combine into one distribution, and how that distribution will change if characteristics of the task change. For example, RACE/A makes it possible to study how and why the RT distribution of Experiment 1 in this article changes as the SOA between the tasks increases.

5.4. Conclusion

Although we implemented RACE/A as an extension to ACT-R (which can be downloaded from <http://www.ai.rug.nl/~leendert/race-a>), it should be stressed that the main points of this article apply to other architectures as well. For example, the cognitive architecture Soar may be extended in a similar way. Soar is a cognitive architecture that also utilizes production rules for task control (Newell, 1990). In addition, Soar also contains a memory system for storing and retrieving facts (Laird, 2008).

We concede that many aspects of embedding accumulator models as we proposed remain to be studied. At the same time, we see great value in the current approach, as it has the potential to increase the power of models of decision making as well as the power of cognitive architectures.

Notes

1. In the simulations reported throughout this article, we used the discrete version of Eq. 4:

$$C_i(t + \Delta t) = (1 - \alpha)C_i(t) + \beta \sum_j S_{ji}A_j(t) + \gamma \sum_k S_{ki}E_k(t) + \varepsilon_i(t).$$

2. The method described by Dragalin et al. (1999) is in fact asymptotically optimal. The method is a generalization of the optimal solution that exists for decisions between two choices, and it approaches optimality for decreasing error rates. For two-alternative choices, RACE/A also simplifies to the optimal model.
3. R code for this linear-mixed effects model as well as the data can be downloaded from <http://csjarchive.cogsci.rpi.edu/Supplemental/>.
4. This model and subsequent models discussed in this article can be downloaded from <http://csjarchive.cogsci.rpi.edu/Supplemental/>.

Acknowledgments

This research was supported by Netherlands Organization of Scientific Research (NWO) ToKeN grant 634.000.002 and by US Office of Naval Research grant N00014-08-10541 and

US Air Force Office of Scientific Research grant FA95500710359 awarded to N. Taatgen. We thank Max Jensch for conducting the experiments; Jelmer Borst, Ronald van Elburg, Glenn Gunzelmann, and Joachim Vandekerckhove for comments on an earlier draft; and the Groningen Cognitive Modeling Group for fruitful discussions on the topics addressed in this article.

References

- Altmann, E. M., & Davidson, D. J. (2001). An integrative approach to Stroop: Combining a language model and a unified cognitive theory. In J. D. Moore & K. Stennings (Eds.), *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society* (pp. 21–26). Mahwah, NJ: Erlbaum.
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, *115*(3), 602–639.
- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*(4), 451–474.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.
- Anderson, J. R., Fincham, J. M., & Douglass, S. (1999). Practice and retention: A unifying analysis. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(5), 1120–1136.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*(4), 703–719.
- Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*(2), 186–197.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*(6), 396–408.
- Anderson, J. R., Taatgen, N. A., & Byrne, M. D. (2005). Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance*, *31*(4), 749–761.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.
- Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database (CD-ROM).
- Bates, D. M. (2005). Fitting linear mixed models in R. *R News*, *5*, 27–30.
- Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*(5), 1059–1082.
- Bogacz, R., & Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation*, *19*(2), 442–477.
- Borst, J. P., Taatgen, N. A., & Van Rijn, H. (2010). The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 363–382.
- 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.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178.
- Brown, S. D., Marley, A. A. J., Donkin, C., & Heathcote, A. (2008). An integrated model of choices and response times in absolute identification. *Psychological Review*, *115*(2), 396–425.

- Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2009). Observing evidence accumulation during multi-alternative decisions. *Journal of Mathematical Psychology*, 53(6), 453–462.
- Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review*, 108(4), 847–869.
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed-processing account of the Stroop effect. *Psychological Review*, 97(3), 332–361.
- Cooper, R. P. (2007). The role of falsification in the development of cognitive architectures: Insights from a Lakatosian analysis. *Cognitive Science*, 31(3), 509–533.
- Dell'Acqua, R., Job, R., Peressotti, F., & Pascali, A. (2007). The picture–word interference effect is not a Stroop effect. *Psychonomic Bulletin & Review*, 14(4), 717–722.
- Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods: Instruments and Computers*, 32(4), 588–615.
- Dragalin, V. P., Tartakovsky, A. G., & Veeravalli, V. V. (1999). Multihypothesis sequential probability ratio tests. I. Asymptotic optimality. *IEEE Transactions on Information Theory*, 45(7), 2448–2461.
- Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, 16(6), 1026–1036.
- Fleetwood, M. D., & Byrne, M. D. (2006). Modeling the visual search of displays: A revised ACT-R model of icon search based on eye tracking data. *Human-Computer Interaction*, 21(2), 153–198.
- Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E. J. (2008). Striatum and pre-SMA facilitate decision making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45), 17538–17542.
- Francis, W. S., Augustini, B. K., & Saenz, S. P. (2003). Repetition priming in picture naming and translation depends on shared processes and their difficulty: Evidence from Spanish-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1283–1297.
- Francis, W. S., & Sáenz, S. P. (2007). Repetition priming endurance in picture naming and translation: Contributions of component processes. *Memory & Cognition*, 35(3), 481–493.
- Glaser, W. R., & Dünghoff, F. J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Learning, Human Perception and Performance*, 10(5), 640–654.
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118(1), 13–42.
- Gray, W. D. (2007). Composition and control of integrated cognitive systems. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 3–12). New York: Oxford University Press.
- Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(3), 483–491.
- Kent, C., & Lamberts, K. (2005). An exemplar account of the bow and set-size effects in absolute identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 289–305.
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, 77(4), 576–588.
- La Heij, W., Happel, B., & Mulder, M. (1990). Components of Stroop-like interference in word reading. *Acta Psychologica*, 73(2), 115–129.
- Laird, J. E. (2008). Extending the Soar cognitive architecture. In P. Wang, B. Goertzel, & S. Franklin (Eds.), *Proceedings of the first conference on artificial general intelligence (AGI-08)* (pp. 224–235). Amsterdam: IOS Press.
- Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning*, 1(1), 11–46.
- Lee, M. D., & Corlett, E. Y. (2003). Sequential sampling models of human text classification. *Cognitive Science*, 27(2), 159–193.

- Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, *11*(2), 343–352.
- Lee, M. D., & Dry, M. J. (2006). Decision making and confidence given uncertain advice. *Cognitive Science*, *30*(6), 1081–1095.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527.
- Logan, G., & Zbrodoff, N. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory and Cognition*, *7*(3), 166–174.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
- MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning Memory and Cognition*, *14*(1), 126–135.
- Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, *15*(2), 197–237.
- McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 471–484.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*(4), 287–330.
- Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms. *Psychological Review*, *104*(1), 3–65.
- Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: II. Accounts of psychological refractory-period phenomena. *Psychological Review*, *104*(4), 749–791.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their interaction* (pp. 1–56). Hillsdale, NJ: Lawrence Erlbaum.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*(2), 266–300.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.
- Ratcliff, R., Hasegawa, Y. T., Hasegawa, R. P., Smith, P. L., & Segraves, M. A. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *Journal of Neurophysiology*, *97*(2), 1756–1774.
- Ratcliff, R., & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review*, *104*(2), 319–343.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333–367.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481.
- Rayner, K., & Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture word interference task. *British Journal of Psychology*, *77*(2), 207–222.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142.
- Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research*, *1*(4), 201–220.

- Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, 115(1), 101–130.
- Schneider, D. W., & Anderson, J. R. (2011). A memory-based model of Hick's law. *Cognitive Psychology*, 62, 193–222.
- Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*, 109(4), 373–392.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3), 161–168.
- Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, 116(2), 283–317.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276–315.
- Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say “Broke”? A model of learning the past tense without feedback. *Cognition*, 86(2), 123–155.
- Taatgen, N. A., Huss, D., Dickison, D., & Anderson, J. R. (2008). The acquisition of robust and flexible cognitive skills. *Journal of Experimental Psychology: General*, 137(3), 548–565.
- Taatgen, N. A., & Lee, F. J. (2003). Production compilation: A simple mechanism to model complex skill acquisition. *Human Factors*, 45(1), 61–76.
- Taatgen, N. A., Van Rijn, H., & Anderson, J. R. (2007). An integrated theory of prospective time interval estimation: The role of cognition, attention and learning. *Psychological Review*, 114(3), 577–598.
- Tamborello, F. P., & Byrne, M. D. (2007). Adaptive but non-optimal visual search behavior in highlighted displays. *Journal of Cognitive Systems Research*, 8(3), 182–191.
- Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, 14(1), 1–36.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592.
- Van Maanen, L., & Van Rijn, H. (2007a). An accumulator model of semantic interference. *Cognitive Systems Research*, 8(3), 174–181.
- Van Maanen, L., & Van Rijn, H. (2007b). Accounting for subliminal priming in ACT-R. In R. Lewis, T. A. Polk, & J. E. Laird (Eds.), *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 1–6). London: Taylor & Francis.
- Van Maanen, L., & Van Rijn, H. (2010). The locus of the Gratton effect in picture–word interference. *Topics in Cognitive Science*, 2(1), 168–180.
- Van Maanen, L., Van Rijn, H., & Borst, J. P. (2009). Stroop and picture–word interference are two sides of the same coin. *Psychonomic Bulletin & Review*, 16(6), 987–999.
- Van Rij, J., van Rijn, H., & Hendriks, P. (2010). Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language*, 37(3), 731–766.
- Van Rijn, H., & Anderson, J. R. (2003). Modeling lexical decision as ordinary retrieval. In F. Detje, D. Doerner, & H. Schaub (Eds.), *Proceedings of the Fifth International Conference on Cognitive Modeling* (pp. 207–212). Bamberg, Germany: Universitaetsverlag.
- Van Rijn, H., Van Someren, M., & van der Maas, H. L. J. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science*, 27(2), 227–257.
- Vickers, D. (1970). Evidence for an accumulator model for psychophysical discrimination. *Ergonomics*, 13(1), 37–58.
- Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, 58(1), 140–159.
- Wagenmakers, E.-J., Van der Maas, H. L. J., & Grasman, R. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14(1), 3–22.

- Wagenmakers, E.-J., Zeelenberg, R., Steyvers, M., Shiffrin, R., & Raaijmakers, J. (2004). Nonword repetition in lexical decision: Support for two opposing processes. *Quarterly Journal of Experimental Psychology A*, 57(7), 1191–1210.
- Wang, X. J. (2008). Decision making in recurrent neural circuits. *Neuron*, 60(2), 215–234.
- Welford, A. T. (1967). Single-channel operation in the brain. *Acta Psychologica*, 27, 5–22.
- Welford, A. T. (1980). The single-channel hypothesis. In A. T. Welford (Ed.), *Reaction time* (pp. 215–252). San Diego, CA: Academic Press.
- Wickelgren, W. (1977). Speed–accuracy tradeoff and information-processing dynamics. *Acta Psychologica*, 41(1), 67–85.
- Zeelenberg, R., Wagenmakers, E.-J., & Shiffrin, R. M. (2004). Nonword repetition priming in lexical decision reverses as a function of study task and speed stress. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30(1), 270–277.
- Zhang, J., & Bogacz, R. (2010). Optimal decision making on the basis of evidence represented in spike trains. *Neural Computation*, 22(5), 1113–1148.