Context effects on memory retrieval

Maanen, Leendert van

Publication date:
2009

Context Effects On Memory Retrieval:
Theory And Applications

This research was supported by Jurriaanse stichting.

This research was supported by BCN

This research was supported by the Netherlands Organisation for Scientific Research under project number 634.000.002

Netherlands Organisation for Scientific Research

ISBN: 978-90-367-3924-5

Cover and Layout: Maarten van Maanen

RIJKSUNIVERSITEIT GRONINGEN

Context Effects On Memory Retrieval:
Theory And Applications

PROEFSCHRIFT

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
maandag 14 september 2009
om 13.15 uur

door Leendert van Maanen
geboren op 6 mei 1978
te Leiden

Promotor:	Prof. dr. L.R.B. Schomaker
Copromotores:	Dr. D.H. van Rijn
	Prof. dr. N.A. Taatgen
Beoordelingscommissie:	Prof. dr. W.D. Gray
	Prof. dr. G.A.M. Kempen
	Prof. dr. J.G.W. Raaijmakers

Contents

Introduction

This thesis is about memory. More specifically, it is about memory for facts such as *Amsterdam is the capital of the Netherlands*, or *The object I am looking at is called a computer screen*. This type of memory will be referred to as *declarative memory*. Declarative memory is an interesting object of study because remembering facts is something humans continuously do; retrieving facts from declarative memory is at the core of cognitive functioning. For example, while typing this page I am continuously trying to remember words that are appropriate given the message I want to convey. At the same time I might remember that I have an appointment later on, and remember the name of the person I have the appointment with. Because declarative memory plays such a central role in human cognition, we can learn a great deal about human behavior if we understand the process by which declarative memory retrievals are driven. Therefore this thesis focuses at understanding the functional process of retrieving facts from declarative memory.

This thesis is also about *information selection*. Because of the over-abundance of information in modern society - caused by an increase in digital storage capacity coupled to an increased accessibility of digital sources - access to information has become cluttered (Brusilovsky & Tasso, 2004). Even though the likelihood that certain information is available is increased (because the amount of information is increased), finding the relevant information has become harder. Therefore, automated assistance to select relevant information out of the abundance of information sources seems a necessity. In a response to this demand, the research field of Information Retrieval has expanded rapidly in recent years (as witnessed for instance by the number of submissions to the annual international ACM SIGIR Conference on Research and Development in Information Retrieval, which is the most important conference in the Information Retrieval field. The number of submissions has expanded in the last decade from 135 in 1999 to 490 in 2007).

The approach we will take in this thesis is to study the similarities between declarative memory retrieval and information retrieval. Very superficially, declarative memory can be perceived as a storage capacity for information that has been encountered during a person's lifetime. Consequently, remembering this information (these declarative facts) for current usage may be analogous to selecting currently relevant information. If we assume that the human cognitive system, through evolution and learning, has developed an optimal solution for this information-selection problem we can develop new information selection algorithms and assist people in handling the information glut (Schenk, 1997) by mimicking the retrieval process of facts from declarative memory.

Multiple researchers have stressed the similarities between information selection and human declarative-memory retrieval. For instance, Anderson (e.g., Anderson, 1989; Anderson & Milson, 1989) demonstrated the parallels between human declarative memory and artificial information-retrieval systems as introduced by Salton and McGill (1983). Salton and McGill characterize information-retrieval systems as systems that (1) consist of sets of files that contain the to-be-retrieved information, (2) have some sort of index of terms that can be used to retrieve the files, and (3) have the possibility to query the system using a subset of these terms. Anderson and colleagues noted that the human declarative memory system shares these characteristics with artificial information-retrieval systems. That is, human memory (1) contains facts (or cognitive chunks, in his terminology), (2) contains relations between these chunks, for instance semantic relations, and (3) has the possibility to retrieve chunks, based on these relations between chunks.

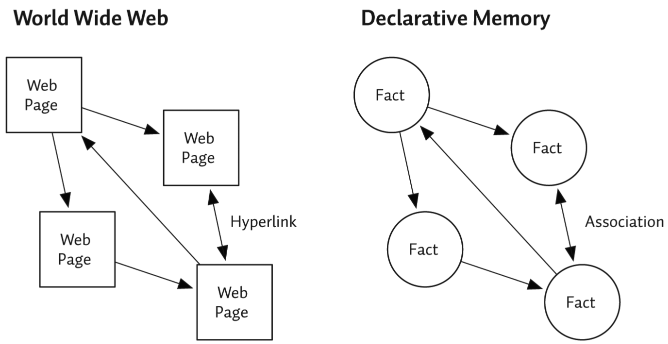


Figure 1.1. An illustration of the similar structure of the World Wide Web and declarative memory. The web consists of web pages that are connected via hyperlinks. In declarative memory facts are stored, that are connected via associations. Figure adapted from (Griffiths, Steyvers, & Firl, 2007).

More recently, Griffiths and colleagues (Griffiths, Steyvers, & Firl, 2007) showed that also search-engine technology parallels human memory behavior. They argue that, given the similar structure of the web and the mind (Figure 1.1), algorithms that are optimized for finding relevant web pages can best predict associations in human declarative memory. Indeed, they show that Google's PageRank algorithm (Page, Brin, Motwani, & Winograd, 1998) outperforms some other algorithms when predicting human responses on a fluency task. In this task, participants are presented with a letter and have to name the first word beginning with that letter that comes to mind. PageRank's predictions map better onto human data than simpler algorithms, specifically word frequency norms and word co-occurrence frequency norms. The ultimate suggestion of this result might be that the way PageRank calculates the relevance of words, given a certain letter, is functionally equivalent to the way human declarative memory "calculates" the relevance of facts in a certain context. This example thus shows that it is useful to study information retrieval algorithms if you are interested in human memory. In this thesis, we will argue that also the opposite direction, that is, applying knowledge of the functional process of human declarative memory to information access, will be beneficial for information retrieval (Part II).

The work reported in this thesis may be separated into two distinct topics. In Part I (Theory) we will study the functional process of declarative memory retrieval, and in Part II (Applications) we will study how functional models of declarative memory retrieval may be applied in information retrieval systems. However, despite this apparent dichotomy both topics involve the concept of declarative memory retrieval, which is the unifying concept for this thesis. Also, both in Part I and in Part II we will argue that declarative memory retrievals are mediated by changes in the environment. That is to say, which declarative fact is retrieved from memory is determined by the personal environment of the cognitive agent. Aspects of the personal environment of the agent that we will study will be the history the agent has with a certain declarative fact and the current context in which the agent attempts to retrieve a certain facts. These two aspects will play an important role in the cognitive models reported in Part I as well as in the application-based studies reported in Part II.

APPROACH

The models of declarative memory we have developed in the context of this thesis are computational cognitive models within the cognitive architecture ACT-R (Anderson, 2007a; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004).

The concept of a cognitive architecture was introduced by Allen Newell as a way of dealing with the plurality of cognitive dichotomies (Newell, 1973). He reasoned that psychological phenomena should not be explained only in terms of unrelated hypotheses that were either confirmed or rejected but rather that psychological phenomena should be explained in each other's context. 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 also important to have a theory of decision processes, and of declarative memory. During the task, the participant has to decide whether he or she has found the target stimulus based on his or her memory of the previous presentation of the stimulus.

Therefore, it makes sense to study different cognitive phenomena within one framework, so that one theory (for instance on visual search) stays consistent with others (for instance on decision making or declarative memory). More recently, Anderson defined the concept of a cognitive architecture as

"a specification of the structure of the brain at a level of abstraction that explains how it achieves the function of the mind" (Anderson, 2007a p. 7).

This definition also takes into account that ultimately the brain is what generates behavior. However, in the context of this thesis this aspect will only play a minor role.

There are three major advantages of cognitive modeling within the constraints posed by a cognitive architecture. The first is that the modeler is constrained in the freedom he or she has in developing cognitive models. Because the architecture consists of a set of central assumptions on cognitive functioning, all models should adhere to these in order to be part of the architecture's body of work (R. P. Cooper, 2007). Cognitive models within an architecture support each other by applying the same central assumptions, thus providing more evidence that a cognitive model might be the correct functional specification of a certain behavior.

The second advantage is that a cognitive architecture gives the modeler the possibility to develop integrated cognitive models. Integrated cognitive models are meant to span multiple aspects of cognition (Gray, 2007b). Ideally, this would comprise models that capture human behavior in a certain task from stimulus perception to response action. For example, an integrated model of the visual search task mentioned before should account for the perception of the target stimulus, temporary storage of the target stimulus, perception of the stimulus array, matching of the stimuli in the array to the stored representation of the target, and issuing a motor command to respond (for instance by a button press). An integrated model can provide quantitative predictions of human behavior. If the model is successful it will accurately predict the behavior of participants engaging in the same task the model was designed for, for instance in terms of response times or accuracy scores.

The third advantage relates to the second. If it is possible to develop integrated models of human behavior in a certain task, then these models can also be deployed as predictors of human behavior in certain applied settings. Examples of these include cognitive models as intelligent agents in computer games (e.g., Shah, Rajyaguru, St. Amant, & Ritter, 2003) or serious games (e.g., C. P. Janssen, Van Rijn, Van Liempd, & Van der Pompe, 2007), cognitive models as simulated users of interfaces (Ritter & Young, 2001), and cognitive models as control models to track a user's attentive state (e.g., Grootjen, Neerincx, & Veltman, 2006).

All three major advantages of a cognitive architecture are utilized in this thesis. The models that appear in Part I take advantage of the central assumptions of the cognitive architecture we use (ACT-R)[1], and are able to provide quantitative predictions of human behavior. The models that appear in Part II can be thought of as integrated models that serve as intelligent agents in information selection tasks.

1. Although one might argue that the Retrieval by Accumulating Evidence in an Architecture (RACE/A) model proposed in Part I actually questions one of the central assumptions of ACT-R (specifically the memory retrieval theory).

Since the cognitive architecture ACT-R (Anderson, 2007a; Anderson et al., 2004) is at the core of the work presented in this thesis, we will now turn to a brief exposition of the central assumptions of ACT-R.

ACT-R

ACT-R is a hybrid cognitive architecture that consists of a set of modules that each process one kind of information. For instance, visual perception is handled by the visual module and motor commands are executed by the motor module. The declarative module is used for storing and retrieving information from declarative memory, the speech module handles the speech output, the aural module handles auditory perception, and the goal and imaginal are used modules for keeping track of sub goals and intentions (Figure 1.2).

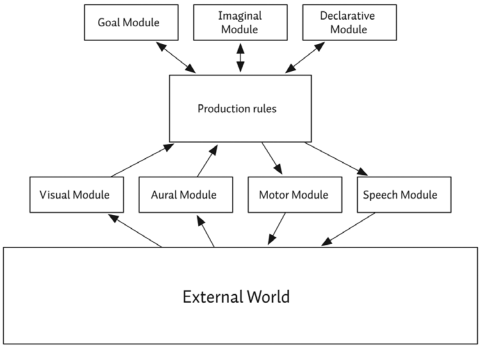


Figure 1.2. Example of modules (blocks) and information streams (arrows) in ACT-R.

Behavior in an ACT-R model emerges from the selection and subsequent execution of production rules. The production rule system communicates with the different modules through a set of buffers that can all contain one chunk of information. If the information that is present in the buffers matches the conditions of a production rule, that rule may be selected to execute its actions. If multiple production rules match, than the rule with the highest utility will be selected (this is referred to as conflict resolution). Production rule actions can be thought of as operations on the buffer contents, such as a request for a new chunk of information from declarative memory, or a request for pressing a button on a keyboard.

Chunks in ACT-R represent simple facts about the world, such as *Amsterdam is the capital of the Netherlands*, or *The object I am looking at is called a computer screen*. Both these example chunks are declarative facts, but the first example can typically be found in the retrieval buffer, and thus represents a fact retrieved from declarative memory, whereas the second example represents a visually observable fact of the world and might be present in the visual buffer. Because information does not occur in isolation chunks also have relations to other chunks. To this end, each chunk may have slots which indicate which other chunks are related.

So far, we have only discussed the symbolic level of ACT-R. Being a *hybrid* cognitive architecture ACT-R also has a subsymbolic level. The subsymbolic equations of ACT-R govern the *activation* dynamics of the chunks and *utility* dynamics of the production rules. We will first briefly discuss the concept of utility in ACT-R before turning to the - in the context of this thesis much more important - concept of activation.

If multiple production rules match the buffer contents this means that ACT-R's central cognition has multiple options that might be executed in order to achieve the current goal. In that case, the best choice is the production rule that has proven to be the most successful one in the past. In other words, the production rule with the highest utility will be selected, following Equation 1.1 (Anderson, 2007a).

$$P_i = \frac{e^{U_i/s}}{\sum_j e^{U_j/s}}$$ (equation 1.1)

2. Note that this equation is similar to the retrieval ratio equation that determines chunk selection in RACE/A, which will be discussed in Chapter 2.

Equation 1.1 denotes the probability that a certain production rule (rule *i*) will be selected.[2] The probability that a production rule is selected if multiple rules are applicable is determined by the utility of all applicable rules (U_j). The parameter *s* plays an important role in this mechanism, because it determines the trade-off between trying new strategies (exploration) and using strategies that have already been applied successfully (exploitation). A low value of *s* means a high probability of re-applying production rules that have been successful in the past; A high value if *s* means a high probability of selecting other production rules, and thus a higher chance of exploring new strategies.

For the chunks in ACT-R a similar system has been developed. All chunks have an activation level that represents the likelihood that a chunk will be needed in the near future. The likelihood is partly determined by a component describing the history of usage of a chunk called the *base-level activation* (B_i in Equation 1.2).

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right)$$ (equation 1.2)

In this equation, t_j represents the time since 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). The idea is that the activation of a chunk decays over time unless attention is shifted to that chunk and its memory trace is strengthened (Figure 1.3). This way, the base-level activation can be used to model both forgetting and learning effects (Anderson & Schooler, 1991).

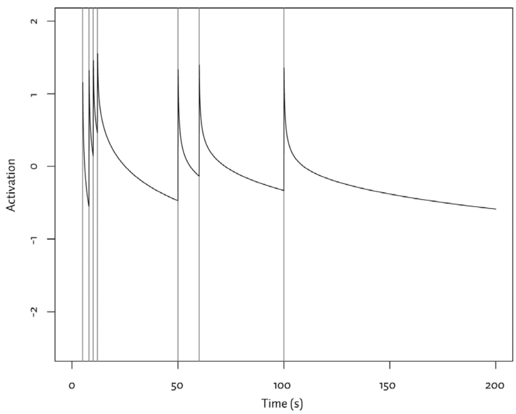


Figure 1.3. The base-level activation of a chunk. The activation decays over time, unless the chunk is being used (indicated by the grey vertical lines), in which case the activation is increased.

The total activation is the sum of the base-level activation, noise (ε in Equation 1.3), and another component describing the influence of the current context (*spreading activation*, Equation 1.3). The spreading activation component is composed of the associative values of other chunks, which are referred to in the slots of a chunk (chunks *j* in Equation 1.3) to chunk *i*, weighed by W_{kj}, representing the importance of various buffers (*k*) and the importance of associated chunks (*j*).

$$A_i = B_i + \sum_k \sum_j W_{kj} S_{ji} + \varepsilon$$ (equation 1.3)

Since in ACT-R it is assumed that the probabilities that chunks will be retrieved from memory are independent the activation values - which have an infinite range - can be scaled to a probability of retrieval according to Equation 1.4 (Anderson, 2007a).

$$P_i = \frac{1}{1+e^{-(A_i-\tau)/s}} \qquad \text{(equation 1.4)}$$

That is to say that if the activation of a chunk *i* (A_i) goes to infinity the probability of it being retrieved goes to 1. Likewise, the probability of a chunk with activation going to minus infinity will go to 0. Two parameters play a role in determining whether a chunk will be retrieved or not. The retrieval threshold τ sets the lower bound for memory retrievals. Chunks with an activation smaller than the retrieval threshold will not be retrieved. Similar to the production selection equation (Equation 1.1), there is a parameter s associated with this equation that controls the trade-off between exploration and exploitation in declarative memory retrievals.

Besides the probability that a chunk will be retrieved from declarative memory ACT-R also predicts the time it will take to retrieve it. If a chunk has a low activation value, indicating that it is not likely that this particular chunk will be needed right now, it is harder to remember, which will be reflected by a long retrieval time. This observation is captured by Equation 1.5, which determines the latency of retrieval, given a certain activation value, with F a scaling parameter:

$$RT = Fe^{-A_i} \qquad \text{(equation 1.5)}$$

In the context of this thesis it is important to note that Equation 1.4 and 5 constitute a ballistic model of declarative memory retrieval. If the actions of the production rule that is selected require a chunk being retrieved from memory, the probability of that chunk being retrieved and the time it takes to retrieve it are deterministic (Brown & Heathcote, 2005). This assumptions will be challenged by the data presented in Part I.

A detailed description of the ACT-R cognitive architecture is provided in (Anderson, 2007a; Anderson et al., 2004).

OVERVIEW

In Part I of this thesis a new mechanism for chunk retrieval will be proposed. The new chunk retrieval mechanism, Retrieval by ACcumulating Evidence in an Architecture or RACE/A, can be thought of as a more fine-grained process-model than the current, deterministic, chunk retrieval model in ACT-R (Equations 1.4 and 1.5 above). RACE/A provides a prediction of the activation dynamics *during* the retrieval process, instead of a ballistic prediction of the retrieval latency and the probability of retrieval. From certain experimental paradigms (for instance, tasks with varying stimulus onsets, e.g., M. O. Glaser & Glaser, 1982; or subliminal perception, e.g., Merikle, Smilek, & Eastwood, 2001), it becomes clear that the default activation equation in ACT-R is not capable of explaining the mechanism by which chunks are retrieved from declarative memory. Chapter 2 will present a general account of the model of memory retrieval processes put forth in this thesis. In addition, it will present experimental evidence that some tasks cannot be modeled within the constraints of the cognitive architecture ACT-R without using RACE/A. The focus of this chapter will be on the interplay between repetition priming (one of the effects successfully modeled by ACT-R) and semantic priming (one of the effects successfully modeled by RACE/A) (Van Maanen, Van Rijn, & Taatgen, submitted).

In Chapter 3, we will validate RACE/A by showing how it accounts for a range of benchmark phenomena. First, Section 3.1 demonstrates how RACE/A accounts for a set of common finding in the lexical decision literature (e.g., Wagenmakers et al., 2008; Wagenmakers, Steyvers, Raaijmakers, Shiffrin, Van Rijn, & Zeelenberg, 2004a). Second, we will present a model of a picture-word interference (PWI) task in which two stimuli dimensions are presented at various onset asynchronies (Section 3.2). The model is able to account for the PWI typical latencies (see also Van Maanen & Van Rijn, 2007b). Third, Section 3.3 shows that RACE/A can also account for effects caused by partially available information, such as in the subliminal priming paradigm (Merikle, Smilek, & Eastwood, 2001; see also Van Maanen & Van Rijn, 2007a).

In Chapter 4 we will present a cognitive model that explains both response times in the Stroop task and response times in a PWI task. By adjusting one parameter, the model can account for the recent finding that the PWI effect is located early in the mental processing stream, while the Stroop effect is late (Dell'Acqua et al., 2007). The chapter demonstrates that RACE/A is not just a theoretical novelty, but can also lead to new insights. The RACE/A enhancement to the ACT-R model described in this chapter was necessary to explain the difference between Stroop and picture-word interference (Van Maanen & Van Rijn, 2008; Van Maanen, Van Rijn, & Borst, submitted).

Part II of this thesis explores the possibility of developing recommender systems based on cognitive models of declarative memory. Recommender systems are applications that assist users in finding relevant or useful information. Often, recommender systems incorporate some form of personalization. This means that by tracking the cognitive state of a user (user modeling), a recommender system can offer recommendations that are relevant or useful for the individual user. By creating user models that predict the declarative memory behavior of individual users, we will be able to predict which facts will be personally relevant or useful.
Chapter 5 will show how the ACT-R declarative memory model can be applied for personalized information retrieval in the cultural heritage domain. It will focus on how the roles of the visitor and the museum can both be appreciated in an online museum. By providing cognitive models that take the role of a museum guide, we will study which aspects of the museum guide's behavior are important for successful artwork recommendations, and we will provide a framework for an online artwork recommender (the Virtual Museum Guide or VMG). This chapter is an extended version of Van Maanen (2007).

In Chapter 6, we will discuss how eye gaze can be used as an input device for such an online artwork recommender. Because the eyes are an important input modality for humans, the point of gaze of museum visitors might express the museum visitor's current interest. In this chapter, we will study how this insight can be used to develop recommender systems that provide personalized information on a specific artwork. We will present results that indicate that gaze and user interest are closely related in the cultural heritage domain. We will conclude by presenting a framework for a complete automated museum guide that may be capable of recommending relevant art and highlighting relevant aspects of the artwork, based on the visitor's point of gaze.

In Chapter 7, we will describe how the declarative memory model of ACT-R can be used for personalized information retrieval of (scientific) abstracts from an indexed database of publications (Van Maanen, Van Rijn, Van Grootel, Kemna, Klomp, & Scholtens, in press), and in Chapter, 8, we will provide the results of a competition between the activation-based recommender developed in the previous chapter and other models developed for the same task (Van Maanen & Marewski, 2009).

Accumulators in context: an integrated theory of context effects on memory retrieval

This paper has been submitted for publication as Van Maanen, L., Van Rijn, H., & Taatgen, N. A. (submitted). Accumulators in context: An integrated theory of context effects on memory retrieval.

INTRODUCTION

Cognitive architectures (e.g., Anderson, 2007a; EPIC, Meyer & Kieras, 1997a; Soar, Newell, 1990; CLARION, Sun, 2006), 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. During the task, participants have to decide whether they have found the target stimulus, based on a memory of the previous presentation of the stimulus. 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).

One of the successes of cognitive architectures (particularly from the architecture ACT-R) is the account of declarative learning. By incorporating a model that estimates the environmental demands on memory (Anderson & Milson, 1989; Anderson & Schooler, 1991), ACT-R models have been able to account for many effects of learning and memory (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998; Anderson, Fincham, & Douglass, 1999; Pavlik & Anderson, 2005). In addition, this declarative memory account has been used to study the interactions between memory and other cognitive processes, such as prospective time estimation (e.g., Taatgen, Van Rijn, & Anderson, 2007; Van Rijn & Taatgen, 2008) and cognitive control (e.g., Altmann & Gray, 2008) or the role of memory in interactive behavior (e.g., Lebiere & West, 1999; West, Lebiere, & Bothell, 2006).

However, one of the drawbacks of the theory of declarative memory retrieval that has been adopted in the cognitive architecture ACT-R is that it does not provide a theory of the *actual* retrieval process. Rather, it provides a prediction of the retrieval time of declarative information, as well as the probability of successful retrieval, under normal conditions. However, circumstances exist where such a *ballistic* model (Brown & Heathcote, 2005; Van Maanen & Van Rijn, 2007b) of declarative memory retrieval does not provide accurate predictions. This is illustrated in Figure 2.1, which shows the retrieval process as captured by a ballistic theory of declarative memory retrieval. At the onset of the retrieval process the time it takes to retrieve a declarative fact from memory (referred to as chunk A in Figure 2.1) as well as the identity of that fact are already known. However, cognition and perception do not stop while retrieving information from declarative memory, and it might very well be that new, relevant information becomes available *during* the interval between retrieval onset and the actual retrieval that may influence the retrieval process.

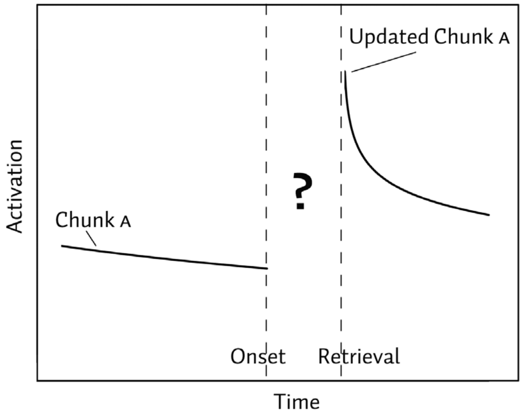


Figure 2.1. Illustration of the ACT-R retrieval process. Activation decays, until the central production system requests information from the declarative system (indicated by the onset of the retrieval). Then, after the time computed by the latency equation (Equation 2.3) has passed, the declarative system reports a retrieval. However, ACT-R has no account of the activation dynamics between these two time stamps.

Consider for example the following experiment. A participant is requested to name a picture that appears on a computer screen. The picture is accompanied by a word from the same semantic category (e.g., a picture of a cat with the word "dog"), which may appear at various time intervals just before or after the onset of the picture. When inspecting the latency data from this experiment (W. R. Glaser & Düngelhoff, 1984, see also Figure 2.8a below), it turns out that the presence of the word interferes with processing of the picture. Moreover, the time interval between the presentation of a word and the presentation of a picture mediates the response latency for the picture. The closer the word precedes the picture the slower the response. Surprisingly however, the maximum interference effect of the word on the picture is when the word *trails* the picture by 100ms.

What this example shows is that when studying interference effects such as these, many task aspects may play an important role. In this case, the asynchronous presentation of stimuli mediates the interference effect, and also the different qualities of words and pictures influence the latency, reflected by the fact that the maximum interference is not at an stimulus onset asynchrony (SOA) of 0ms (when word and picture are presented simultaneously), but at an SOA of 100ms. Complex interference patterns such as this cannot be fully explained by traditional architectural models that focus on a broad range of tasks. In the "Asynchronously presented stimuli" section we will further discuss how these two aspects (asynchronous presentation and stimulus quality) determine response latencies in this task.

Many specialized models exist that specifically address the interference issues in declarative memory, however. For example, a large body of work is devoted to understanding the decision dynamics in two-choice reaction time tasks (Ratcliff & Smith, 2004). In these models, the decision process is thought of as a process in which evidence for two response options is sampled, until a decision for one or the other has been reached. The decision time is determined by the length of this "deliberation process" (Busemeyer & Townsend, 1993, p. 432), and will be influenced by the accrual rate of the evidence for the response options. Thus, if all evidence points in one direction, the deliberation will be fast, and the decision time will be short.

In this chapter, we propose to integrate such a sampling process in a cognitive architecture, which we will refer to as Retrieval by Accumulating Evidence in an Architecture (RACE/A). This way, a number of new memory related phenomena can be explained by the architectural approach. Because in RACE/A it is possible to dynamically adapt the sampling process to new information, it becomes possible to model tasks in which asynchronies between stimulus presentations exist, something that is currently not possible in architectural models. For instance, in a dual-task in which the interval between the tasks governs the response latency (as for example in Psychological Refractory Period experiments), RACE/A explains how the decision process depends on the interval change (Van Maanen, Van Rijn, & Borst, submitted, as well as Experiments 1 and 2 of the current chapter). For these situations, RACE/A can provide quantitative model fits.

RELATED THEORIES OF MEMORY RETRIEVAL DYNAMICS

Previous models of memory retrieval have focused on the functional process underlying simple decision making (e.g., Ratcliff & McKoon, 2008) or perceptual identification (e.g., Usher & McClelland, 2001). These models are often referred to as sequential sampling models (Ratcliff & Smith, 2004). In sequential sampling models, the discrimination between mental representations is thought of as a mechanism that accumulates the likelihood that a certain mental representation is the intended one. Typically, there is a boundary, either fixed or relative to another accumulator, above which the representation is discriminated and may be used in another cognitive process. Accumulation depends on the quality or the quantity of a stimulus, either absolute or relative to other stimuli. Because the latency in these kinds of paradigms depends for a large extent on when the boundary for a specific accumulator is reached and accuracy depends on which accumulator reaches the boundary first, sequential sampling models provide an elegant explanation for speed-accuracy trade-offs often observed in cognitive tasks (Ratcliff & Smith, 2004).

Retrieval from declarative memory also involves discriminating between different mental representations. Thus, memory retrieval could also be described as a process in which the likelihood is accumulated that a certain mental representation is the intended one. Sequential sampling models therefore provide an explanation of many memory related phenomena (e.g., Ratcliff, 1978).

The three most important parameters that determine behavior in sequential sampling models are (Figure 2.2, Wagenmakers, van der Maas, & Grasman, 2007):

- A starting point of accumulation (*z*)
- Match boundaries (*a* and *b*)
- Mean drift rate (*v*)

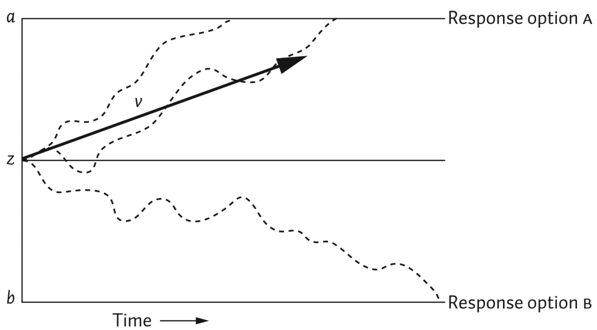


Figure 2.2. Illustration of a diffusion model. The response time is the time needed to reach one of the decision boundaries.

The match boundaries *a* and *b* in these kinds of models represent the two response options for a participant in the tasks that are modeled with the sequential sampling models. For instance, in lexical decision the match boundary represents the amount of accumulated evidence to give a “word” response and the non-match boundary represents the amount of evidence needed to give a “non-word” response.

The position of the starting point (*z*) relative to the match boundaries determines the prior likelihood of a match and a non-match. For example, if the starting point is closer to match boundary *a* than to match boundary *b*, the accumulation needed to cross *a* is less than the accumulation necessary to cross *b*. In this case, in the absence of a any drift towards *a* or *b*, the likelihood of reaching *a* is higher than reaching *b*. Manipulation of this parameter has been used to model participants' prior expectations on the probability of stimuli, for instance the probability of non-words in a lexical decision task (Wagenmakers et al., 2008). In the model of Wagenmakers et al. a high non-word probability was modeled by setting *z* to a lower value. This meant that crossing the non-word boundary was faster than the word-boundary because the accumulation process was shorter, which is visible in the data as well.

The third important parameter, mean drift rate, indicates the average speed of accumulation. A high value indicates a faster accumulation (a high drift). This parameter has for instance been manipulated to account for stimulus discriminability effects (Usher & McClelland, 2001). Thus, highly discriminable stimuli may be modeled by a high drift in either direction, and stimuli that are more difficult to discriminate may be modeled with a lower drift rate.

One of the drawbacks of the classical diffusion model is that it only accounts for two response options (a match and a non-match). Other memory retrieval models have been proposed that overcome this. For example, Usher and McClelland (2001) proposed a sequential sampling model for perceptual choice tasks in which each response option is represented by an accumulator, but in which the drift rates are dependent. Apart from accumulation caused by stimuli (the mean drift rate), the drift is also determined by lateral inhibition from other accumulators and decay. In this model, the time course of a perceptual choice is determined by the likelihood that a stimulus leads to one response, as well as the likelihoods of other responses.

Another well-known memory retrieval model is the REM model (Shiffrin & Steyvers, 1997). In this model, the retrieval process is thought of as a continuous Bayesian decision process, in which the odds that a particular decision will be made depend on the ratio of likelihoods between the response options. For instance, for lexical decision, the likelihoods for the "word" and "non-word" responses are considered to be a function of features of the stimulus. If the stimulus resembles a word, the likelihood of the "word" response is higher than if the stimulus consists of a completely randomized letter string. Certain instances of the REM model also include an aspect of sequential sampling (e.g., Norris & Kinoshita, 2008; Wagenmakers et al., 2004a). In these models, the likelihoods of the response options continuously drift and a decision is based on the current likelihood ratio in the system. In this way, the REM model accounts for pseudo-homophone effects in lexical decision under deadline or signal-to-respond conditions (Wagenmakers et al., 2004a).

These accounts have provided much insight in how retrieval from declarative memory works. However, computational models that are derived from these theoretical accounts often only model a single retrieval event. These models fail to appreciate that retrieving declarative knowledge from memory does not stand alone, but is always part of the execution of a particular task. Cognitive architectures on the other hand provide a theory of task execution (Newell, 1990). However, the explanation provided by these models is not always at the level of detail of sequential sampling models. RACE/A reconciles both approaches.

COGNITIVE ARCHITECTURES

From the many variants of a cognitive architecture that exist (Anderson, 2007a; e.g., ACT-R, Anderson et al., 2004; Soar, Laird, Newell, & Rosenbloom, 1987; EPIC, Meyer & Kieras, 1997a; Meyer & Kieras, 1997b; Newell, 1990; Rosenbloom, Laird, & Newell, 1993; CLARION, Sun, 2006; Sun, 2007), ACT-R is the theory with the most emphasis on declarative memory retrieval (e.g., Anderson et al., 1998; Anderson, Fincham, & Douglass, 1999; Anderson & Reder, 1999). However, we claim that ACT-R's theory of declarative memory is too static to account for the competition and interference effects in declarative memory retrieval that we address. In what follows, we will first introduce the cognitive architecture ACT-R and then clarify why in certain cases the current declarative memory theory in ACT-R is insufficient. Then we will introduce our new proposal for declarative memory retrieval, RACE/A.

ACT-R

ACT-R is a hybrid cognitive architecture in which behavior in a task can be described by a sequence of production rule executions. The rules specify which actions to execute given certain conditions. To execute a production rule, the conditions are matched against the current information state, which is represented by a set of buffers, each containing one piece of information. Which information is present at a certain point in time is determined by each of the specialized modules, that each process one kind of information. For instance, visual perception is handled by the visual module, and motor commands are executed by the motor module. 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 (Figure 2.3). 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.

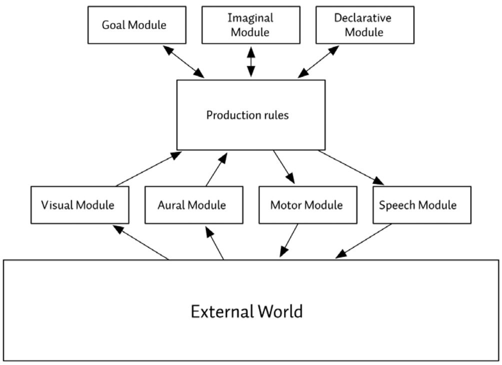


Figure 2.3. Modular layout of ACT-R. Boxes denote information-processing modules, arrows present transfer of information.

Thus, the presence of information determines which production rule is selected and executed. Both the presence and absence of stimuli can modify the buffer content and determine the selection of production rules, and the actions that are executed as part of a previous production rule. For instance, a production rule's actions may contain a request to retrieve certain information from memory, which will be stored in the retrieval buffer after it has been retrieved.

Declarative information in the ACT-R cognitive architecture is represented by chunks. These are simple facts about the world, such as *Amsterdam is the capital of the Netherlands*, or *The object I am looking at is a computer screen*. Both these example chunks are declarative facts, but the first example can typically be found in the retrieval buffer and thus represents a fact retrieved from declarative memory, whereas the second example represents a visually observable fact of the world and might be present in the visual buffer.

All chunks in declarative memory have an activation level that represents the likelihood that a chunk will be needed in the near future. The likelihood is partly determined by a component describing the history of usage of a chunk called the *base-level activation* (B_i in Equation 2.1).

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \quad \text{(equation 2.1)}$$

The base-level activation represents the theory that declarative memory is optimally adapted to the environment (Anderson & Schooler, 1991). 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 Equation 2.1, t_j represents the time since the jth 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, noise (ε in Equation 2.2), and another component describing the influence of the current context (*spreading activation*, Equation 2.2). The RACE/A theory will extend ACT-R by substituting the spreading activation component of the activation by a component describing spreading activation as part of the retrieval process. The spreading activation component is composed of the associative values of other chunks, which are referred to in the slots of a chunk (chunks j in Equation 2.2) to chunk i, weighed by W_j, representing the importance of associated chunks (j).

$$A_i = B_i + \sum_j W_j S_{ji} + \varepsilon \quad \text{(equation 2.2)}$$

The assumption in ACT-R is that chunks that are temporarily available to central cognition (that is, chunks that are present in the buffers) increase the probability that related chunks will be needed. The associations that exist between two chunks (S_{ji}) reflect the pattern of co-occurrences of the two events that these chunks represent (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.

An important reason to compute the need probability of a chunk (captured by the activation value) in ACT-R is to predict the time it will take to retrieve chunks. If a chunk has a low activation value, indicating that it is not likely that this particular chunk will be needed right now, it is harder to remember, which will be reflected by a long retrieval time. This observation is captured by Equation 2.3, which determines the latency of retrieval, given a certain activation value, with F a scaling parameter:

$$RT = Fe^{-A_i} \quad \text{(equation 2.3)}$$

However, this approximation of the declarative retrieval process suggests that information that becomes available after the declarative retrieval is initiated, but before the predicted retrieval time, cannot influence the time course anymore (Figure 2.1). We will refer to this type of retrieval model as *ballistic* (Van Maanen & Van Rijn, 2007b), to indicate that once initiated, the memory retrieval process can no longer be influenced.

RETRIEVAL BY ACCUMULATING EVIDENCE

In this section, we will develop the dynamics of the Retrieval by ACcumulating Evidence in an Architecture (RACE/A) memory retrieval theory. We will first present the equations that govern the activation dynamics and show how they relate to ACT-R concepts. As RACE/A is integrated in a cognitive architecture, we will show next how RACE/A in the architecture can account for simple reaction time experiments. Third, we will present data and a model of a task that can only be explained with the integrated account.

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 (Equation 2.1). The short-term dynamics are mediated by spreading activation from other chunks and the presence or absence of perceptual stimuli. During a retrieval process (e.g., the interval between Onset and Retrieval in Figure 2.4), the activation of chunks that match a set of retrieval conditions (chunks A and B) gradually accumulates until a certain decision criterion (explained below) has been reached. The chunk that has been decided upon (chunk A) will be retrieved from declarative memory, and the accumulation of activation stops. Because no new activation is being accumulated, the short-term component of the activation of all chunks decays. However, given that the usage history of the retrieved chunk has been altered (because it is currently being used), the chunk's long-term component is being increased in such a way that it greatly exceeds the current level of short-term activation. For this reason, the net activation of each chunk in the system can be described as

$$A_i = \max(B_i, C_i) \qquad \text{(equation 2.4)}$$

indicating that the activation of a chunk is the maximum of the need probability of a chunk (reflected by B_i) and the accumulating evidence for that chunk (reflected by C_i). Similar to most sequential sampling models, the short-term activation dynamics can be represented by a starting point, a drift, and a decision boundary, which will be discussed below.

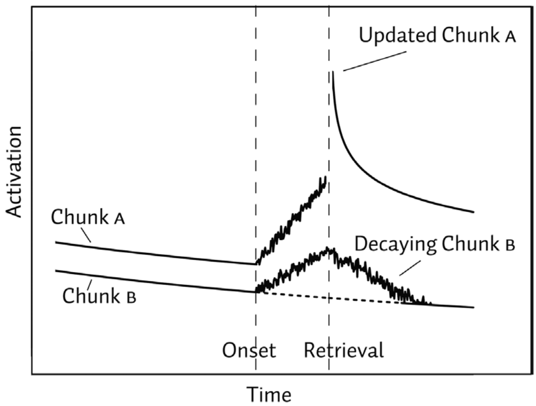


Figure 2.4. The RACE/A dynamics during a typical memory retrieval process. The two chunks compete for retrieval, and chunk A wins and is retrieved. Chunk A's base-level activation is updated (and decays), while chunk B's accumulated activation decays again.

STARTING POINT

The starting point of the accumulation reflects the prior probability that a chunk is needed. This is reflected by ACT-R's base-level activation equation (Equation 2.1), which incorporates the usage history of a chunk. Chunks with a high base-level activation start the accumulation of activation at a higher starting point, and are thus more likely to be retrieved from memory.

DRIFT

Drift in RACE/A is the reflection of the current demands of the environment. Thus, drift is a function of stimuli, as well as the currently active declarative facts. All facts and stimuli, which will collectively referred to as *sources of activation*, continuously spread excitatory activation towards associated chunks. This means that a chunk that has more sources of activation (more evidence) or sources with more activation ("stronger" evidence) will accumulate faster than a chunk 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 (indicated decaying chunk B in Figure 2.4). The drift in RACE/A is also determined by a logistically distributed noise sample, that adds stochasticity to the system.

These considerations are reflected by Equation 2.5, which may be referred to as the drift equation.[3] The drift equation captures the dynamics of short-term activation (C) of one chunk (chunk i) over time.

$$dC_i = [-\alpha C_i + \beta \sum_j S_{ji} A_j + \varepsilon]dt$$

(equation 2.5)

3. In the simulations reported throughout this paper, we used the discrete version of Equation 2.5: $C_i(t+\Delta t) = (1-\alpha)C_i(t) + \beta \sum_j S_{ji} A_j(t) + \varepsilon$

In this equation, the decay of short-term activation is expressed by α, which should be a value in the range [0,1] to create decay. The spreading activation component is a sum of the activation of other chunks (A_j), weighted by the associations that exist with chunk i (S_{ji}). The spreading activation component is scaled by a factor β that determines the overall accumulation speed. The noise is expressed by ε. To summarize, the decay parameter α together with scaling factor β determine the average drift of the chunks in the system. However, the chunk that receives the most spreading activation from sources of activation will (in the absence of noise) be the first to reach the decision boundary.

DECISION BOUNDARY

The decision boundary in RACE/A is relative to the activation of competitors in the system. This choice reflects the insight that if multiple memory representations are relevant, responding becomes more difficult (Hick, 1952; Luce, 1986). This is reflected by Equation 2.6, which expresses the conditions under which a decision will be made. If the activation of a certain chunk (chunk i in Equation 2.6) exceeds the activation of all competitors (j, including i), in the system by a certain ratio θ (referred to as the decision ratio), then that chunk is retrieved from memory. The duration of the retrieval process constitutes the interval between the onset of the retrieval process (when the request for a retrieval is made) and the moment at which the decision is made.

$$\frac{e^{A_i}}{\sum_j e^{A_j}} \geq \theta$$

(equation 2.6)

The relative likelihood of one chunk is the Luce ratio for that chunk (Luce, 1963).

One important special case for RACE/A is when there is no competition. Often, there is only a single chunk that matches a retrieval request. Consider for example an adult who wants to retrieve the answer to the problem 3 + 4 = ?. The activation of competing answers is very low, because adults are very experienced in this task and rarely make errors. Therefore, this results in a retrieval process in which there is practically no competition.

This special case is addressed in (Anderson, 2007a, Appendix 3.1), in which ACT-R's latency predictions are compared with an instance of the diffusion model. Assuming no noise (and no competitors), the retrieval time in RACE/A can be described as a ballistic model, in which the retrieval time is a function of the distance between the decision boundary and the Luce ratio for the chunk. Figure 2.5 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. This ensures that a RACE/A model can fit experimental data set in which no competition effects are to be expected, similar to default ACT-R models.

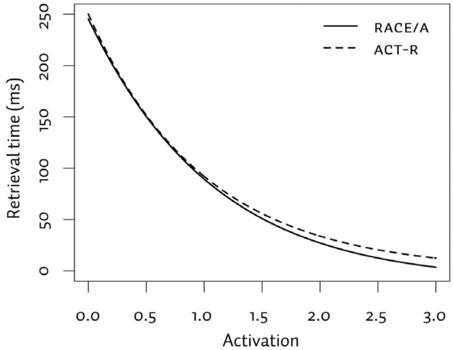


Figure 2.5. RACE/A and default ACT-R latency predictions for various activation levels. For RACE/A, these predictions only hold if there are no competitors.

SIMPLE SIMULATIONS

Before discussing memory retrievals using RACE/A in the broader scope of complex task execution, it is important to understand the activation dynamics of the 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. In addition, no noise is added to the system. Figure 2.6a (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 activates only one chunk. That chunk accumulates activation, whereas the other remains at its minimum value, which is the starting point (the base-level activation). Figure 2.6a (right) presents the relative Luce ratios for both chunks, with the dotted horizontal line indicating the decision ratio θ. Because there are two equally probable chunks in this competition, the Luce ratio at the start of the process is 0.5 for both chunks, but quickly changes in favor of the accumulating chunk, until it crosses the decision ratio.

Figures 2.6b presents the activation and Luce ratios for two chunks, of which one receives activation from a stimulus, but the other has a higher starting point. This is for instance the case if one chunk has a higher base-level activation because it is more familiar, or because the chunk has been recently retrieved. Under these conditions, retrieval takes longer because the Luce ratio of the activated chunk is smaller due to the higher initial activation from the other chunk.

In Figure 2.6c, a situation is depicted in which one chunk is activated later than the other. This will be an important case in the picture-word interference model described in the next section. Even though the retrieval process of the first chunk has already initiated, activation of the second chunk at a later moment in time influences the Luce ratio and thus increases the retrieval time of the first chunk. It should be noted that in order for the first chunk to be retrieved before the second, spreading activation from the stimulus to this chunk should be at least as high as spreading activation to the other chunk. Otherwise, the secondly activated chunk will be retrieved. This is what happens if a target stimulus trails a to be ignored stimulus in time, as in the picture-word interference experiment described below.

The last simulation addresses a condition in which one of the chunks is only activated for a short duration (Figure 2.6d). This is what happens in masked priming, 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. 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 chunks accumulates further. Again, the retrieval time is increased as compared to the default situation in Figure 2.6a, because the activation of the decaying chunk still influences the Luce ratios. The simulations depicted by

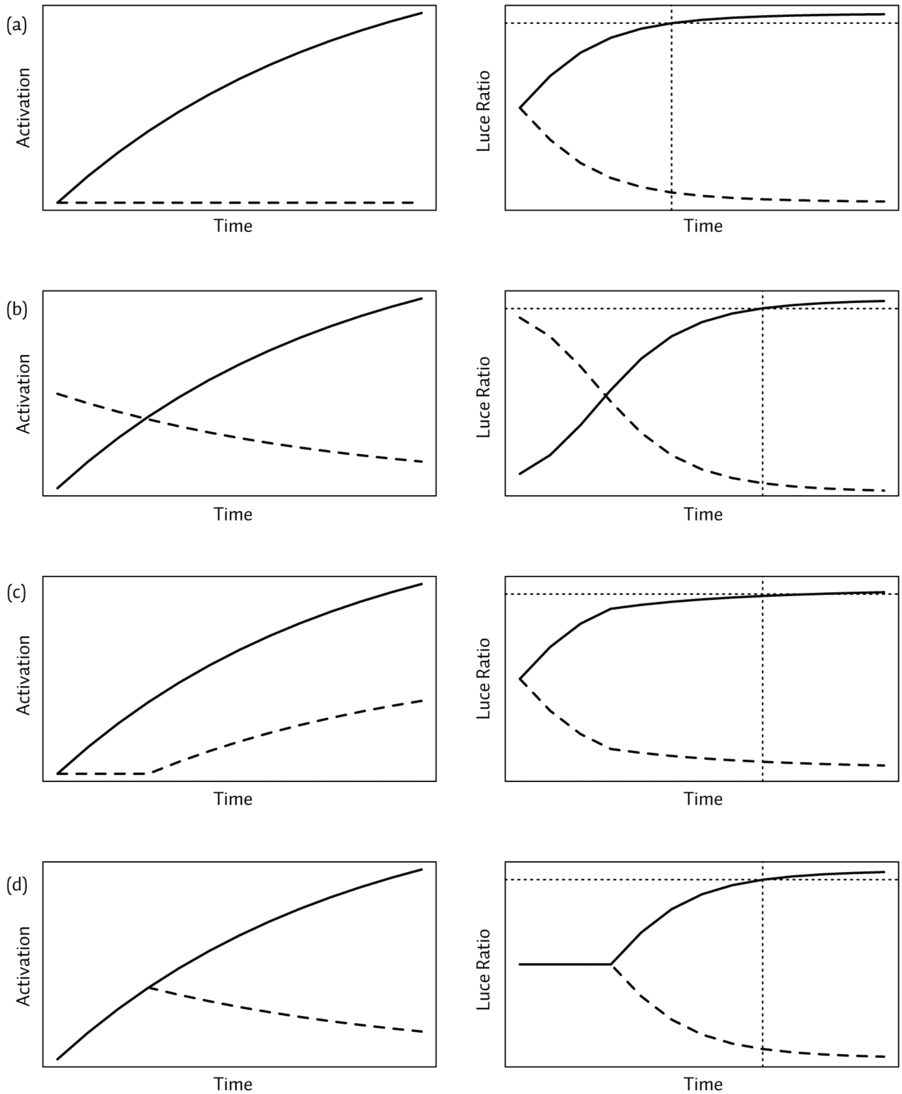


Figure 2.6. Race/a dynamics for four typical retrieval processes. Left: Activation of two chunks. Right: Luce ratios of two chunks. The vertical dotted lines indicate the retrieval times. See text for explanation of the differences between Panels a-d.

Figure 2.6c and Figure 2.6d are particularly difficult to explain with a ballistic theory of memory retrieval, because they simulate conditions in which information becomes available *during* the retrieval process.

To summarize these simple simulations, the Luce ratio of 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. If the chunks are associated, then the behavior is similar, but the retrieval times are increased even more (Figure 2.7, note that the x-axis differs from Figure 2.6). Under these conditions, both chunks accumulate due to spreading activation from the other chunk (compare Figure 2.6a with Figure 2.7a). This results in slower accrual of the Luce ratios and therefore a longer retrieval. This explains the semantic gradient effect found in picture-word interference studies (e.g., W. R. Glaser & Düngelhoff, 1984; Klein, 1964; Rayner & Springer, 1986; Smith & Magee, 1980). If two concepts are associated, for instance because they belong to the same semantic

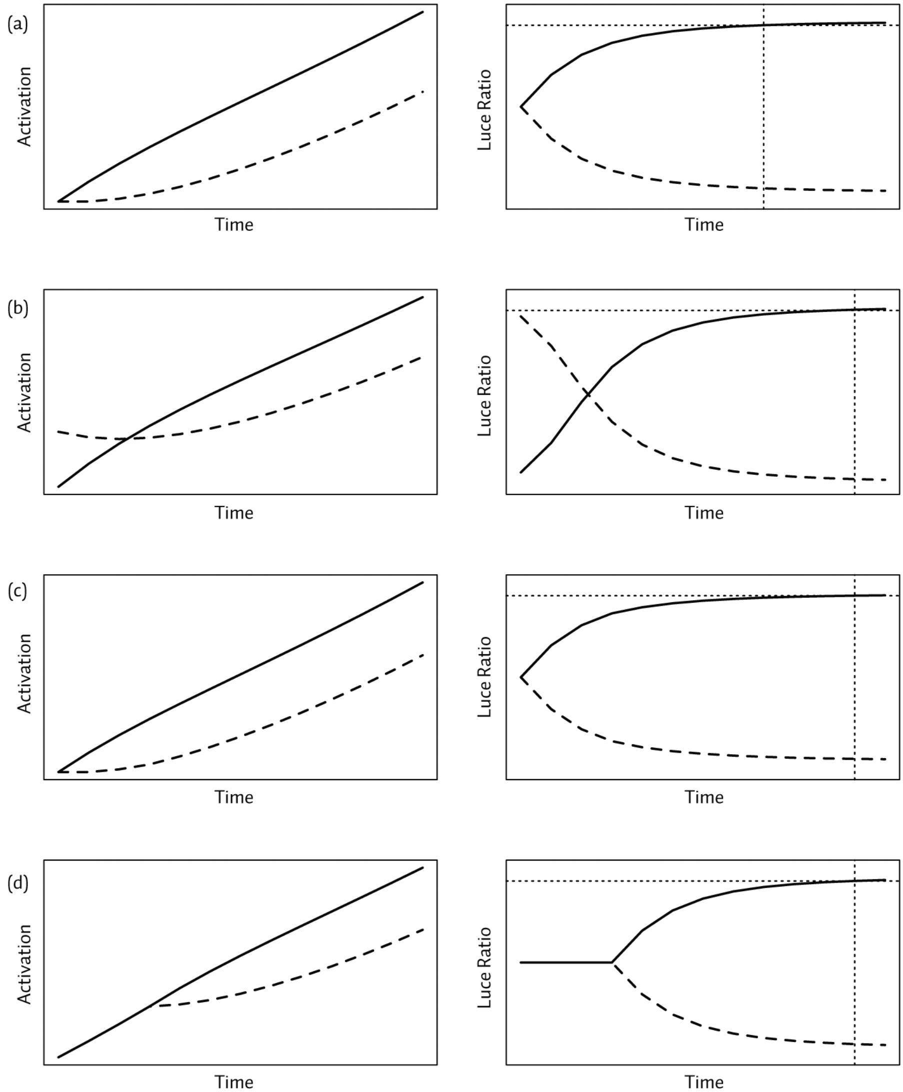


Figure 2.7. RACE/A dynamics for four typical retrieval processes, with spreading activation between the chunks. Left: Activation of two chunks. Right: Luce ratios of two chunks. The vertical dotted lines indicate that a chunk has been retrieved. Note that the axes differ from Figure 2.6.

category, the will excitate each other, increasing decision boundary and thus the retrieval time and the response latency. A full RACE/A model of picture-word interference will be presented in the next section.

ASYNCHRONOUSLY PRESENTED STIMULI

In this section, we will discuss a picture-word interference experiment performed by Glaser and Düngelhoff (1984, Experiment 1). As this is a prime example of an experiment that is difficult to explain with a ballistic model of memory retrieval. After introducing the experiment, we will provide RACE/A model fits to this data set.

Glaser and Düngelhoff's (1984) Experiment 1 consists of two tasks: One is naming a picture, while a word is presented (picture naming task); the other consists of reading a word while a picture is presented as distractor (word reading task). In both tasks, the distractors are presented on different SOAs relative to the target stimulus (either the picture or the word, depending on the task). Thus, if a distractor is presented at negative SOA, it is presented *before* the target stimulus. At positive SOAs, the distractor is presented *after* the target stimulus. The first condition is one in which both distractor and target stimulus refer to the same concept (e.g., a picture of a house versus the word *house*). This is referred to as the congruent condition. In two other conditions, distractor and target stimulus refer to different concepts. In the related condition the concepts belong to the same semantic category (e.g., a picture of a church versus the word *house*), in the unrelated condition the concepts do not belong to the same semantic category (e.g., a picture of a cat versus the word *house*). As control condition target stimuli were presented together with meaningless letter strings (picture naming task) or rectangles (word reading task).

Glaser and Düngelhoff (1984) show that interference decreases as the semantic relation of the word to the picture decreases, which is known as the *semantic gradient effect* (Klein, 1964; Lupker, 1979; Rayner & Springer, 1986; Smith & Magee, 1980). They also show *facilitation* in the congruent condition, meaning that latency is decreased when both target and distractor stimuli refer to the same concept (Bajo, 1988; Ehri, 1976). The most interesting finding for the current discussion is that the size of the interference effect is mediated by the temporal distance between target and distractor. If the onset asynchrony of the two stimuli is small, the interference is high. However, the interference does not maximize at an SOA of 0 ms, indicating simultaneous presentation of the two stimuli, but rather at an SOA of 100ms, indicating that interference is highest if the distractor is presented slightly after the target. A final effect observed in PWI studies is a clear *asymmetry* between the picture-naming task and the word-reading task: The semantic gradient and facilitatory effect virtually disappear in the word-reading task, while they were visible in the picture-naming task. The data of Glaser and Düngelhoff 's Experiment 1 are provided in Figure 2.8a.

PICTURE-WORD INTERFERENCE MODEL

We will now turn to a RACE/A model that explains the effects found in the picture-word interference paradigm. As RACE/A is tightly integrated with the ACT-R cognitive architecture, we will present the model in ACT-R terminology. Our model of picture-word interference comprises two chunk types: lemma chunks and concept chunks. The concept chunks can be regarded as representations of semantic properties of a certain concept. Chunks of the lemma type can be regarded as sets of orthographic and syntactic properties of a word (cf., Levelt, 1989).

Two types of stimuli can be presented to the model, pictures and text. If a stimulus is presented, it spreads activation to related chunks in declarative memory (Equation 2.5). Because of the familiarity of the depicted concepts, we assume that all pictures have been attended equally often. Thus, spreading activation from the pictures is kept at a constant value, which is a parameter of the model. However, 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 textual stimuli is higher than from pictures.

Figure 2.9 summarizes the connections that exist between chunks in this model. Text engages directly on the lemma chunks, while for pictures, conceptual information needs to be retrieved before the lemma of the picture can be retrieved. This reflects the finding that words can be processed without a conceptual level (W. R. Glaser & Glaser, 1989; La Heij, Happel, & Mulder, 1990), that is, the meaning of a word is not necessary for pronouncing the word. By contrast, an interpretation of the depicted concept is necessary for naming a picture.

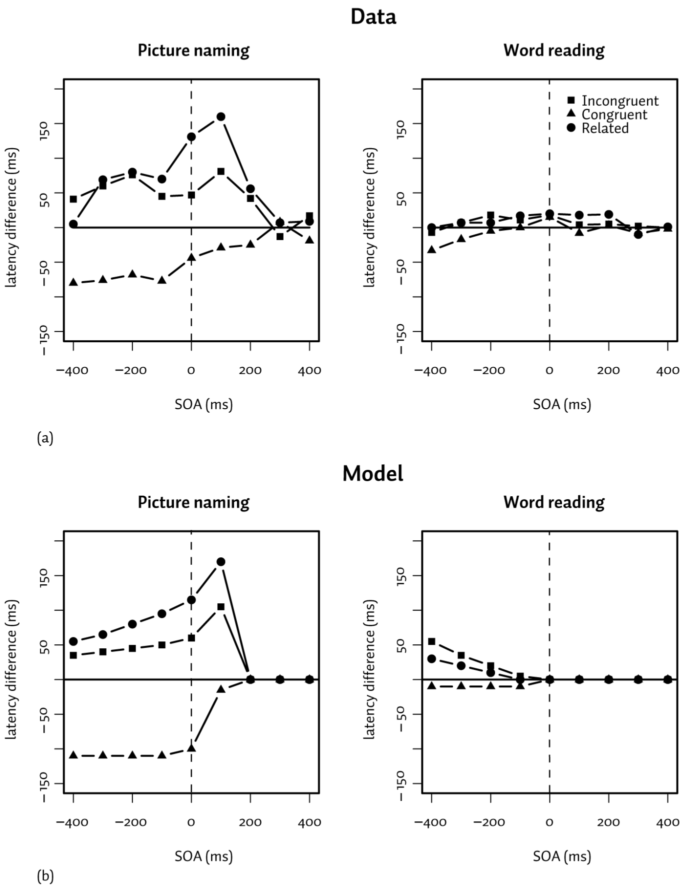


Figure 2.8. (a) PWI data from (W. R. Glaser & Düngelhoff, 1984, Experiment 1) and (b) model fit of the PWI model.

When a stimulus is presented, the model accumulates activation for concept and lemma chunks. After lemma retrieval, the model can commence word-form encoding and pronunciation of the response. For explanatory purposes, we assume for this data set that these stages do not influence the latency differences. Note however that in subsequent models these stages are included, as they will play an important explanatory role.

For both the picture-naming task and the word-reading task, we presented the model with the same conditions as the participants; a neutral, congruent, related, and unrelated condition, with SOAs of -400, -300, -200, -100, 0, +100, +200, +300, and +400ms, similar to the original Glaser and Düngelhoff (1984) experiment. In the neutral condition, no distractor stimulus was presented. That is, no text chunk was in the visual buffer at any moment. Glaser and Düngelhoff presented the subjects with a non-word distractor and a non-picture distractor respectively in the picture-naming and word-reading control condition. These were chosen in such a way as to minimize the amount of picture or word processing. Assuming a successful operationalization by Glaser and Düngelhoff, we simulated this condition by not presenting a distractor in the neutral condition. In the congruent condition, the distractor and the target stimulus both refer to the same concept. In the unrelated condition, distractor and target stimulus refer to different concepts.

When the picture is presented in the neutral condition, it spreads activation to its associated concept. At the same time, all concept chunks spread their activation to the lemma chunks. Because there is no distractor present, the most likely lemma to be retrieved from memory is the one related to the depicted item. Because there is no competition from other sources apart from the standard activation of the lemma representations, retrieval is relatively quick.

For the congruent condition, the distractor word is the name of the picture. If the distractor and target stimulus are both presented at the same time, then not only the effect from the picture contributes to the activation of the desired lemma, as in the neutral condition, but also activation spreading from the word. Therefore, retrieval will be faster than in the neutral condition. If the SOA is negative, that is, if the word is presented before the picture, the desired lemma is already active, because of spreading activation from the word to the lemma. The size of this effect correlates with the distance between the stimuli. For (positive) SOAs larger than +100 ms, the desired lemma is already retrieved before the word can contribute spreading activation. Therefore, no difference between the neutral and congruent conditions is predicted for these SOA levels.

In the unrelated condition, the word and the picture do not refer to the same concept. If both stimuli are presented simultaneously, a competing lemma is activated by the distractor word before the target lemma is activated. This will result in increased competition. At large negative SOAs, the word will be processed completely and rejected as a response option. Because of subsequent decay, there will be less interference at negative SOAs. The degree of interference is negatively correlated with the SOA, because shorter intervals will result in more active word-related lemmas when the picture is presented and thus more competition (cf., Figure 2.7b). At positive SOAs, the target lemma will become active before the distractor stimulus is presented. Thus, less time is available for the distractor to interfere (cf., Figure 2.7c). An interesting consequence of the asynchrony in the model between pictures and words in lemma activation is that interference is highest at a SOA of +100ms. Because pictures first activate conceptual representations, a slightly later presentation of a word will result in maximum competition between the lemmas. This patterns is supported by the data (Figure 2.8a).

If the target and distractor stimuli are semantically related, the distractor lemma is not only activated by the word, but also by the conceptual representation of the picture, mediated by a related conceptual representation. Thus, higher associations result in stronger competition and more interference.

The experimental setup for the word-reading task was equal to that of the picture-naming task. The only difference was that participants were instructed to read the words, hile ignoring the pictures. In general, pronouncing a written word is faster than naming an item in a picture (W. R. Glaser & Düngelhoff, 1984). As said, our model explains this by a shorter processing route for word-like stimuli and faster encoding of word-like stimuli. The faster encoding is reflected by higher activation of the text chunks than the picture chunks in the visual buffer.

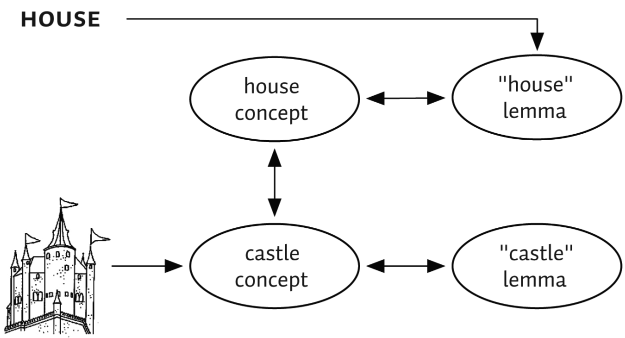


Figure 2.9. Connectivity between memory representations in the PWI model. In this example, a line drawing of a castle activates the conceptual representation of a castle, while the word "house" directly activates the house-lemma.

MODEL RESULTS & DISCUSSION

Figure 2.8 presents the results of our picture-word interference model. Because the model does not capture the vocalization aspects of the task, Figure 2.8 presents the latency differences between the neutral condition and the other conditions to account for differences in competition during memory retrieval. The Root Mean Square Error (RMSE) of the model fit is 28 ms (Here and throughout this article, we will only report RMSE and not the R^2 of the model fit, because for current purposes it is a more insightful measure of goodness-of-fit than R^2).

Because of the shorter processing route for words and faster encoding, there is less time for picture-induced competion to influence the word-reading process. This results in the typical Stroop asymmetry described by Glaser and Düngelhoff (1984) (Figure 2.8, RMSE=18ms). However, at negative SOAs, our model does predict a small interference effect, which is not in the data. Because of the longer processing route, it takes about 400ms for the picture-related distractor lemma to become fully activated. Therefore, at a negative SOA of 300-400 ms, the competition posed by the distractor lemma is maximal at the onset of the word.
Interestingly, in the Stroop variant of this experiment (M. O. Glaser & Glaser, 1982), a small but significant interference effect is observed at negative SOAs over 300ms, similar to our model prediction for PWI. This indicates that the Stroop effect could be explained by a longer processing account only. However, since generally PWI and Stroop are considered instances of the same process (e.g., Cohen, Dunbar, & McClelland, 1990; MacLeod, 1991; Roelofs, 2003; Van Maanen & Van Rijn, 2007b; Van Maanen, Van Rijn, & Borst, submitted), the issue of interference in word reading remains unsolved. We did not examine this issue further, because the model presented here is aimed at providing a general account of interference effects in memory retrieval.

A ballistic model of picture-word interference would not have been able to account for the SOA effects in PWI. Consider the latency equation of ACT-R (Equation 2.3): Given the ballistic nature of declarative memory retrieval, a standard ACT-R model cannot explain the interference and facilitation effects that are observed at small positive SOAs, since the latency is already determined at the onset of the target stimulus. At negative SOAs, the retrieval initiated by the distractor might not yet be finished once the retrieval initiated from the target stimulus is requested. Therefore, in a model that does not consider the retrieval process, no influence from the distractor stimulus would be predicted, resulting in identical predictions for all conditions.

MASKED STIMULI

The previous section showed that RACE/A can account for the effects of asynchronous presentation of stimuli at very short intervals. The PWI model is an example of how RACE/A processes information that becomes available *after* a memory retrieval process has already initiated. In this section, we will discuss a model that demonstrates how RACE/A accounts for conditions in which information is not available for the full duration of a trial. More specific, we will discuss a model of a classical masked priming experiment (Marcel, 1983, Experiment 3).

The experiment consisted of a Stroop task in which participants were asked to name the color of a slide, while a color word was flashed in front of the slide for a brief duration. Four prime type conditions were tested: Color congruent, color unrelated, neutral, and no-word. In the congruent condition, the prime was the name of the target color, whereas in the unrelated condition the prime was the name of another color. In the neutral condition, the prime was a non-color word. The no-word condition presented the mask only. Thus, no prime was presented. Two awareness conditions were tested. One in which the prime was masked (unaware condition) and one in which the prime was presented for 400 ms (aware condition). The presentation duration of the masked prime was chosen in such a way that participants could not discriminate between the presence or absence of a prime. Prime and cue were presented simultaneously (Marcel, 1983).

Participants had to respond to the color patches by pressing a button associated to one of the colors. The aware condition replicated the typical patterns of interference and facilitation as commonly found in the Stroop paradigm. In the unaware condition, a smaller interference effect was found as compared to the aware condition (Figure 2.10a).

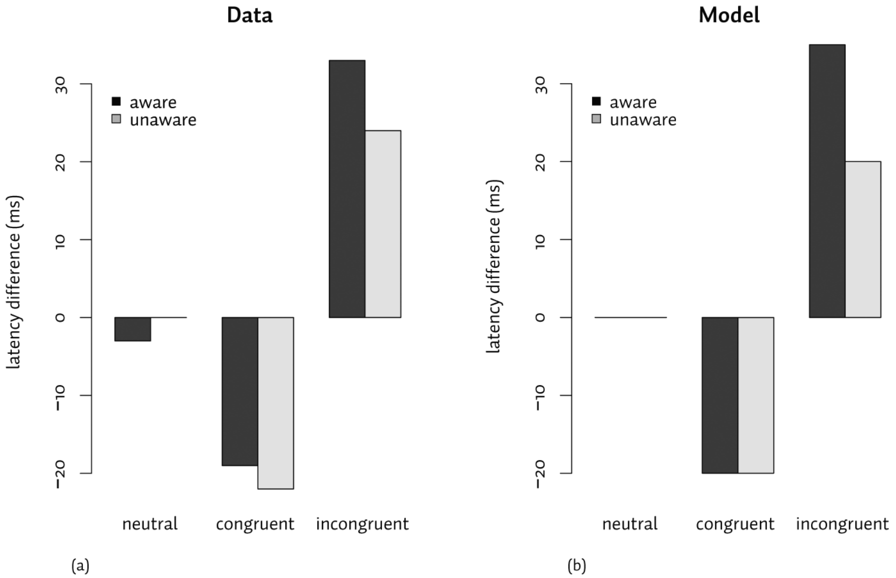


Figure 2.10. (a) Masked priming data from (Marcel, 1983, Experiment 3) and (b) model fit of the masked priming model.

MASKED PRIMING MODEL

The masked priming model is very similar to the PWI model. Concept chunks and lemma chunks contain the conceptual information on color and the syntactic information on color words, respectively. However, because the original experiment involved a response using a button press, we included extra chunks that represented information on which button was associated with which concept. These chunks will be referred to as motor-mapping chunks, because they map a conceptual representation onto a motor response (Figure 2.11).

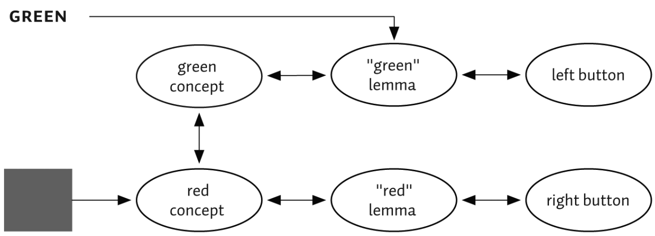


Figure 2.11. Connectivity between memory representations in the masked priming model. The connectivity is similar to the PWI model, but for the response a chunk representing the concept-to-button information is activated.

The rest of the model was similar to the picture-word interference model. Thus, textual features of a stimulus (such as the word "red") activate lemma chunks. Color patches activate concept chunks that represent color information. Concepts and lemmas spread activation to each other, as well as concepts that are from the same semantic category, such as two colors.

We estimated the presentation duration of the prime in the unaware condition at 70 ms, which is in the range of durations used in the original experiment (30-80 ms, Marcel, 1983). With this duration, the prime's Luce ratio did not exceed the decision boundary, indicating that the model remained unaware of the presence or absence of the prime.

MODEL RESULTS & DISCUSSION

Figure 2.10 presents latency differences (RMSE = 2ms). The model predicts similar effects of awareness as are present in the data, as well as the normal Stroop effects. Because the prime (the color word) is only present for a short duration in the unaware condition, it spreads less activation towards the associated lemma. Therefore, the competition between color concepts is less (because the lemma is also less activated than in the aware condition), resulting in a smaller interference effect in the unaware condition than in the aware condition.

In the congruent condition, the model predicts no effect of awareness. This results from a floor effect in the retrieval time. Because the color patch already activates the concept enough for retrieval, the added value of the congruent word is diminished and therefore the presentation duration has no effect on the latency.

As in picture-word interference, an extra stimulus in masked priming influences the retrieval of a target stimulus. From a symbolic perspective, stimuli have to be active in order to engage in cognitive processing. In ACT-R, this means that a stimulus has to be present in a buffer in order to influence other cognitive processes. Stimuli that are presented for such short times as are common in masked priming paradigms will not become available in the visual buffer, because ACT-R assumes an attention shift to the stimulus which takes a certain amount of time, exceeding the presentation duration of the prime (Anderson, Matessa, & Lebiere, 1998). In addition, if a prime were present in the visual buffer it would be available to conscious processing. In that case, the model would no longer be an accurate description of the task. Thus, the ballistic nature of the declarative retrieval theory in the cognitive architecture prohibits accurate modeling of masked priming experiments.

EXPERIMENT 1: REPETITION PRIMING

The previous models addressed competition effects in relatively straightforward conditions. In each trial, one chunk was retrieved, and the latency was a function of the competitive effects. The experiment and model presented in this section have a more complex structure. In this experiment, a picture-word interference (PWI) task is performed as the second, main task of a psychological refractory period (PRP) design. In addition to the normal PRP setup, which will be discussed shortly, we introduced a systematic repetition of the presented items. This way, this experiment requires the interaction of RACE/A - to account for the PWI interference effects – with ACT-R's long-term declarative memory theory to account for the repetition effects, and ACT-R's production rule and buffer system to account for the general task setup.

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 main task). The interval between the stimulus onsets of the two tasks is manipulated (SOA). A typical finding, known as the PRP effect (Telford, 1931) is a negative correlation between SOA and response latency on the main task, Responses to the first task are typically unaffected by varying the SOA. The PRP effect has been explained (McCann, 1992, but see Meyer & Kieras, 1997b; e.g., Pashler, 1994; Welford, 1967, 1980) by the assumption that both tasks share a cognitive resource that can only be used by one task at a time. Thus, the second task is delayed because the first task still requires a critical resource (Figure 2.12). As the interval between the tasks increases, the delay becomes smaller, resulting in a faster main task response.

Applying the PRP design to PWI, (Dell'Acqua et al., 2007) have shown that the effect size of picture-word interference diminishes with decreasing PRP-SOA[4]. 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 a small interval between the first and the second task generates a large delay (referred to as "cognitive slack"), in which the interference can be resolved. If the interval increases, the delay of the second task decreases, and therefore the interference becomes apparent in the reaction times (for a different interpretation of this data, see Van Maanen, Van Rijn, & Borst, submitted).

4. To distinguish the SOA manipulations in PWI and PRP context, we will refer to the SOA in the PRP context as PRP-SOAs. The SOA between word and picture, which was part of the PWI design discussed before, was always zero in these and following experiments.

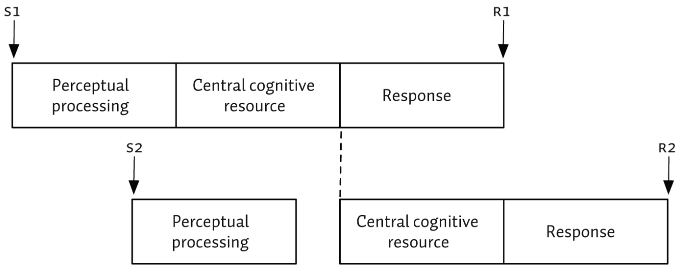


Figure 2.12. Gant diagram of the PRP effect. The top bar indicates processing of the first task. The bottom bar indicates processing in the second task. S1: stimulus of task 1; S2: stimulus of task 2; R1: response on task 1; R2: response on task 2.

The different SOAs of the PWI task show different effects because of short-term changes in priming-related activation. Another possible cause for differences in retrieval latency is the number of repetitions of a particular target: the more repetitions, the shorter the latency (repetition priming). It has been shown that repetition of items results in priming effects that span intervening trials (e.g., Becker, Moscovitch, Behrmann, & Joordens, 1997; Joordens & Besner, 1992; Scarborough, Cortese, & Scarborough, 1977), suggesting that this form of priming affects long-term memory, rather than short-lived residual activation from a previous trial. We predict that the stronger long-term representation of the stimuli decreases the interference from related and unrelated distractor words, because the starting points of the accumulation process are higher on each repetition, resulting in faster retrieval process and thus less interference.

METHODS

Participants

23 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.

Stimuli

49 images were taken from the PD/DPSS image set (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 the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993) and

matched to the related distractors with respect to word length (plus or minus 1 letter) and word frequency (plus or minus 10%).

The tones for the primary task consisted of a 300Hz, 600Hz, and 1200Hz tone, similar to the experiment conducted by Dell'Acqua et al. (2007).

Design

For each participant, an experimental list was created in which each image was combined with each relatedness condition (related and unrelated) and every PRP-SOA (100ms, 350ms, and 800ms). 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: (1) The same condition (relatedness or PRP-SOA) did not occur more than twice in a row, and (2) the same tone did not occur more than twice in a row.

Procedure

Each trial started with the presentation of a fixation cross for 1500ms followed by the tone-classification tone for 150ms. After the PRP-SOA, the PWI stimulus was presented. The word and the picture that formed the PWI stimulus were presented simultaneously (that is, all trials were presented with a PWI-SOA of 0ms). The participants were instructed to always respond to the tone first and then to the PWI-stimulus. The tone had to be classified as either low, medium, or high pitch 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. If participants failed to answer in the correct sequence, a screen informing them of the correct procedure was presented.

The participants were tested individually. First, each participant practiced the tone classification task in isolation. Second, a set of PWI stimuli was presented in single task setting and the participant was instructed to name the picture and ignore the word. 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. After this practice block, the actual experiment started. The experimental block was preceded by two filler trials that were the same for all participants and that were not analyzed.

RESULTS

One participant was excluded from the analyses because of an excessive error pattern (52% erroneous trials). Trials on which the remaining 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/ PRP-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 Figure 2.13. 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. Relatedness, PRP-SOA level, and 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 inter-subject and inter-item variability. An ANOVA on the mixed-effects model (Baayen, Davidson, & Bates, 2008) showed that the following factors contributed significantly to the participants' reaction times: PRP-SOA β = -0.401; $F(1,6122) = 741$; $p < 0.001$, relatedness $\beta = 45.5$; $F(1,6122) = 10.3$; $p = 0.001$, repetition $\beta = -48.2$; $F(1,6122) = 246$; $p < 0.001$, and a three-way interaction between SOA, relatedness, and repetition $\beta = 0.0529$; $F(1,6122) = 4.10$; $p = 0.04$. Further analysis of the effects per SOA level revealed that this three-way interaction is caused by the decreasing interference over repetition at a PRP-SOA of 800ms (relatedness times repetition, $\beta = 21.9$; $F(1,2059) = 5.20$; $p = 0.02$). This interaction was not significant at other PRP-SOAs (Fs < 1). To summarize, this experiment replicates the typical PRP effect, in which response latency on the main task are negatively correlated with PRP-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 800ms PRP-SOA, the repetition effect was smaller for trials with related distractors as compared to the 100ms PRP-SOA.

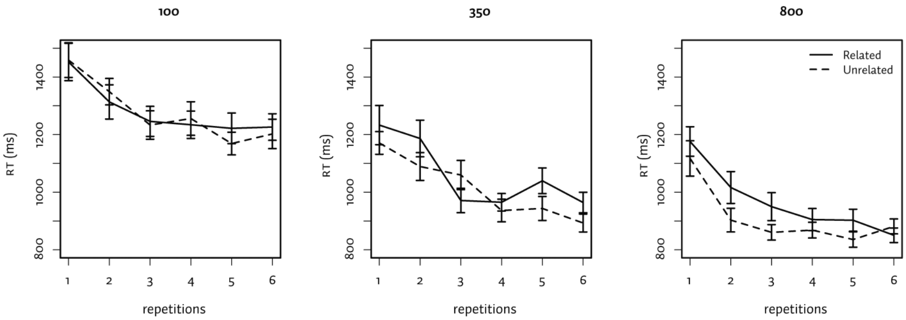


Figure 2.13. Data from Experiment 1. Each panel represents the latency on each PRP-SOA level for all repetitions of the picture.

DISCUSSION

The difference between the related and the unrelated condition is reflected by a difference in response latency, the interference effect. Since in the related condition picture and word are semantically related (and 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., W. R. Glaser & Düngelhoff, 1984; Van Maanen & Van Rijn, 2007b). In the model, this effect of competition is accounted for by the retrieval ratio presented in Equation 2.5.

The size of the interference effect decreased with decreasing SOA, as well as with repetition of pictures. 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 caused 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).

All three panels of Figure 2.13 show latency curves very similar to the power law of learning, 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 ACT-R. This way, the starting points of all competing memory representations will be set at a principled value, because the starting values are determined by a validated theory of long-term memory processes, as well as precise predictions of the time course of the memory retrievals, both within trial and between trials.

MODEL OF EXPERIMENT 1

The low-level PWI aspects of the model of Experiment 1 are identical to the earlier discussed models: The model retrieves conceptual representations from memory if a picture-like stimulus is presented and retrieves a lemma representation from memory if a word-like stimulus is presented. Because lemmas spread activation to the conceptual representations that relate to them, the presentation of a distractor word causes interference at the conceptual level. The relative decision boundary that determines retrieval from memory becomes harder to reach for the conceptual representation of the picture, increasing the retrieval time. The different activation levels of the target chunk versus competing chunks 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, making it even harder to reach the relative decision boundary. 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 Figure 2.6a and Figure 2.7a, 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, than it retrieves a motor program to articulate the desired response.

In addition to mechanisms to account for the PWI-aspects of the experiment, the model also contains mechanisms to account for the tone-classification task. If a tone is presented, the model processes auditory information and retrieves a memory trace that encodes the appropriate stimulus-response mapping (that is, which button to press given the perceived tone). Because of the task instruction to withhold the vocal PWI response until the tone classification is made, the model includes a control state that ensures that the retrieval of the response lemma (as part of the PWI task) does not start until the retrieval of the response stimulus response mapping (as part of the tone classification task) has been completed. As soon as the response mapping is retrieved, 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 in 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.

Model results and discussion

Similar to Experiment 1, we excluded outliers that were more than three standard deviations of the mean of each PRP-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 an error is returned. Consecutively, the model tries to retrieve the desired chunk again, resulting in prolonged response latencies.

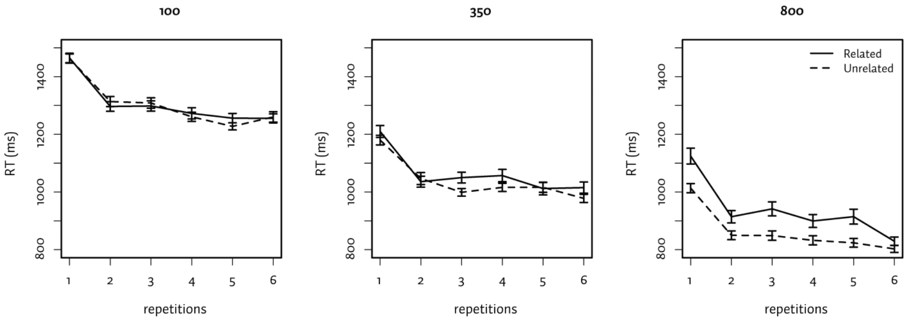


Figure 2.14. Model fit for Experiment 1.

Figure 2.14 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 & Saenz, 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 (e.g., Wagenmakers, Zeelenberg, Steyvers, Shiffrin, & Raaijmakers, 2004b; Zeelenberg, Wagenmakers, & Shiffrin, 2004).

As an extra validation of the model, Figure 2.15 shows the models fit to the main PRP effects (RMSE = 20 ms); a decreased latency as a function of increased SOA (for all relatedness conditions), and a decreased interference effect with increased SOA.

EXPERIMENT 2: TASK STRATEGY

Experiment 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.

This synergistic advantage of the integration of RACE/A with a cognitive architecture will be further demonstrated in Experiment 2. Participants in Experiment 2 were asked to perform exactly the same two tasks as in Experiment 1. However, this time an extra condition is included in which the word names the picture. This condition will be referred to as the congruent condition. This relative minor manipulation has the potential to change the outcome of the

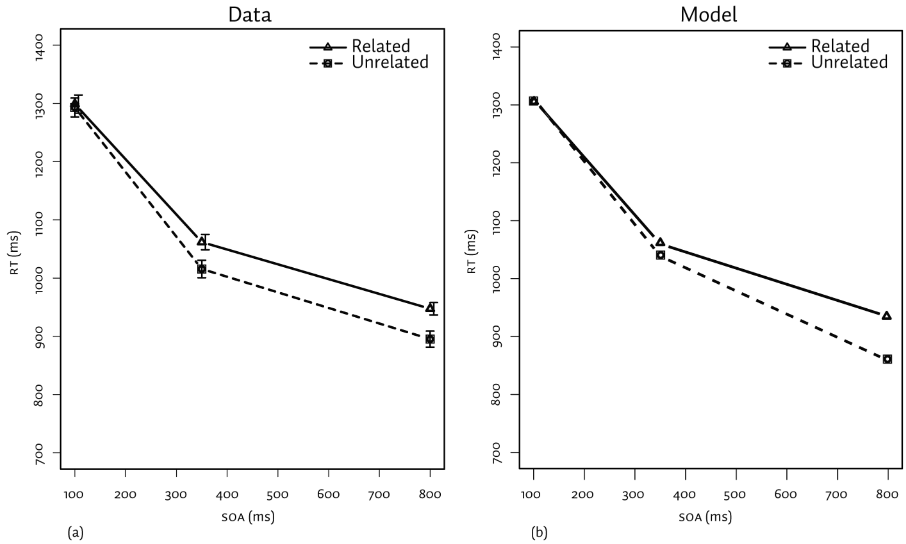


Figure 2.15. (a) Model fit for Experiment 1, aggregated over repetitions. (b) Data of Experiment 1, aggregated over repetitions.

task significantly. In the context of Experiment 1 it is assumed that the participants tried to minimize the processing of the words as much as possible, since in all conditions the word hampered task execution (that is, naming the picture). However, in Experiment 2 it might be beneficial to follow a different strategy, because in one third of the trials (the congruent trials), reading the word actually facilitates the correct answer. Because of the increased benefit of word processing as opposed to Experiment 1, more interference is predicted, which should be observable in a larger latency difference between the related and the unrelated condition.

METHODS

Participants

22 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.

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.

Design

The design was the same as in Experiment 1.

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.

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 Figure 2.16b. Experiment 2 is analyzed analogous to Experiment 1. Thus, we fitted a linear mixed effects model with relatedness, PRP-SOA, and repetition as fixed effects and participant and picture as random effects. An ANOVA on the factors revealed that there were main effects of PRP-SOA level ($\beta = -0.460$; $F(1, 8947) = 1192$; $p < 0.001$), relatedness ($\beta_{related} = 408$; $\beta_{unrelated} = 238$; $F(2, 8947) = 188$; $p < 0.001$), and repetition ($\beta = -22.0$; $F(1, 8947) = 315$; $p < 0.001$). Repetition also interacted with the other factors (repetition times relatedness interaction: ($\beta_{repetition\ x\ related} = -43.7$; $\beta_{repetition\ x\ unrelated} = -16.7$; $F(2,8947) = 9.23$; $p < 0.001$ and repetition times PRP-SOA interaction: $\beta_{repetition\ x\ SOA} = -0.00586$; $F(1,8947) = 18.7$; $p < 0.001$) and there was a significant three-way interaction ($\beta_{repetition\ x\ SOA\ x\ related} = 0.0639$; $\beta_{repetition\ x\ SOA\ x\ unrelated} = 0.0352$; $F(2,8947)=8.41$; $p<0.001$). However, the main finding in Experiment 2 is that the PRP-SOA between tone presentation and PWI presentation does not have an effect on the interference effect size (PRP-SOA times relatedness interaction: $F(2,8947) < 1$).

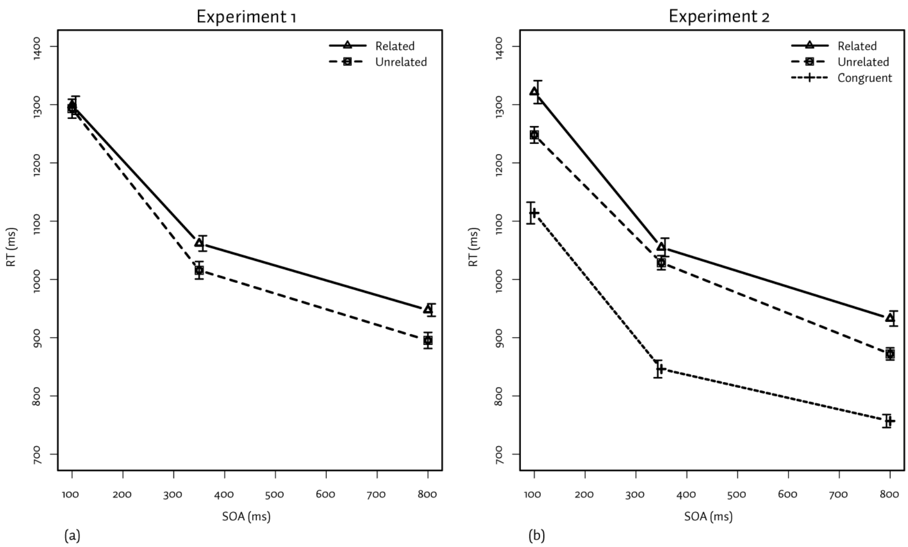


Figure 2.16. Data from (a) Experiment 1 and (b) Experiment 2.

DISCUSSION

The typical PRP speed-up is present, suggesting a correct operationalization of the PRP design and indicating the presence of "cognitive slack". Therefore, the observation that 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 cognitive slack time. If participants process the word more often or to a higher level than in Experiment 1, 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. 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, Roelofs, & Meyer, 1999; McClelland, 1979) in which activation from different stages spreads to each other, affecting processing later in the task. In this experiment, cascading activation means that the increased interference may partly be due to competition in the response stages, and is therefore not affected by the cognitive slack time, because the response is always *after* the delay. The cognitive model discussed below will demonstrate that one way of accounting for these effects is by assuming a different strategy than the one taken in Experiment 1.

MODEL OF EXPERIMENT 2

The model of Experiment 2 is the same as the model of Experiment 1, with the exception of the repetition priming mechanisms to keep the model tractable. The model performs the same task as in Experiment 1, but we assumed that participants in Experiment 2 were more likely to process the word than in Experiment 1. To account for this strategy, the model retrieves the lemma activation associated with the word stimulus upon word presentation. This differs from the strategy in the model of Experiment 1 in which the word was not actively retrieved and stored in the retrieval buffer until the response selection stage.

If word and picture are congruent, this new strategy means that lemma information on the picture is already available when the model requires it. It can therefore directly continue with retrieving relevant wordform information. In this case, the model performs two steps in parallel (conceptual and lemma retrieval), which accounts for the faster responses in the congruent condition. However, if word and picture are not congruent (that is, in the related or unrelated condition), the model initially retrieves the incorrect lemma. The reason for this is that that 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. Next, 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 boundary. Therefore, retrieval in the related condition is slower than in the unrelated condition. To validate that a different strategy has been adopted in Experiment 2 as compared to Experiment 1, we also ran a model *without* the reading strategy.

Model results and discussion

Figure 2.17b presents the fit of the model to Experiment 2 (RMSE = 74 ms). Similar to the empirical data, the model shows an increased interference effect at an PRP-SOA of 100ms as compared to Experiment 1. Also, the model shows the standard PRP effect.

Figure 2.17a presents the model's fit without the reading strategy on the previous experiment (Experiment 1, RMSE = 48 ms). Here, the model shows the dependence of the interference effect (the difference between the relatedness conditions) on the PRP-SOA between the tasks. If we calculate how well the no-reading-strategy model fits with Experiment 2, we find a much worse fit (RMSE = 392ms). Likewise, if we compare latency predictions of the reading-strategy model of Experiment 2 with the data of Experiment 1, we also find

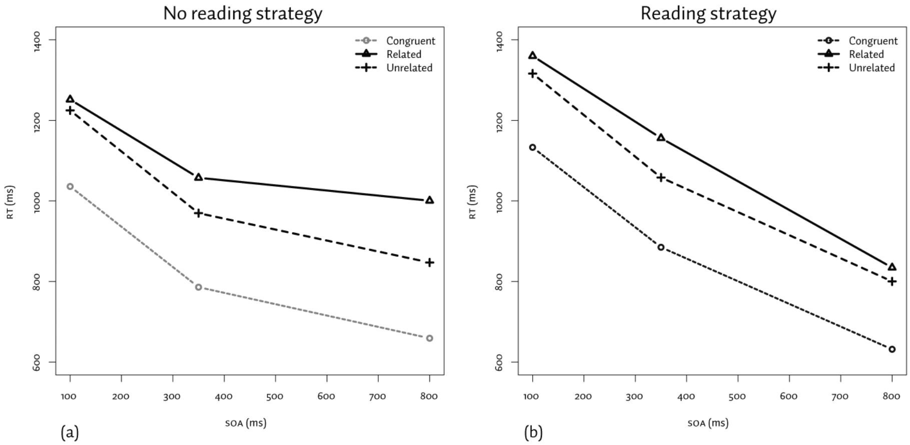


Figure 2.17. Model fits of Experiment 2. (a) The panel shows simulations of the model without reading the word upon presentation. The gray dotted line indicates the hypothetical condition in which word and picture are congruent. (b) Model fit of Experiment 2, with reading the word upon presentation.

a worse fit (RMSE = 514ms, with congruent condition excluded). Thus, the two simulations combined show that in Experiment 1 participants are more likely to ignore the word as much as possible, whereas in Experiment 2 they are more likely to read the word upon presentation. This difference in strategy accounts for the difference in response latency patterns observed between Experiments 1 and 2.

GENERAL DISCUSSION

SUMMARY

In this paper, we have provided a theory of context effects on declarative memory. The theory, called Retrieval by Accumulating Evidence in an Architecture or RACE/A makes three important assumptions. First, RACE/A assumes that memory retrieval is a dynamic process, in which the probability of retrieval of information is continuously updated. Second, RACE/A assumes that memory retrieval consists of a competitive process, in which potentially relevant pieces of information influence each other's probability of retrieval. Third, RACE/A takes into consideration that memory retrievals are embedded in a larger cognitive system. Although architectures or models that share the first two assumptions are not uncommon, RACE/A is the first account in which the detailed accounts of memory retrieval are embedded in a larger cognitive system that enables explanations on all levels of processing. Thus, these three assumptions make it possible to account for a large range of phenomena with a single theory.

First, RACE/A explains how latencies are a function of the distributed presentation of stimuli. Because the probability of retrieval is continuously updated, the cognitive system can adapt its retrieval process to new information that becomes available during the retrieval process (as in the PWI model), or adapt the retrieval process when information is no longer available (as in the masked priming model).

Second, RACE/A provides an explanation for interference effects found in the Stroop paradigm. The interference that exist between the target stimulus and the distractor stimulus is explained as a competition between chunks, in which both chunks related to the target stimulus and chunks related to the distractor stimulus accumulate activation, which results in lower Luce ratios and slower retrieval. Two major characteristics of the Stroop effect, the *semantic gradient effect* and the *Stroop asymmetry*, naturally follow from RACE/A.

The semantic gradient effect refers to the finding that the degree with which a distractor is semantically related determines the size of the interference effect (Klein, 1964). For picture-word interference this is apparent from the latency difference that exist between the related and unrelated distractors (see also Figure 2.8a and W. R. Glaser & Düngelhoff, 1984; Lupker, 1979). In RACE/A, distractors that have a stronger relationship spread more activation towards each other, which leads to decreased Luce ratios and increased competition. In the PWI model and the repetition-priming model, this property of RACE/A explains the difference between the related and unrelated conditions.

The asymmetry between word reading responses and color naming (or picture naming) responses that is present in the Stroop paradigm is explained by the different processing routes that are often hypothesized for words and pictures and colors (e.g., W. R. Glaser & Glaser, 1989; La Heij, Happel, & Mulder, 1990; Roelofs, 1992). Whereas naming colors or words requires retrieval of conceptual information from declarative memory, pronouncing a word can be done without access to the meaning of that word. Thus, the longer route for colors or pictures creates an asynchronous start of accumulation for the potential responses. As words are also processed faster than colors or pictures, the correct lemma is retrieved before the color or picture concept increases the activation of associated lemmas, effectively precluding any interference effects for word reading in Stroop or PWI tasks.

The third class of phenomena that are accounted for by RACE/A relate to the interplay between long-term memory and short-term retrieval processes. Because in RACE/A memory retrievals are embedded in a general theory of cognition, RACE/A can explain the interference dynamics in complex experimental designs. For example, the different interference patterns that Experiment 1 and Experiment 2 provide is explained by assuming that participants use a different strategy in Experiment 2, which boosts the activation of the distractor in such a way that the interference is increased.

WHY NOT JUST RETRIEVAL MODELS?

Many tasks can be thought of a sequence of declarative memory retrievals. For instance, word production can be thought of as a process in which the speaker first need to retrieve a message from memory, then lexical information, then morpho-phonological information finishing with a phonetic code (cf., Levelt, Roelofs, & Meyer, 1999). Each of these stages comprises one or more memory retrievals, that all influence each other. Moreover, when producing multiple words, each new word is influenced by the memory processes of the previous words. This is similar to a simple experiment with a sequence of trials, all of which comprise one or more memory processes. Often this results in (undesired) between-trial effects.

A cognitive architecture provides a framework in which between-trial effects can be studied, because it provides a theory on how declarative memory processes interact – with each other as well as with other cognitive processes. In addition, the architecture provides a theory on the temporal dynamics of a task or a sequence of trials. For a theory in which memory retrieval is presented as an sequential sampling process with asynchronous retrieval onsets, such as RACE/A, this is a crucial feature, because it provides a validated basis for the starting point values of the sampling process. The benefits of an architecture become clear when examining the cognitive models developed for the experiments reported in this paper. The PWI model demonstrates that a competitive memory process as RACE/A proposes can account for the interference effects observed in typical PWI experiments. However, to account for the complex interference patterns observed in Experiments 1 and 2, it is necessary to incorporate the repetition priming mechanism in the model. In addition, theories of cognitive control as well as visual and aural perception as well as of manual and vocal responses are necessary to provide a quantitative prediction of the response latencies in Experiments 1 and 2.

In contrast, consider a study of a lexical decision task that had been repeated over 5 days, Dutilh et al. (in press; see also Wagenmakers, in press) found the typical speed-up in performance usually observed in practiced behavior. A diffusion model analysis of this task revealed that the speed-up could be attributed in part to the non-decision part of the response times. That is, the speed-up could be explained by processes other than the actual lexical decision, but rather by the perceptual or motor response processes that are also part of the task execution. We hypothesize that a RACE/A model analysis would have provided a similar result, but with RACE/A it would have been possible to theorize on the mechanism underlying the speed-up in non-decision time. One explanation would be that proceduralization of the motor responses would account for the speed-up. In RACE/A this process may be accounted for by production compilation (Taatgen & Anderson, 2002; Taatgen & Lee, 2003), in which multiple production rules are combined in one new one. This speeds up task execution because less production rules have to be selected in order to finish a trial.

This example shows that integrating an accumulator model of memory retrieval in a cognitive architecture may enhance the explanations derived from the model. In this case, an extra explanation for *why* the non-decision time decreases might be provided. In the cognitive models of the experiments reported above, the architectural structure provides a theory of cognitive control (Altmann & Gray, 2008; Salvucci & Taatgen, 2008) , as well as theories on perception and action. While these theories may not be extremely detailed, they provide enough detail to understand memory and decision behavior in various task contexts (Gray, 2007a).

WHY NOT JUST ACT-R?

There have been previous attempts to model the Stroop effect in ACT-R (e.g., Altmann & Davidson, 2001; Juvina & Taatgen, 2009; Lovett, 2002, 2005). Thus, the question arises why a more detailed retrieval theory is needed in the first place. However, these models describe retrieval in a Stroop task as a ballistic process, with a *retry-mechanism* that checks if a retrieved chunk matches already retrieved information; If not, the retrieval is retried. Because words are processed faster than colors, the probability of an incorrect retrieval is highest if word and color do not match. If they do match, the response to the word is also the correct response, so no retry is required. Stroop interference is in this sense a function of the mean number of retrieval attempts before the correct answer has been retrieved. Although on average this may result in correct latency predictions, the predicted latency distributions essentially consist of a set of unimodal distributions, distributed around the time required to perform one retrieval, two retrievals, and so on, plus the time required for other processing steps. This does not seem to be a correct interpretation of the Stroop effect, since generally the latency distribution is considered to be unimodally (Heathcote, Popiel, & Mewhort, 1991).

A previous more general solution to interference effects in ACT-R is competitive latency. In this account, the predicted retrieval time is scaled according to the activation of competitors. If there are multiple competitors or relatively active competitors, retrieval times are decreased (Equation 2.7).

$$RT = F\frac{\sum_{j \neq i} e^{A_j}}{e^{A_i}} \qquad \text{(equation 2.7)}$$

While in some situations this equation may give an accurate prediction of response time (e.g, Van Rijn & Anderson, 2003), there are some modeling-technical difficulties. However, the competitive latency equation remains a ballistic model of response time and is therefore not able to account for the kind of interference effect that we address in this paper. Similar to the default latency prediction in ACT-R (Equation 2.3), this model cannot account for the effects of asynchronously presented stimuli or masked stimuli.

IMPLICATIONS FOR THE STROOP EFFECT AND PICTURE-WORD INTERFERENCE

The PRP paradigm has been used to elucidate the mechanisms underlying the Stroop effect and picture-word interference (Dell'Acqua et al., 2007; Fagot & Pashler, 1992; Van Maanen, Van Rijn, & Borst, submitted). Fagot and Pashler (1992) performed a similar experiment as the Experiment 1 reported above, but instead of a picture-word interference task they used a Stroop task. In their experiment, participants were requested to classify a tone as either high or low and subsequently name the color of a Stroop stimulus. Fagot and Pashler tested two Stroop conditions, a congruent one in which word and color refer to the same color concept, and an incongruent one in which word and color refer to different color concepts. To remain consistent with our terminology, we will refer to this second condition as the related condition. The results of this experiment show that the difference between the congruent and related condition (the Stroop interference) is *not* dependent on the PRP-SOA between the tone onset and the Stroop stimulus onset. Assuming a central processing bottleneck (Pashler, 1994) or a cognitive control structure (Meyer & Kieras, 1997b; Salvucci & Taatgen, 2008), this result suggests that interference in a Stroop task takes place after the bottleneck stage, because otherwise the interference would have been absorbed in the cognitive slack time at small PRP-SOAs. This result is different from Experiment 1 as well as from (Dell'Acqua et al., 2007), were interference was found to be affected by PRP-SOA, suggesting that interference in PWI is absorbed in the cognitive slack time.

Dell'Acqua et al. argue that this difference between PWI and Stroop (absorption in cognitive slack time versus no absorption in cognitive slack time) is an indication that the mechanism that causes Stroop interference differs from the mechanism that causes picture-word interference (but see Van Maanen, Van Rijn, & Borst, submitted, for another interpretation of these results). However, the PWI conditions contrasted in our Experiment 1 and in the Dell'Acqua study are the related and the unrelated conditions, whereas the Fagot and Pashler (1992) study contrasts the related and the congruent conditions. Experiment 2 provides a new perspective on this discussion, because here a congruent PWI condition is included. Given that in Experiment 2 the difference between the congruent and related condition was not mediated by PRP-SOA, we might conclude that the difference between picture-word interference and Stroop was a difference in task conditions, and not caused by a different processing stage. However, this conclusion is still speculative, since Experiment 2 included a third condition as well. As our experiments demonstrate, the latency patterns in PRP experiments depend on the participants' expectations as much as on constraints on cognitive processing.

RACE/A shows that an accumulator model in the context of a cognitive architecture provides accurate latency predictions, that can account for the interactions that exist between competitive processes in memory retrieval, long-term learning effects, and cognitive control.

Evaluation of the RACE/A Integrated Model of Memory Retrieval

Section 3.2 of this chapter has been published as Van Maanen, L., & Van Rijn, H. (2007). An accumulator model of semantic interference. Cognitive Systems Research, 8(3), 174-181.

Section 3.3 of this chapter has been published as Van Maanen, L., & Van Rijn, H. (2007). Accounting for subliminal priming in ACT-R. In R. L. Lewis, T. A. Polk & J. E. Laird (Eds.), 8th International Conference on Cognitive Modeling (pp. 1-6). Ann Arbor, MI.

INTRODUCTION

In the previous chapter Retrieval by ACcumulating Evidence in an Architecture (RACE/A) was introduced. This theory is aimed at explaining competitive effects during memory retrieval. The competition between different memory traces is evoked by association between the competing traces. Thus, if two traces spread activation to each other and they are both possible candidates for a particular retrieval from memory, they will compete for retrieval.
Competition in memory retrieval seems to be ubiquitous in behavior. For example, every decision one makes based on previous knowledge involves memory retrieval. To demonstrate that many interference phenomena are competitive effects in memory retrieval, we will discuss a series of different experiments in this chapter that show the breadth of RACE/A. The most obvious effects that RACE/A should account for are effects in lexical decision. The accumulator models on which RACE/A is inspired are often intended at lexical decision experiments (e.g., Wagenmakers et al., 2008). We will first shortly introduce the RACE/A model of lexical decision (Section 3.1), and then elaborate on two RACE/A models that account for competitive effects that are less straightforward. These are the asynchronous presentation of stimuli (Section 3.2) and the masked or subliminal priming paradigm (Section 3.3).

If the onset of stimulus presentation is asynchronous, the interference effects that stimuli have on each other is a function of the onset asynchrony. That is, the closer the stimuli are presented in time, the more effect there is. For example, in picture-word interference, in which participants are required to name a picture, while ignoring a word that is presented at an asynchronous onset (either before or after the picture), the effect size is correlated with the onset (but not monotonically, see Section 3.1 and W. R. Glaser & Düngelhoff, 1984). While this effect is something that may be studied in a regular accumulator model, the architectural foundation of RACE/A provides a theoretical framework for the onset timing of the various memory retrievals that play a role in the picture-word interference paradigm.

If the presentation of a stimulus is only present temporary, such as if the stimulus is masked (e.g., Marcel, 1983) or in a signal-to-respond paradigm (e.g., Wagenmakers et al., 2004a), than the effect size of interference effects is also affected. The model in Section 3.3 describes this situation. The model fits data from a Stroop task in which the words are presented subliminally. That is, the words are presented for such a short time that participants report not being aware of the presence or absence of the word. The color patches on which they have to respond are continuously present. Smaller Stroop effects than in the default experiment are found under these conditions (Marcel, 1983), which is explained by RACE/A by less time to accumulate activation for the word, resulting in a higher Luce ratio for the color.

A RACE/A MODEL OF LEXICAL DECISION

In lexical decision, participants are presented with a letter string and are required to indicate whether the string forms a correct word (e.g., BALK for Dutch or CREAM for English) or not (e.g., BALC and CEARM). Lexical decision is often ascribed to a sampling process in which evidence for the hypothesis that the letter string is a word is sampled, until a threshold has been reached. If the threshold is reached, the participant will provide a "word" response. If a certain fixed amount of time (a deadline) has passed, the participant will respond "non-word" (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger & Jacobs, 1996). These models can be interpreted as a serial search of memory for a word matching the letter string.

Contrary to these deadline models, in the RACE/A model of lexical decision, the non-word responses are driven by an extra accumulator that samples the likelihood that the current letter string is not a word. This means that when time passes, the likelihood of a non-word response increases. If the Luce ratio of the non-word crosses the decision boundary (as discussed in Chapter 2), the model responds with a non-word response. If the Luce ratio of the word-response crosses the decision boundary, the model responds "word".

Because the drift rate in the model is determined by the familiarity of the word (in case of "word" responses), words that have a high frequency of occurrence in natural language are recognized faster, resulting in faster responses. Figure 3.1 presents the fit of the lexical decision model on data from an experiment by Glanzer and Ehrenreich (1979). In the experiment, Glanzer and Ehrenreich studied the difference in response time between high frequent and low frequent words, as well as the response time associated with non-word responses. They found that on average, participants respond faster on high frequent words than on low frequent words, and slowest to non-words.

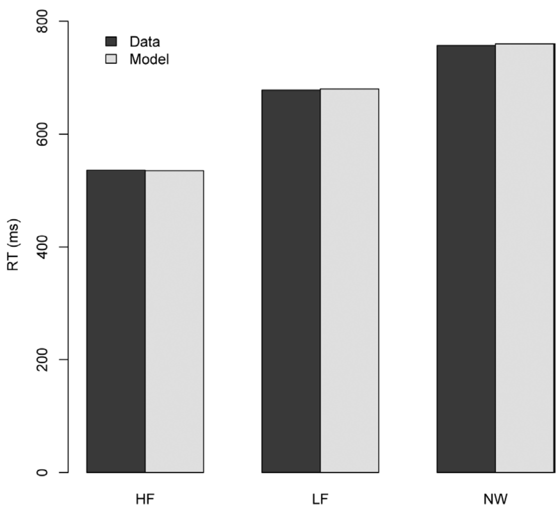


Figure 3.1. Fit of the lexical decision model to a data set by Glanzer and Ehrenreich (1979). HF: High frequent words; LF: Low frequent words; NW: non-word letter strings.

The model accounts for the basic lexical decision word frequency effect by assuming higher base-level activation values for high frequent than for low frequent words. This results in a faster decision process, and therefore shorter response times.

The model is very straightforward and in line with previous accumulator models of lexical decision (e.g., Brown & Heathcote, 2008; Ratcliff, Gomez, & McKoon, 2004; Wagenmakers et al., 2008; Wagenmakers et al., 2004a), but it accounts for an extra finding that most accumulator models of lexical decision cannot explain (but for an exception see, Wagenmakers et al., 2008).

If the stimuli are words that are very infrequent, that is, word that are extremely rare in natural language use, then participants remain able to accurately respond to these, but at a significant time cost: The average response time under these conditions exceeds the average response time for non-word responses (Wagenmakers et al., 2008). Models that incorporate a strict deadline for the accumulation process fall short here, because these would assume that the non-word responses are only given if the word search has terminated. Therefore, the mean latency of non-word responses is predicted not to exceed the mean latency of correct word responses. These model predictions are not in agreement with Wagenmakers et al.'s findings.

The RACE/A model of lexical decision does account for the effects found by Wagenmakers et al. (2008). In the RACE/A model, the probability of not retrieving a word accumulates, similar to the probabilities of finding a particular word. Because the decision on what to retrieve from memory depends on the Luce ratios of the words and the non-word representation, the response latency for very low frequent (VLF) words can be extended beyond the average non-word response time. In case of a VLF trial, the memory trace of the VLF word accumulates very slowly, but so does the non-word representation. The Luce ratio of the VLF word increments thus very slowly, increasing the decision time. In case of a non-word trial, only the non-word representation accumulates. Because no word representations (including the VLF word) interfere, the decision time may be shorter than on VLF trials (and so is the latency). By contrast, the non-word decision time is longer than the decision time for words that are more frequent. For these words, the decision time is fast because the accrual rate is much higher than the accrual rate of the non-word representation. Figure 3.2 presents the fit of the model on the data from Wagenmakers et al. (2008). Besides the median latency, the model also captures the latency distributions to some extent, as indicated by the plotted quantiles.

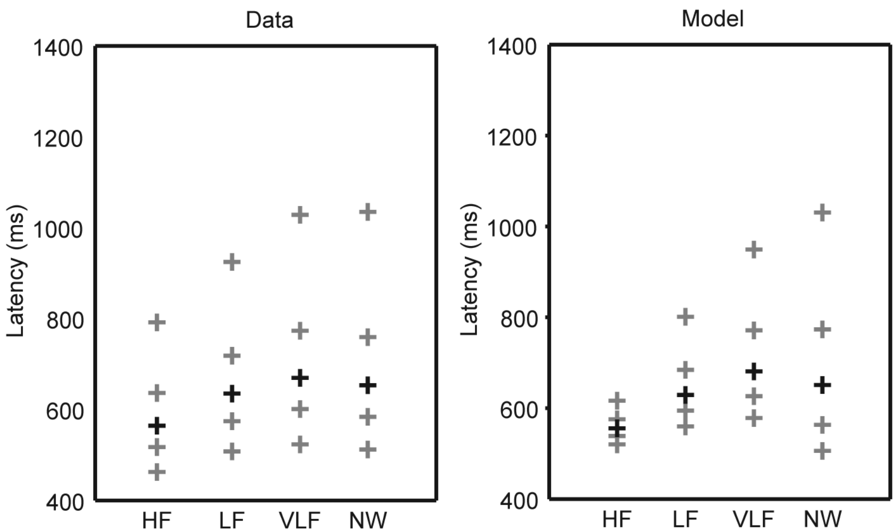


Figure 3.2. Fit of the lexical decision model to a data set of Wagenmakers **et al.** (2008). The figure shows the 10%, 30%, 50% (black markings), 70%, and 90% quantiles of the latency distributions. HF: High frequent words; LF: Low frequent words; VLF: Very low frequent words; NW: non-word letter strings.

A MODEL OF STIMULUS ONSET ASYNCHRONIES IN PICTURE-WORD INTERFERENCE

INTRODUCTION

Often, symbolic models of cognition can be thought of as giving a *stroboscopic* account of cognition. By illuminating a process such as a movement with a stroboscope, the movement is sliced into discrete steps that together represent the original, continuous, movement. However, information about the movement is lost when the stroboscope does not flash, and an observer will not be aware of how the movement develops during these brief flash intervals. Similarly, symbolic models of cognition reflect a continuous cognitive process on a higher level of analysis, but on a lower level of analysis, analogous to a single flash of the stroboscope, these models provide a discrete account of that process. In most cases, interpreting the higher level of analysis as a continuous process is sufficient for understanding cognitive functioning, but in some tasks, the underlying discrete account might provide a misinterpretation of the process.

As a general example of such a task, consider the way retrieval of memory chunks is modeled in the ACT-R architecture of cognition (Anderson et al., 2004; Anderson & Lebiere, 1998). Retrieval latency is based on the activation of the to-be-retrieved memory chunk:

$$RT_i = Fe^{-A_i} \qquad \text{(equation 3.1)}$$

Equation 3.1 states that the retrieval time (*RT*) of a chunk (*i*) is inversely proportional to the exponentionally scaled activation of that chunk (A_i), with *F* a scaling parameter. If a retrieval request is made to the declarative memory system, the activations of all chunks are compared, and the highest is selected for retrieval. The latency is calculated according to the above equation and, after the appropriate amount of time has passed, retrieval of that chunk is reported. Even if new information is presented between the retrieval request and the actual retrieval, the retrieval result and latency cannot be influenced.

However, many experiments show that information that is presented shortly before or after a target stimulus can influence both the timing and accuracy of the task at hand (e.g., MacLeod, 1991; Neely, 1991). In a picture-word interference task for example, participants respond slower in the picture-naming task when a distractor word is presented, even if that distractor word is presented shortly *after* the target stimulus.

Since ACT-R has been successfully applied to numerous memory related tasks (e.g., Anderson et al., 1998; Pavlik & Anderson, 2005; Taatgen & Anderson, 2002), it should also provide an explanation of picture-word interference phenomena. However, given the ballistic nature of the way memory retrieval is currently modeled in ACT-R, the question becomes how ACT-R can be extended to include interference phenomena on very short latencies. In this paper, we will present a means to extend the ACT-R architecture of cognition to incorporate these interference effects. While we extended the memory system of ACT-R, we have made sure that the main characteristics of the tested and proven declarative memory equations were not altered. This way, we made sure that our approach towards semantic interference fits in with a broader theory of cognition, while at the same time we add a new phenomenon to the subset of cognition that ACT-R can account for.

A candidate explanation for semantic interference effects comes from the field of choice behavior modeling. In sequential sampling models of simple choice behavior, the choice between candidates is modeled by competition between candidates. Sequential sampling is based on the idea that choosing one option over the other is based on sampling of inherently noisy neural representations of these choices, until one has sampled enough evidence to be chosen (Ratcliff & Smith, 2004). The RACE (Retrieval by ACcumulating Evidence) model

presented in this chapter is very similar to a specific instance of sequential sampling models: The leaky competing accumulator model as discussed by Usher and McClelland (2001).
RACE is implemented using the same basic principles as the leaky competing accumulator model: (a) it consists of a set of non-linear stochastic accumulators, all of which represent one memory chunk that can be retrieved. (b) The activations of the accumulator units are increased by external input and recurrent activation, but are decreased by lateral inhibition and decay. However, the actual implementation of some aspects differs, most importantly different activation and evidence accumulator functions, both of which have been adapted to fit RACE in the ACT-R framework.

RACE ARCHITECTURE

The name RACE (Retrieval by ACcumulating Evidence) reflects both the accumulation of evidence for memory representations and the competition between memory chunks during retrieval: The comparison with a race between chunks seems appropriate in this respect.
The activation levels of memory chunks in RACE consist of two components: A long-term component that governs the global activation of chunks and a short-term component that comes into play during the retrieval process. The long-term component is represented by the ACT-R base-level activation equation (Anderson et al., 2004):

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \qquad \text{(equation 3.2)}$$

where t_j is the time since the jth presentation of a memory chunk and d is the parameter that controls decay, which is fixed at 0.5, as is common practice for ACT-R models (Anderson et al., 2004). The idea is that memory decays over time unless attention is shifted to a memory chunk and its activation is strengthened.

RACE's short-term component, called accumulated activation (C_i, to avoid confusion with the general symbol for activation A_i used in ACT-R), is continuously computed from the moment that a request for retrieval of a chunk is made. The accumulated activation of chunks changes as a consequence of positive and negative influences from other chunks. Chunks from the same chunk type inhibit each other, thereby competing for accumulated activation increase. Chunks of different chunk types excitate each other, spreading their activation in the classical sense (Collins & Loftus, 1975). Thus, by continuously updating positive and negative spreading activation, some chunks may reach a level of activation at which retrieval can take place.

The accumulated activation can be described as a system of two dependent equations (Equations 3.3 and 3.4 presented below). As stated earlier, these equations incorporate the basic assumptions of Usher and McClelland (2001), but are adapted to fit in the ACT-R framework.

$$E_i^k(t) = \sum_{j \notin k} e^{A_j(t-1)} S_{ji} - \sum_{l \notin k} e^{A_l(t-1)} S_{li} \qquad \text{(equation 3.3)}$$

The system functions as follows: At every time step, positive associative values (reflected by the first term of Equation 3.3) and negative associative values (second term of Equation 3.3) towards a memory chunk are computed and the difference is calculated. This is called the net evidence ($E_i^k(t)$) of chunk *i* of chunk type *k* at a certain time *t*. Since relative - not absolute – activation values are what count in ACT-R, an exponential scaling is applied to calculate net evidence. Also, both positive and negative associative values are weighted by the associative strengths (S_{ji} and S_{li}) that exist between sources of activation and the chunk *i*. There are two types of sources of activation in RACE: Chunks (*l* in Equation 3.3) of the same chunk type (*k*) spread negative activation to each other, while chunks (*j*) of different types spread positive activation. This is analogous to neurobiological findings from which it is clear that lateral inhibition between cortical representations of visual stimuli (Kastner, De Weerd, Desimone, & Ungerleider, 1998) as well as excitatory projections to other cortical layers (Callaway, 1998) exist. Note that most ACT-R models do not place constraints on the functional role of chunk types (although it does pay a role in production compilation, Taatgen, 2005).

$$C_i(t) = C_i(t-1) + e^{\beta E_i(t)-1} - d^{\mathrm{acc}} \cdot \ln T \qquad \text{(equation 3.4)}$$

At each point in time, the net evidence determines the accumulated activation growth (Equation 3.4). Accumulated activation increases exponentially according to the amount of net evidence and a scaling factor β. If net evidence is negative (that is, more inhibition than excitation), then growth is negative. At all time steps, evidence decays with (represented by the second term of Equation 3.4), in which T is the time since the start of the accumulation and d^{acc} a decay parameter. This way, accumulation decay in RACE resembles decay in the ACT-R optimized learning equation (Anderson & Lebiere, 1998).

The activation of a chunk at any time is the sum of base-level and accumulated activation, plus a small normally distributed noise sample. A chunk is retrieved if this total activation crosses the accumulation threshold.[5] The retrieval latency is defined as the time between the retrieval request and the time that the total activation of a matching chunk reaches this accumulation threshold.

If no evidence is sampled, accumulated activation decreases because of decay. Therefore, continuous evidence-based positive reinforcement is necessary for successful retrieval, and absence of positive evidence results in prolonged retrieval latencies or retrieval failures.

5. The accumulation threshold is a different concept from the retrieval threshold in default ACT-R. Where the retrieval threshold determines the minimum activation at which a chunk may be retrieved, the accumulation threshold determines the amount of activation at which a chunk is retrieved.

PICTURE-WORD INTERFERENCE

One of the most well-known experimental paradigms in cognitive psychology is the Stroop-task (Dyer, 1973; Stroop, 1935), where, in the original setup, participants have to either name the color a word is written in, or read the word, which is always a color name. It turns out that naming the color is much more difficult than reading the word – especially if color and word of a single stimulus do not correspond – as is reflected in increased reaction times and decreased accuracy in the color naming condition. The Stroop-task can be regarded as an instance of a more general class of experiments that demonstrate interference effects in various naming tasks between pictorial stimuli and word-form stimuli. These experiments are generally called picture-word interference experiments (W. R. Glaser & Düngelhoff, 1984; MacLeod, 1991). In the case of the Stroop-task, the pictorial stimulus is the word color.

We tested the RACE model in a picture-word interference task, using two tasks and four different conditions, similar to the experimental setup in Glaser and Düngelhoff (1984, Experiment 1). One task consisted of reading a word (target stimulus) while a picture is presented as distractor; the other task consisted of naming the depicted item (target stimulus), while a word is presented (distractor). In both tasks, the distractors were presented at different SOAs (Stimulus Onset Asynchronies). If a distractor was presented at a negative SOA, it was presented before the target stimulus. At positive SOAs, the distractor was presented after the target stimulus. Figure 3.3 presents stimuli examples of the different conditions. The first condition (Figure 3.3a) was one in which both target and distractor stimulus refer to the same concept. This is referred to as the concept-congruent condition. In two other conditions, target and distractor stimulus refer to different concepts. In the category-congruent condition the concepts belong to the same semantic category (e.g., a picture of a house and the word *church* were presented, Figure 3.3b), in the incongruent condition the concepts do not belong to the same semantic category (e.g., a picture of a house versus the word *cat*, Figure 3.3c). In the neutral condition the target stimuli were accompanied by non-word or non-picture distractors, respectively, to minimize the amount of processing of the distractor stimulus (Figure 3.3d and 3.3e).

Glaser and Düngelhoff (1984) found that interference is highest in the category-congruent condition, which is known as the semantic gradient effect. They also showed facilitation in the concept-congruent condition, meaning that latency is decreased when both target and distractor stimuli refer to the same concept. A third effect they report is a clear asymmetry between the picture naming task and the word reading task. The semantic gradient and facilitatory effect virtually disappear in the word reading task, but they are prominent in the picture-naming task.

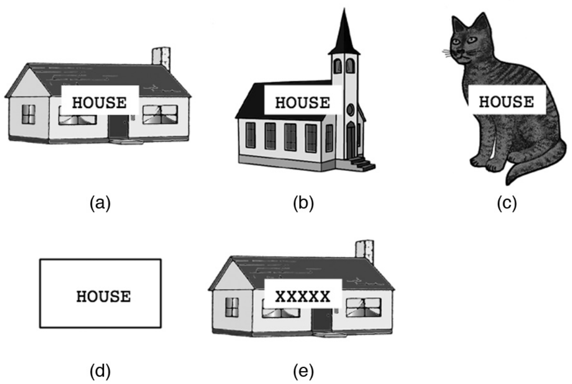


Figure 3.3. Example stimuli. (a) Concept-congruent condition, (b) category-congruent condition, (c) incongruent condition, (d) neutral condition in word reading, (e) neutral condition in picture naming. The example images are taken from the image set by Rossion and Pourtois (2004).

PICTURE-WORD INTERFERENCE MODEL

We will begin our discussion of the picture-word interference model with a review of the WEAVER++ model of speech production (Levelt, Roelofs, & Meyer, 1999). WEAVER++ is similar to RACE in some ways, but lacks the integration in a cognitive architecture that we provide. WEAVER++ has a similar memory structure as RACE, and a similar activation accumulation mechanism as RACE. In WEAVER++ however, it remains unclear how the model is connected with other aspects of cognition besides language production. Although WEAVER++ and previous versions of that theory have been demonstrated to fit an impressive number of data sets (e.g., Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992, 1997, 2003), it lacks a unified account of cognition, that for instance ACT-R does provide. RACE's integration in the ACT-R framework ensures that our account can be naturally integrated in models of other aspects of cognition. One example of this is the subliminal priming model described by Van Maanen and Van Rijn (2007a).

Word production in WEAVER++ goes through a sequence of stages, one of which is the retrieval of the to-be-spoken word from semantic memory. In WEAVER++ this response selection stage (choosing a lemma) is followed by response programming and execution stages. Since our focus has been on the retrieval process, these vocalization aspects of the task are not included in our model of picture-word interference. The RACE mechanism is similar to the mechanism proposed for the lexical selection stage in WEAVER++ (1992).

The lexical processing stage from WEAVER++ is modeled as follows: A network of conceptual nodes is connected to a network of lemma nodes. The conceptual nodes convey meanings, and are connected with labeled links. For instance, The concept DOG(X) represents the meaning of the noun *dog*, and has a labeled connection of the type IS-A to the concept node ANIMAL(X), indicating that a dog is an animal (Roelofs, 1992). The nodes in the lemma network represent the syntactical dependencies of the concept nodes. Each lemma node has a labeled SENSE link to the corresponding concept node, labeled links to syntactic properties – grammatical gender, syntactic category. The links between concept nodes and between concept and lemma nodes differ in their connection strength, indicating a difference in accessibility. Via a spreading activation mechanism, activation of one node influences activation of neighboring nodes. Activation is also mediated by decay.

If the ratio of the activation of one lemma node against the activations of the others exceeds a predefined (relative) criterion, selection of that lemma node takes place, and WEAVER++ will proceed with the retrieval of the morpho-phonological properties of that lemma.

Analogous to WEAVER++, our model of picture-word interference comprises three chunk types (Figure 3.4): Icons, lemmas, and concepts. The concept chunks can be regarded as representations of semantic properties. Chunks of the icon type represent iconographic instances of the stimuli. This might be similar to Roelofs (1992) object-form memory store. Chunks of the lemma type can be regarded as sets of both orthographic and syntactic properties of a word. Note that this is a simplification of Roelofs' (1992) model, in which the response selection stage (choosing a lemma) is followed by response programming and execution stages. Since our focus is on the retrieval process, these vocalization aspects of the task are not included in our model of picture-word interference.

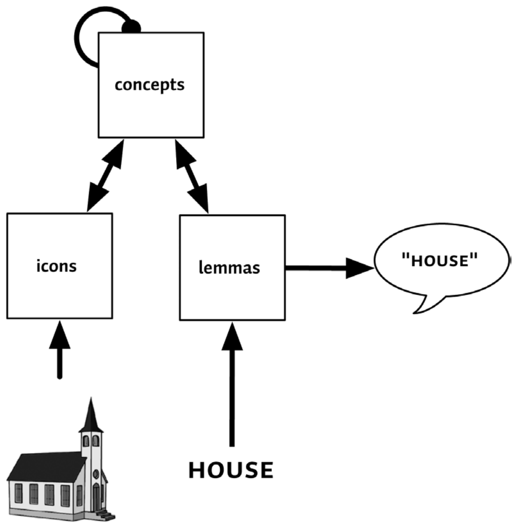


Figure 3.4. Processing route for pictures and words in the picture-word interference model. The route for words is shorter, since words do not require concept retrieval. Interference takes place between concept type chunks.

Positive activation is spread between chunks of different types. That is, icons spread to concepts and vice versa, and lemmas spread to concepts and vice versa. As in Roelofs' (1992) model, no direct spreading activation was allowed between lemmas and icons. The concept chunks also have negative associations between them and spread negative activation to each other.

At different SOAs, distractor stimuli were presented to the model, except in the neutral condition in which only a target stimulus was presented. We tested the same four conditions as Glaser and Düngelhoff (1984) did. The only deviation from the original experiment was the neutral condition: Glaser and Düngelhoff presented the participants with a non-word distractor and a non-picture distractor respectively in the picture naming and word reading neutral condition. As said, these were chosen in such a way as to minimize the amount of picture or word processing as possible. Assuming a successful operationalization by Glaser and Düngelhoff, we simulated this condition by not presenting a distractor in the neutral condition.

In the concept-congruent condition, the distractor consisted of a word stimulus referring to the same concept as the target, but of a different stimulus type (as in Figure 3.3a). When activation spreads through the model, both distractor and target activate the same chunks, but not in the same order. The word will activate its associated lemma directly, whereas the picture will first activate the associated icon and concept chunks. In the category-congruent condition and the incongruent condition, the distractor and the target refer to different concepts. However, in the category-congruent condition, associations between chunks representing these concepts exist, reflecting the fact that the target and distractor stimuli belong to the same semantic category.

The distractors were presented at SOA times relative to the onset of the target stimulus of -400, -300, -200, -100, 0, +100, +200, +300, and +400ms, similar to the original Glaser and Düngelhoff experiment. The stimuli presentations were modeled as a fixed increase in activation of the lemma or icon type chunks during the period that a stimulus was presented.

Since the task was a verbalization of either the picture name or the word, a trial was finished when the stimulus-designated lemma was retrieved or after two seconds, indicating a retrieval failure.

In the picture naming task, the model predicts the following behavior: In the concept-congruent condition with negative SOAs, a distractor word is presented before the target picture. The word activates a lemma chunk, which increases the activation of the associated concept chunks, but inhibits the activation increase of other lemma chunks. The higher activation of the concept chunks increases the activation of the associated icon chunks. Thus, after the distractor is presented, all chunks that are involved in naming the picture (one icon, one concept, and one lemma chunk) have an increased activation. When the target is presented, all concept-congruent chunks have a higher activation as compared to the stimulus onset in the neutral condition, and thus a shorter retrieval latency. In the concept-congruent condition with positive SOAs, the same process occurs, but to a lesser extent since the distractor lemma's activation has less time to influence the activation of the target lemma before it is retrieved: The picture has already increased the target lemma's activation before the word is presented.

In the incongruent and category-congruent conditions (both at negative and positive SOAs), the activation of the chunks that are activated by the distractor interferes with the activation of the chunks that are activated by the target, because the target and the distractor stimulus activate different sets of chunks.

RESULTS

Figure 3.5 summarizes the results of our simulation studies.[6] The figure represents the latency differences in different conditions relative to the neutral condition. Since the focus of our model is on the retrieval part of the picture naming and word reading tasks, we can compare the latency differences between the different conditions from the model to the data. The observed latencies from the data set also comprise timing effects from other subtasks, such as pronunciation or perceptual encoding.

Negative values in Figure 3.5 indicate faster retrievals than in the neutral condition, and positive values indicate slower retrieval than in the neutral condition. The qualitative effects

6. An R implementation of RACE and the picture-word interference model can be retrieved from http://www.ai.rug.nl/cogmod/models.

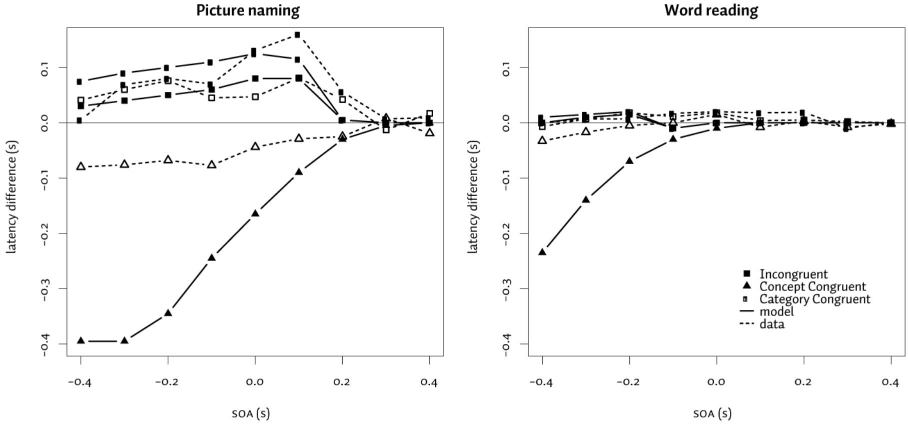


Figure 3.5. Simulation results and experimental data (Glaser & Düngelhoff, 1984, Experiment 1) for the picture naming task and the word reading task.

observed in Glaser and Düngelhoff (1984 Experiment 1) can be seen in the predicted latency differences from the RACE model. The semantic gradient effect can be observed by noting the different relative latencies of the category-congruent and incongruent conditions. The higher latencies in the category-congruent condition compared to the incongruent condition indicate that higher associations between concepts result in stronger inhibition.

The facilitatory effect in the picture-naming task is also apparent, although the effect appears to be too large. Our explanation for this increased effect is that the activation of the target lemma chunk is too high if the target stimulus is presented, probably caused by too little decay after the previous retrieval initiated by the distractor stimulus. Another consideration might be that the base-level activation goes to near infinity directly following a retrieval, thus causing too much increase in activation of the target lemma chunk. This may also explain the observed effect in the simulation of the word-reading task. Thus, it seems that the base-level activation may not be a good measure of the level of activation of a chunk at these short time intervals. Because RACE is intended as an extension of the ACT-R framework, it did not seem appropriate to change the way in which the global, long-term activation is computed.

When comparing the differences between the two simulated tasks (picture naming and word reading), the asymmetry observed by Glaser and Düngelhoff is also shown by RACE. We explain this asymmetry by two effects: The shorter processing route and the faster encoding of word type stimuli. As Roelofs (W. R. Glaser & Glaser, 1989; La Heij, Happel, & Mulder, 1990; 1992) noted, pronouncing words does not require retrieving a concept from memory, therefore processing word type stimuli can be much faster than processing pictorial stimuli. In Figure 3.3 it can be seen that the route in our model from a picture to the associated lemma is much longer than from the word to the associated lemma: Two intermediate steps have to be taken (that is, processing of icon chunks and processing of concept chunks), before the lemma chunk is retrieved. Thus, potentially interfering pictures do not activate lemma chunks before the target lemma is retrieved. Only at high negative SOAs an effect can be seen (Figure 3.5), because under that condition there is enough time for the distractor stimulus to activate the inhibiting distractor lemma and interfere with retrieval of the target lemma.

A difference in encoding speed is incorporated to account for the observation that word recognition is an automated process and picture recognition is not. Without this difference, the model processes pictures nearly as fast as words, and picture naming in the incongruent condition is as fast as in the neutral condition. The faster encoding of words reinforces the effect that the lemma associated with a word is processed before the lemma associated with the picture stimulus is retrieved.

DISCUSSION

We have shown that a sequential sampling model can account for the time course of memory retrieval during asynchronously presented stimuli. This is an extension of the results from Usher and McClelland (2001) in which they only investigated accumulators with equal onset times. Moreover, RACE can be regarded as an extension of the ACT-R theory of cognition. It combines the long-term base-level activation equation of ACT-R with a short-term accumulated activation used for retrieval. RACE replaces the retrieval mechanism in ACT-R represented by Equation 3.1.

The general fit of our model of picture-word interference is quite reasonable, thereby indicating that the RACE equations can provide for semantic interference effect in memory retrieval. Retrieval in the concept-congruent condition seems to be too fast, however. We hypothesize that this is a result of the way we modeled the global, long-term activation component, namely by using ACT-R's base-level activation equation. Because in the concept-congruent condition the target and distractor stimuli both refer to the same chunks, retrieval of these chunks - caused by the distractor stimulus - increases their base-level activations. The high activation of these chunks will result in a very short latency for the retrieval caused by the target stimulus. It seems that the base-level activation equation is a better predictor of activation at slightly larger time scales, when the retrievals are more spaced. This is supported by the observation that retrievals in most ACT-R models are temporally more separated than ~100ms. Further research in the interaction between the base-level activation and accumulated activation seems necessary to correct for the fast retrievals in the concept-congruent condition.

Also, our model accounts for facilitatory effects. In line with the findings from Glaser and Düngelhoff (1984), RACE predicts that semantic facilitation occurs if target and distractor both refer to the same concept.

In the past, ACT-R models of semantic interference effects have been proposed (Altmann & Davidson, 2001; Lovett, 2002, 2005). The WACT model (Altmann & Davidson, 2001) seems similar to RACE at first sight, since it combines ACT-R with insights from the WEAVER++ model. However, WACT describes retrieval in a Stroop task as ballistic, but with a *retry-mechanism* that checks if a retrieved lemma chunk matches already retrieved conceptual information; If not, retrieval is retried. Thus, WACT accounts for inhibitory effects by multiple retrievals caused by retrieval failures. As such, WACT is a perfect example of a stroboscopic account of cognition. Retrieval latency for one trial can be the latency associated to one retrieval attempt, or two retrieval attempts, or many, but nothing in between. Therefore, the distribution of reaction times predicted by WACT is clustered around the time it takes for one or multiple retrieval attempts. This does not correspond with the general assumption that participants' reaction times in the Stroop-task are unimodally distributed (Heathcote, Popiel, & Mewhort, 1991).

NJAMOS (Lovett, 2002, 2005) is an ACT-R model of the Stroop task that theorizes that the Stroop effects are due to utility differences in the production rules for word and picture recognition. In the model, a general production rule is assumed that fires if a stimulus is present that has some word-like qualities, irrespective of the current task (color naming or reading). In addition, a more specific production rule is assumed that only fires if the task is color naming. The second rule has a lower utility than the first, meaning that the system has a preference to execute the first rule over the second. Therefore, in most cases (because of noise over the utility values), the second rule will only be selected after the first production has fired. In those cases the first rule has not completed the task successfully, because the task was color naming, not word reading. The second rule thus has to fire to complete the task. This two-step procedure for color naming is intended to explain the Stroop asynchrony between reading and color naming.

If the color and the word are congruent, the chunk that encodes the word-like features of the stimulus spreads activation to the chunk that encodes the color information. The activation of that chunk will increase, resulting in the facilitatory effect at negative SOAs. In the incongruent condition, negative spreading activation is introduced to explain the interference effects.[7]

7. Note however that this is an undiscussed deviation from standard ACT-R, where spreading activation is intended to be positive because it represents the increased likelihood of needing one chunk when another chunk is present.

At small positive SOAs (e.g., +100ms), NJAMOS also predicts a divergence between the latencies for different conditions, although smaller than that observed in the data (Lovett, 2002). Given the ballistic retrieval latency equation of ACT-R, it seems that these latencies can only be explained by averaging over several trials. That is, either one (the general rule) or two production rules (both general and specific) will fire, resulting in a bimodal distribution of the data. Again, the distribution of the Stroop-latencies does not seem bimodal (Heathcote, Popiel, & Mewhort, 1991).

We suggest that not fully processed words at small positive SOAs might explain this difference between model and data. Perhaps a combination of the utility-based explanation Lovett proposes combined with RACE will produce a better fit to the data.

The picture-word interference experiment shows that the RACE model can be a useful extension of the ACT-R architecture of cognition. However, one crucial feature of RACE is not supported by the ACT-R architecture. In RACE, all chunks in declarative memory spread activation to all other chunks. ACT-R assumes that only chunks that are presently in the buffers spread their activation (Anderson et al., 2004).[8] Global spreading-activation was not included in the architecture because it appeared that no second-order priming effect exists, indicating that spreading of activation through declarative memory was not necessary (Anderson, 1990). However, more recent evidence suggest a second-order priming effect, although very weak, that cannot be explained by assuming only first-order associations between prime and stimulus (Livesay & Burgess, 1998). Therefore, we consider this deviation of ACT-R theory reasonable.

8. In ACT-R 6.0, chunks in all buffers can spread activation, as opposed to ACT-R 5.0, in which only chunks in the goal buffer could be a source of activation.

Experiments using subliminal primes indicate that priming may also occur when a prime is not fully processed (Marcel, 1983; Merikle, Smilek, & Eastwood, 2001), which hints that priming already occurs before chunks in the buffers are fully identified. A dynamical activation mechanism such as RACE may provide accurate modeling accounts for this observation. In RACE, activation of chunks – either in the buffers or in declarative memory – always affects the activation of other chunks, even before the accumulation threshold is reached and a chunk might be retrieved. A RACE model of a subliminal priming task has been shown to account for retrieval latencies typically observed in these kinds of tasks (Van Maanen & Van Rijn, 2007a).

This section demonstrates how RACE can account for picture-word interference phenomena. We believe that RACE can account for all effects that involve semantic interference or facilitation. Using RACE, an explanation of these effects can be provided on a higher level of abstraction than connectionist modeling, because it is integrated in a full cognitive architecture. This way, RACE combines insights from multiple levels of abstraction.

A MODEL OF SUBLIMINAL PERCEPTION

INTRODUCTION

Successful behavior depends for a large part on having declarative knowledge available at the right time. Humans are therefore continuously retrieving declarative facts from long-term memory storage, based on their continuously updated perception of the environment. The continuous character of perception is reflected in the memory retrieval process, as can for instance be observed in the retrieval latencies of psychonomic experiments in which stimuli are asynchronously presented (e.g., picture-word interference, W.R. Glaser & Düngelhoff, 1984) or in experiments in which the presentation durations of stimuli are manipulated (e.g., subliminal priming, Marcel, 1983). A cognitive model of declarative memory retrieval should also reflect the continuous character of the input on which memory retrievals are based. However, current cognitive architectures such as ACT-R (Anderson et al., 2004) or Soar (Newell, 1990) cannot satisfactorily account for this (Van Maanen & Van Rijn, 2006; Van Maanen & Van Rijn, 2007b). Retrieval by ACcumulating Evidence (RACE) is a model that does describe the *process* of retrieving one or more chunks of information from memory. In RACE, memory retrieval is not considered ballistic, but is rather thought of as a process in which the likelihood that a piece of information will be needed for successful behavior is continuously estimated. Therefore, the likelihood estimate can be continuously adapted to the changing environment.

RACE can be perceived as an interaction of ideas from cognitive architectures that rely on symbol manipulation (Anderson et al., 2004; Newell, 1990) and ideas from sequential sampling models (Ratcliff & Smith, 2004; Usher & McClelland, 2001). The architectural nature is clear from the cognitive constraints imposed on RACE. In the current implementation of the theory, we constrained RACE by adopting the rational approach that is intrinsic to the ACT-R cognitive architecture (Anderson et al., 2004). However, the *subsymbolic* computations that drive declarative memory retrieval are rooted in sequential sampling.

This paper will describe how RACE is implemented in the ACT-R architecture of cognition and will present a RACE model of a subliminal priming task. We will discuss which features of RACE naturally align with ACT-R, and which features of RACE seem to contrast with ACT-R. We chose to implement RACE as an extension to ACT-R because of ACT-R's widespread use in the cognitive modeling world (see for instance the web site of the ACT-R community: http://actr.psy.cmu.edu). More importantly however, adopting an existing general approach towards cognition will reduce the proliferation of different cognitive theories (Newell, 1990), and will constrain theorizing about RACE. A third reason for choosing ACT-R as a modeling framework is that the way ACT-R defines retrieval latency has difficulties with modeling semantic interference (Van Maanen & Van Rijn, 2006; Van Maanen & Van Rijn, 2007b). Extending ACT-R with RACE might solve this issue.

ACT-R

A prominent theory that explains behavior at the symbol manipulation level is the ACT-R architecture of cognition (Anderson et al., 2004). Because RACE is implemented as an extension to ACT-R, we will give a very short overview of the architecture, concentrating on these aspects of the theory that relate to declarative memory retrieval.

ACT-R is a cognitive theory in which production rules operate on declarative memory and the environment. Production rules are conditions-actions pairs whose actions are executed if their conditions are met. To determine which production rule's actions will be executed, ACT-R contains a set of *buffers* of which the content is matched against the conditions of each production rule. If multiple production rules are applicable – meaning that, given the buffer contents, multiple sets of actions may be performed – the production rule with the highest utility will be selected, a process called *conflict resolution*. By default, the buffers represent the current goal of the system, the current perceptual state, and a declarative fact that is currently in the focus of attention, that is, that is recently retrieved from long-term memory. Other buffers may be defined if necessary for the task at hand (as has for instance been done for prospective time interval estimation, Taatgen, Van Rijn, & Anderson, 2007). The content of a buffer is a chunk: a symbolic unit that represents a simple fact, such as *The capital of Canada is Ottawa*, or *The object I am attending is green and spherical*. Both these example chunks are declarative facts, but the first example can typically be found in the *retrieval buffer*, and represents a fact that has been retrieved from long-term memory, whereas the second example represents a visually observable fact of the world, and might be present in the *visual buffer*. In the context of this paper, we are primarily interested in the way ACT-R incorporates retrieval of chunks from long-term memory, although we not necessarily want to constrain RACE to declarative memory retrieval.

All chunks have an activation level that represents the likelihood that a chunk will be needed in the near future. The likelihood is in part determined by a component describing the history of usage of a chunk called the *base-level activation* (B_i in Equation 3.5).

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \qquad \text{(equation 3.5)}$$

In this equation, t_j represents the time since 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). The idea is that the activation of a chunk decays over time unless attention is shifted to that chunk and its activation is increased. This way, the base-level activation can be used to model both forgetting and learning effects (Anderson & Schooler, 1991).

The total activation is the sum of the base-level activation and another component describing the influence of the current context (*spreading activation*, Equation 3.6). The spreading activation component is the sum of strengths of association from chunks *j* to chunk *i*, weighed by W_{kj}, representing the importance of various buffers (*k*) and of associated chunks (*j*).

$$A_i = B_i + \sum_k \sum_j W_{kj} S_{ji} \qquad \text{(equation 3.6)}$$

A more detailed description of the ACT-R cognitive architecture is provided in (Anderson et al., 2004; Anderson & Lebiere, 1998).

RACE MODEL OF MEMORY RETRIEVAL

RACE is a proposal for a new retrieval mechanism in ACT-R. In RACE, retrieval of a chunk is thought of as a process in which the likelihood that a chunk will be needed given the current context is continuously estimated. This is different from ACT-R, were the context can influence the retrieval of a chunk only at the onset of a particular retrieval request. Note that the continuous aspects of ACT-R's base-level learning equation (Equation 3.5) are retained in RACE. The continuous updating of context-based activation is similar to the account presented in the leaky competitive accumulator model described by Usher and McClelland (2001).

Also similar to ACT-R, the accumulation process in RACE is influenced by various sources of evidence. Increases in activation may be caused by the current context, which may be formed by the current buffer contents, or other chunks that are currently active. Via a spreading activation mechanism these chunks provide evidence for the likelihood that other chunks will be needed. That is, they increase the activation of these chunks.

Another source of evidence for the likelihood that a chunk will be needed is the history of usage of that chunk. Frequently or recently used chunks are more likely to be used again in the near future. In RACE, this is reflected by the starting point of the accumulation process. The level of activation at which accumulation starts is determined by the base-level activation of ACT-R, which reflects the frequency and recency of the usage of a chunk (Anderson & Schooler, 1991). To preserve the temporal nature of the evidence for a chunk, the accumulated RACE activation is subject to continuous decay. Activation of a chunk thus decreases if not enough evidence for that chunk is present. Since the context may change over time, the accumulation process is not determined when a retrieval process is initiated (the retrieval onset), but may also change. Therefore, incoming information or the removal of information from the buffers may influence which chunk will be retrieved.

Activation values represent the *relative* likelihood that a chunk may be needed (Anderson & Lebiere, 1998), which means that the level of activation at which a chunk has been retrieved should also be defined *relative* to the activation of other chunks. Therefore, RACE uses a *retrieval ratio* that determines how much the activation of a particular chunk must stand out against the total activation of all competing chunks. This is analogous to the relative stopping rule described by Ratcliff and Smith (2004, cf., ACT-R's former competitive latency mechanism, discussed in Van Rijn & Anderson, 2003). If multiple chunks match the criteria of the retrieval request, the chunk that reaches the retrieval ratio first will be retrieved. In these cases, the eligible chunks compete for retrieval. If the activation levels of multiple chunks increase, the total activation of the system also increases, making it more difficult for a chunk to reach the retrieval ratio. This feature of RACE will prove to be important in explaining differences in retrieval latency, for example in the model of subliminal priming explained later in this paper.

So far, we described the general idea of the RACE model of memory retrieval. In this section, the exact implementation of RACE will be presented and how RACE relates to the ACT-R architecture.

The accumulated activation component of RACE is described by the following equation:

$$C_i(t+\Delta t) = d^{\text{acc}} C_i(t) + \beta \sum_{j \in k} C_j(t) S_{ji}$$

(equation 3.7)

This equation reflects the idea that the accumulated activation of a chunk at a certain moment in time ($C_i(t+\Delta t)$) is determined by the level of accumulated activation one time step ago ($C_i(t)$), summed with spreading activation from other chunks; that is, the accumulated activation of other chunks ($C_j(t)$) weighed by strengths of association between these chunks and the chunk i (S_{ji}). At retrieval onset, accumulation starts with the history-based evidence, which is the current base-level activation. Thus

$$C_i(\text{retrieval onset}) = B_i(\text{retrieval onset})$$ (equation 3.8)

Accumulated activation decays away, the speed of which is controlled by the parameter d^{acc}. A smaller value of d^{acc} results in faster decay. The parameter β in Equation 3.7 controls the amount of influence of the context. Although in ACT-R activation can have a negative value, we have chosen in our current implementation to ignore the spreading activation from very small – that is, negative – activation values for reasons of computational efficiency.

By continuously updating spreading activation towards a chunk, the chunk may reach a level of activation at which retrieval can take place. The time at which retrieval takes place is the first moment after the start of accumulation at which the following inequality holds:

$$\frac{e^{A_i}}{\sum_j e^{A_j}} \geq \theta$$

(equation 3.9)

This means that for a chunk to be retrieved (*i* in Inequality 3.9) the activation should be high with respect to all competing chunks (*j*). Because ACT-R activation values represent the *relative* likelihood that a chunk will be needed, an exponential scaling is applied to eliminate effects from possible negative values, as is common in ACT-R equations.

Perhaps a clarification is needed on the notions base-level activation (B_i, defined in Equations 3.5 and 3.8) and accumulated activation (C_i in Equation 3.7). To incorporate frequency and recency effects in the retrieval process, the accumulation of activation starts at the current level of base-level activation (Equation 3.8). During a retrieval process however, activation is estimated according to Equation 3.7. At retrieval, the base-level activation of the retrieved chunk is also increased to account for the recent encounter with the retrieved chunk, because at the next retrieval attempt the base-level activation is again used as the starting value of the accumulation process.

The question arises which of the two activations (B_i or C_i) is a better predictor of the likelihood that a chunk will be needed. We believe that at very short time intervals – such as the SOAs from the subliminal priming experiment discussed below – accumulated activation better aligns with the empirical data. However, at longer time intervals, base-level activation has been shown to give good predictions (e.g., Anderson et al., 1998; Anderson & Schooler, 1991). Because in the subliminal priming task and model described below prime and target are retrieved in a very small time window, focusing on accumulated activation only will suffice to model the priming effects. Therefore, for this model the base-level activation values were kept constant over all chunks.

SUBLIMINAL PRIMING

In this section, we will discuss the task we modeled using RACE: a subliminal priming study by Marcel (1983). Also, we will discuss why this particular task is interesting given the specific nature of RACE. In subliminal priming tasks, primes are presented that are not consciously perceived by the participant. Usually, primes are presented for a very short period and are followed by a visual mask, so that participants can not discriminate between the presence and absence of a prime (Marcel, 1983; Merikle, Smilek, & Eastwood, 2001). Marcel (1983) showed that under these circumstances priming effects persisted. His Experiment 3 describes a Stroop-task in which words are presented as primes, and color patches are presented as cues. Participants had to respond to the color patches by pressing a button associated to one of the colors. He found the same kind of interference and facilitation as usual in the Stroop paradigm, but a smaller effect for the subliminal primes than for the consciously perceived primes (Figure 3.8 presents the latencies that Marcel observed). Marcel concluded that subliminal primes have an effect on latency, even though participants are not aware of their presence.

Four prime conditions were tested by Marcel (1983 Experiment 3): Color congruent, color incongruent, neutral, and no-word. In the congruent condition, the prime was the name of the target color, whereas in the incongruent condition the prime was the name of another color. In the neutral condition, the prime was a non-color word that was also unrelated to colors. The no-word condition presented the mask only. Thus, no prime was presented. The condition in which the prime was subliminal was called the unaware condition. In the aware condition, by contrast, the presentation duration was 400ms. Both prime and cue were presented at the same time.[9]

From a symbolic perspective, stimuli have to be considered as symbols in order to engage in cognitive processing. In ACT-R, this means that a stimulus has to be present in a buffer. However, stimuli that are presented for such short durations as common in subliminal priming paradigms do not reach the visual buffer. ACT-R assumes an attention shift to the stimulus before an object can be encoded as a symbolic chunk, which takes a certain amount of time, estimated at 185ms (Anderson, Matessa, & Lebiere, 1998). This exceeds the presentation duration of the prime in the unaware conditions (which is 80ms at maximum, Marcel, 1983). In ACT-R models, stimuli that are presented for less than the time it takes to shift attention can therefore not influence central cognition. The way ACT-R deals with stimulus durations is all or none. Either the stimulus has been presented not long enough, and the stimulus is not perceived at all, or it is fully is perceived. Consequently, symbolic theories of cognition cannot account for subliminal priming data. The next section will show how RACE deals with the short presentation durations typical in subliminal priming tasks.

9. In the original experiment, Marcel included also another condition with a Stimulus Onset Asynchrony between prime and cue. This condition is similar to the picture-word interference study by Glaser and Düngelhoff (1984), which has previously been discussed by Van Maanen and Van Rijn (2006; 2007b). Therefore, it is not included here.

SUBLIMINAL PRIMING MODEL

The subliminal priming model comprises three chunk types, as outlined in Figure 3.6: Lemmas, concepts, and motor mappings. The concept chunks can be regarded as representations of semantic properties. Chunks of the lemma type can be regarded as sets of orthographic and syntactic properties of a word. The motor mapping chunks represent the information which button to press for which color.

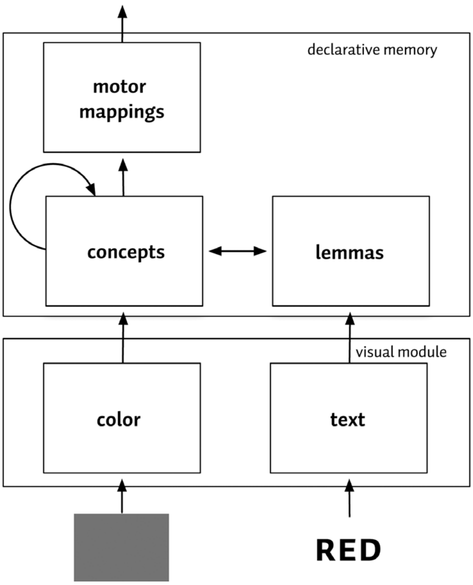


Figure 3.6. The flow of activation in the congruent condition of the subliminal priming model.

Now, for example in a no-word condition, the cue (being a color patch) spreads activation to its associated concept, which spreads activation to the associated motor mapping resulting in a button press. A similar *flow of activation* will occur in the other conditions, albeit that because of the presentation of a prime word, lemma chunks will also be activated. The activation of multiple motor mapping chunks causes competition in RACE, because the retrieval ratio is harder to reach with multiple accumulating chunks.

Before the experiment, Marcel determined for each participant the critical presentation duration for which participants could not discriminate between presence and absence of a prime (see Marcel, 1983 for details of the procedure). The presentation durations he found ranged from 30 to 80 ms. We used the presentation duration as an extra parameter in fitting the model to the data, with the constraints that its value should be in the range that Marcel found and that the activation of the prime chunk would not exceed the retrieval ratio (Inequality 3.9). Because the primes in the original experiment were visually masked, we assume that the presentation duration is equal to the time that the prime is available to the visual system.

Parameter	Value
A_{color}	1.8
A_{text}	1.5
β	.255
d^{acc}	.72
Δt	5 ms
θ	.81
aware presentation duration	400 ms
unaware presentation duration	70 ms

Table 1: Estimated parameter values for the subliminal priming model.

Table 3.1 presents all relevant parameters for the subliminal priming model. The presentation duration of primes in the aware condition is 400ms, as in the original experiment. The unaware presentation duration was estimated at 70ms, serving as the model's critical presentation duration. This duration depends on the RACE parameters presented in bold-face in Table 3.1. These parameters were not estimated for this experiment, but rather copied from a RACE model of picture-word interference (an updated version of Van Maanen & Van Rijn, 2006; Van Maanen & Van Rijn, 2007b). Hence, the only parameters presented here that were estimated for this model were the activation of the words (A_{text}) and of the color (A_{color}). The association values (S_{ji}) between chunks are presented in Figure 3.7.

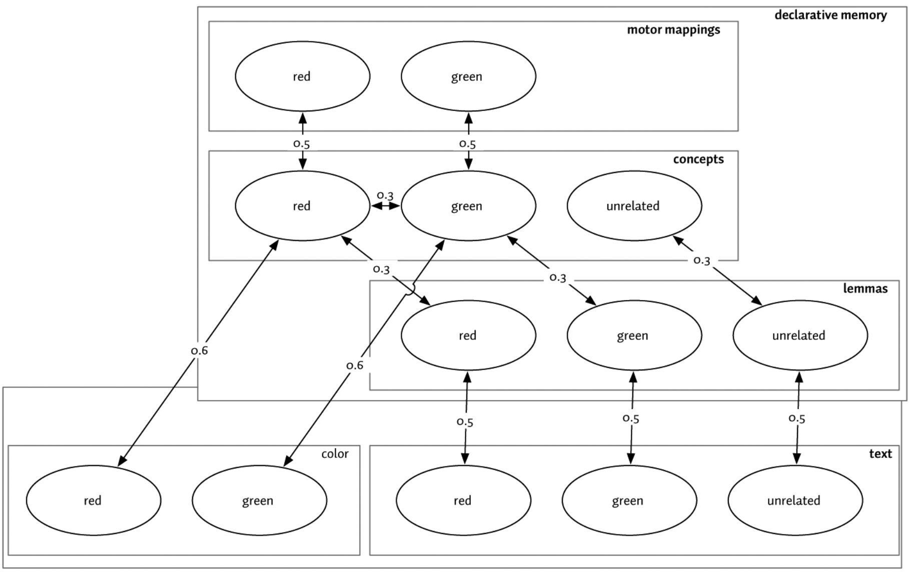


Figure 3.7. Associative values between different chunks in the subliminal priming model.

RESULTS

The results of the subliminal priming model are presented in Figure 3.8. We present here differences in latency relative to the no-word condition as this model only captures the memory retrieval process, which comprises the time course from the start of retrieval of a chunk until the retrieval of the motor mapping chunk. The model captures quite nicely the effects observed in the data[10] by Marcel (1983) ($R^2 = 0.987$).

10. As the variance in the original data cannot be deduced from the published results, a sensible formal comparison is not possible.

In the unaware conditions, only the target chunks reached the retrieval ratio, and no other chunks. Therefore, these chunks are the only ones that are consciously perceived by the model. The model thus remained unaware of all other chunks, as is required for these conditions.

An explication of how the activation flows through the model will be insightful. We split this up in four sections, each describing one condition.

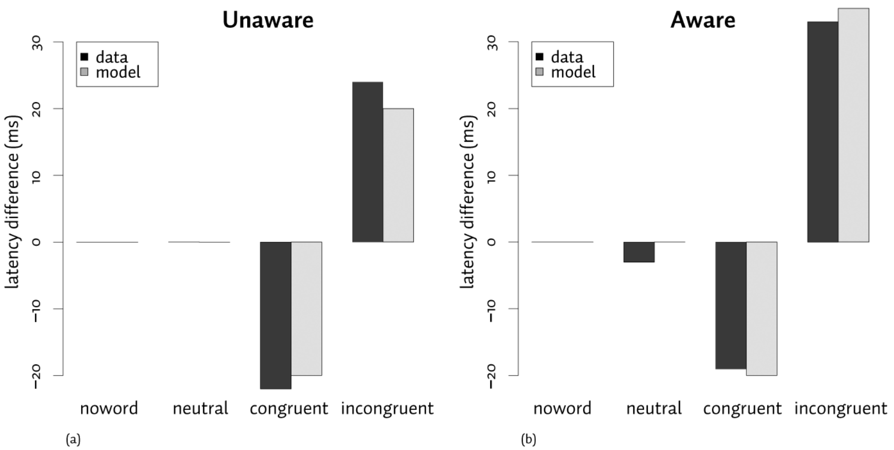


Figure 3.8. Comparison of the latencies found by Marcel (1983) (a) and the latencies predicted by the subliminal priming model (b). Shown here are the latency differences relative to the no-word condition.

Neutral

In the neutral condition, there is no competition between motor mappings, because there is no button associated with the neutral word. The activation cascades through the network similarly to the no-word condition, because no association exists between the neutral word and the target color, both at the lemma level and at the concept level. Therefore, the activation of the motor mapping associated with the target color increases similarly to the no-word condition because the activation of *all* the motor mapping chunks increases as in the no-word condition.

As an example, Figure 3.9 gives the activation accumulation in the neutral unaware condition. The activation of the neutral word lemma increases, but, due to the short presentation duration, it does not reach the retrieval ratio. This indicates that the neutral word does not reach awareness.

NO-WORD

The no-word condition is similar to the Neutral condition, because no distractor stimulus is present, resulting in the same behavior of the model as in the Neutral condition. Because in the no-word condition, there is no distractor, there is no difference between aware and unaware.

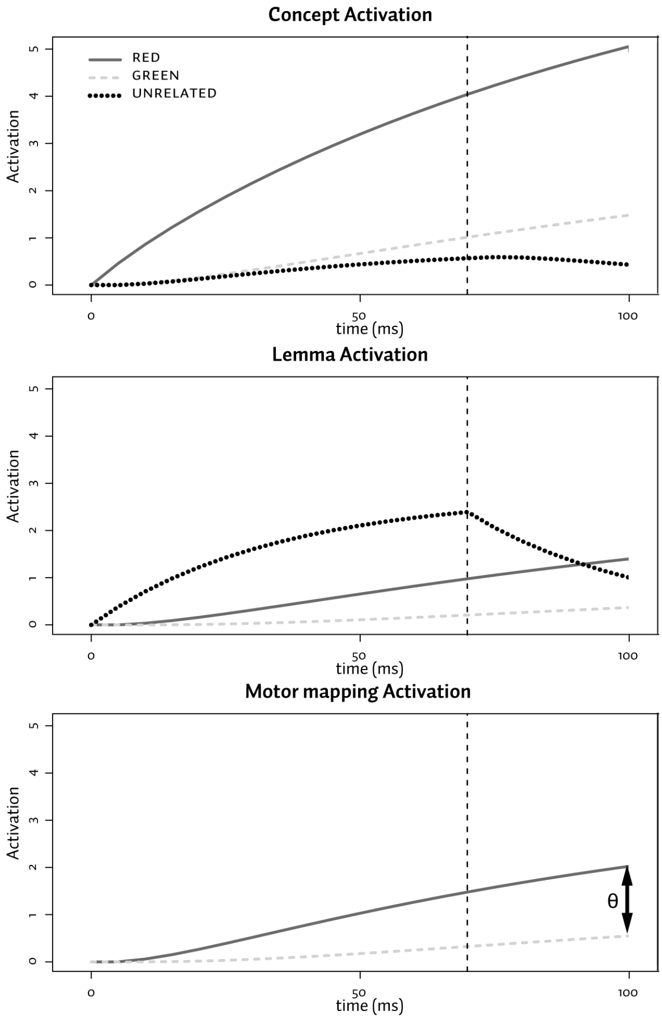


Figure 3.9. Activation accumulation in the Neutral Unaware condition. The stimuli are a red color patch and an unrelated prime word. The vertical dotted line indicates when the presentation of the prime ends. At this point, the activation of the related lemma has not crossed the retrieval ratio. After 100ms the red motor mapping chunk does cross the retrieval ratio.

Congruent

Both target and distractor stimuli activate the same concept: the color chunk directly, the text chunk mediated via the lemma chunk. Spreading activation towards the associated motor mapping chunk is therefore higher than in the Neutral and No-word conditions, resulting in faster retrieval.

Incongruent

Because both target stimulus and distractor activate a motor mapping, competition for retrieval takes place at the motor-mapping level. Higher activation for competing chunks means that it is more difficult to cross the retrieval ratio, leading to longer retrieval latencies. The effect is strongest in the aware condition, representing the longer presentation duration of the prime, and thus the longer accumulation of activation of prime-related chunks.

DISCUSSION

A difficult question when modeling cognitive tasks that deal with awareness is how awareness is defined within the model. We chose to set a strict boundary for awareness, the retrieval ratio. When a chunk reaches the retrieval ratio, it becomes available inside the buffers. We assume that people are aware of chunks that are currently in the buffers (Taatgen, in press), and not aware of chunks that have not yet reached the activation needed to enter the buffers.

Race involves a direct connection between information in the external world (that is, the visual module) and the activation values of declarative chunks in declarative memory. In this respect, race deviates from act-r, in which all visual information must be mediated by the visual buffer. However, since the visual buffer is associated with awareness as chunks appearing in the visual buffer enter the declarative system, another pathway must be present to account for the subliminal priming data modeled in this paper. We hypothesize that the connections in race from the visual module to the declarative memory module may represent part of the ventral visual pathway, that is known to involve connections from striate cortex (associated with act-r's visual module) to temporal brain regions (associated with the declarative memory module Anderson et al., 2004).

The model of subliminal priming discussed in this paper demonstrates that race can account for the retrieval latencies observed by Marcel (1983, Experiment 3). By using standard race parameter values, the fit of our model to the data set of Marcel was quite good. In combination with previous models of declarative memory retrieval that use race (Borst & Van Rijn, 2006; Van Maanen & Van Rijn, 2006; Van Maanen & Van Rijn, 2007b), this suggests that race might be regarded as a general model of declarative memory retrieval. The added value of the race model is that it gives a rational account of how the process of declarative memory retrieval develops. Even the effects on declarative memory retrievals of changes in the world that last only milliseconds can now be taken into account.

Stroop and picture-word interference are two sides of the same coin

This chapter is submitted for publication as Van Maanen, L., Van Rijn, H., & Borst, J. P. (submitted). Stroop and picture-word interference are two sides of the same coin.

INTRODUCTION

Over the years, the idea that the picture-word interference (PWI) effect and the Stroop effect are two manifestations of the same process has gained wide support. For instance, MacLeod's influential review on the Stroop effect (MacLeod, 1991) also discusses picture-word interference, and even lists the picture-word task in his list of "eighteen major empirical results that must be explained by any successful account of the Stroop effect" (MacLeod, 1991, Appendix B, p. 203). Recently, Dell'Acqua, Job, Peressotti, and Pascali (2007) have argued that the underlying processes of both effects are different, challenging the assumption of a single underlying process. In this paper, we will present a computational model that, on the basis of a single process, can explain both the traditional phenomena discussed by MacLeod (1991) and the data of Dell'Acqua et al. (2007).

In the Stroop paradigm, participants are presented with a word written in a specific color, and are instructed to either read the word, or name the color the word is printed in (MacLeod, 1991; Stroop, 1935). A typical Stroop experiment usually consist of three conditions: a congruent condition in which the word and the color refer to the same color concept (e.g., the word "red" written in red ink), an incongruent condition in which the word and the color refer to different color concepts (e.g., the word "red" written in green ink), and a neutral condition in which only the text or the color is displayed. This is often operationalized by presenting a set of X's in red ink (for color naming) or the word "red" printed in black ink (for word reading).

Three experimental findings in a Stroop study are extremely robust. First, color naming in the incongruent condition is slower than color naming in the neutral condition. That is, participants who are required to name the color of the ink are slower if the word describes a *different* color than if they have to name a color of a row of X's. This is often referred to as semantic interference, since many accounts of this effect assume that the relation in meaning between the color of the ink and the word itself causes the interference that results in reaction time differentiation (e.g., Cohen, Dunbar, & McClelland, 1990; W. R. Glaser & Glaser, 1989; Klein, 1964). A second finding in the Stroop literature is that naming the color of the ink of a word that describes the *same* color (congruent condition) is faster than the neutral condition. This will be referred to as semantic facilitation (but see MacLeod & MacDonald, 2000, for an alternative explanation of semantic facilitation that is based on accidental reading of the word). A third important observation is that both semantic interference and semantic facilitation disappear when the participants are not instructed to name the color of the stimulus, but instead are asked to pronounce the word, irrespective of the color the word is printed in. This so-called Stroop asynchrony is often explained by the difference in processing speed between colors and words (e.g., Cohen, Dunbar, & McClelland, 1990; Roelofs, 1992), or the difference in automaticity between color-naming and word-reading (e.g., Lovett, 2005).

The picture-word interference task (PWI) typically has a similar set-up to the Stroop task. Participants are presented with a picture on which a word is super-imposed, and are instructed to either name the picture, or read the word. As in the Stroop task, semantic interference, semantic facilitation and Stroop asynchrony are usually observed in the PWI task (W. R. Glaser & Düngelhoff, 1984). Semantic interference is observed if participants are required to name a picture on which a category-member word of that picture is super-imposed (e.g., a picture of a dog with the word "cat" inscribed in it) as their responses are slower compared to the condition in which a picture without a superimposed word has to be named or to the condition in which the superimposed word is completely unrelated (e.g., a picture of a dog with the word "desk" superimposed). Semantic facilitation is observed if, by contrast, the picture and the word refer to the same concept (e.g., a picture of a cat with the word "cat" superimposed) as the participants are faster. The Stroop asynchrony can be observed in the disappearance of these effects if the task is to read the word instead of naming the picture.

At a more general level, these effects are often aggregated in terms of a single Stroop or PWI-effect. This more general effect is the difference in performance for trials in which there is an interfering stimulus present versus trials in which there is no interfering stimulus. For the Stroop task, this entails the difference between trials in which a color word is written in different colored ink and trials in which a color word is written in similar colored ink. For the PWI task, the effect is the difference between trials in which the picture is presented with a semantically related, but different word superimposed versus trials in which picture and word are completely unrelated.

Given the similarity of both tasks, it is not surprising that both tasks have often been explained in similar terms. In fact, in our previous work in which we presented a computational model of PWI, we took for granted that picture-word interference is just an instance of the Stroop effect. (Van Maanen & Van Rijn, 2007b). Other computational models also present a single mechanism that underlies both PWI and Stroop performance (e.g., Cohen, Dunbar, & McClelland, 1990; Lovett, 2005; Roelofs, 1992, 2003).

Most of the theoretical accounts of Stroop-like effects have focussed on finding the locus of the interference effect in the mental processing stream. For instance, many theorists assume that interference is caused by a competition between different response options (Cohen, Dunbar, & McClelland, 1990; e.g. Dyer, 1973; and W. R. Glaser & Glaser, 1989; Kuipers, La Heij, & Costa, 2006; Lovett, 2005; Roelofs, 1992, 2003). In this view, interference is caused by an increased difficulty of selecting the appropriate response in the incongruent condition. Often, this has been attributed to either a difference in the speed of processing between the stimulus dimensions or to a difference in automaticity between the stimulus dimensions (MacLeod, 1991).

Some theoretical accounts assume an interference effect during an early stage. In particular, Dell'Acqua et al. (2007) argue that the picture-word interference effect is caused by a competition that occurs *before* an appropriate response is selected. With respect to the Stroop effect, many studies now suggest that the Stroop effect may be manifested at multiple stages (e.g., De Houwer, 2003; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005; Van Veen & Carter, 2005).

Recently it has been suggested that the Stroop effect and picture-word interference are *not* caused by the same process (Dell'Acqua et al., 2007). In particular, it has been argued that the loci of the two interference effects in the mental processing stream differ. Many studies suggest that the locus of the Stroop effect is on the level of response selection (e.g., Fagot & Pashler, 1992; Kuipers, La Heij, & Costa, 2006; MacLeod, 1991; Roelofs, 2003). That is, Stroop interference occurs because an incorrect response possibility that is triggered by the distracting feature of the stimulus (the word), interferes with the correct response that is triggered by the target feature (the color). Dell'Acqua et al. argue that this is not the case for PWI. Their result suggests that picture-word interference is on the level of the perceptual encoding. Based on these results, Dell'Acqua et al. concluded that although the effects in both tasks seem similar, they are in fact caused by different mechanisms.

This paper presents a computational cognitive model of both the Stroop and the PWI effect, reconciling the recently observed differences between Stroop and PWI with the general view that PWI is an instance of the Stroop effect. In this model, both Stroop and PWI effects are accounted for by the same underlying mechanism. In other words, the model provides additional evidence for the view that the underlying mechanism that causes interference in both tasks is the same (cf. Cohen, Dunbar, & McClelland, 1990; Lovett, 2005; Roelofs, 1992, 2003), whilst still being able to explain the observation that participants respond differently to Stroop and PWI stimuli under certain circumstances (Dell'Acqua et al., 2007).

ANALYZING THE LOCUS OF INTERFERENCE

To analyze the time course of the Stroop effect, Fagot and Pashler (1992, Experiment 7) studied whether the Stroop effect would persist in a psychological refractory period (PRP) design. In a PRP design, participants are required to perform two tasks concurrently. The main manipulation in PRP designs is an asynchrony in onset between both tasks. Usually, the instruction is to first give the response associated with the stimulus that was displayed first. If the stimulus onset asynchrony (SOA) between tasks is relatively long, the processing of the first task is finished before the stimulus of the second task is presented. On the other hand, if the SOA is short, processing associated with the first stimulus might not be finished yet when the second stimulus appears. The typical observation with short SOAs is that the response to the second task is delayed (Telford, 1931). This delay is often interpreted as evidence for a bottleneck in processing. The central bottleneck theory assumes that there exists a processing stage during which only a single process can proceed concurrently (e.g., Pashler, 1994). Therefore, the secondly presented task will be delayed by processing of the first task. By testing the effect of different SOAs the time course of processing can be unraveled.

Fagot and Pashler (1992) used the PRP paradigm in the context of the Stroop task. They presented participants with a simple tone classification task as primary task, and a Stroop task as secondary task. Fagot and Pashler hypothesized that if the Stroop effect – operationalized as the latency difference between an incongruent Stroop stimulus and a congruent Stroop stimulus – would be caused by the perceptual encoding of the stimulus, the Stroop effect would disappear at short SOAs. This would be the result of the delay in execution of the second task. The delay causes a gap in processing of the second task in which the interference could be resolved. Fagot and Pashler found no effect of SOA on the magnitude of the Stroop effect, which they interpreted as evidence that the locus of the Stroop effect is located relatively late in the stream of mental processing. This finding, they argued, is in line with a response selection account for the locus of the Stroop effect, because response selection is also late in the mental processing stream.

In a similar experiment, Dell'Acqua et al. (2007) answered the same question for the PWI effect. Dell'Acqua et al. conducted an experiment very similar to the experiment ran by Fagot and Pashler (1992), but instead of a Stroop task, participants were presented with a PWI task. In this experiment, the interference was operationalized as the latency difference between a semantically related word-picture pair and an unrelated word-picture pair. Interestingly and surprisingly, SOA mediated the PWI effect: shorter SOAs were associated with smaller PWI effects. This indicates that the inference originates from the initial stages of processing, as with short SOAs most interference is captured in the PRP-induced delay, but with long SOAs the delay does not account for all interference. This finding is in line with the view that the locus of PWI is early, possibly during perceptual encoding of the stimulus.

Based on this result, Dell'Acqua et al. (2007) argue that the difference between their findings and Fagot and Pashler's (1992) are "incompatible with the often reiterated principle that the PWI effect comes about for limitations of the cognitive system that are analogous to those causing the Stroop effect" (Dell'Acqua et al., 2007, p. 720). They conclude that their analysis "favor[s] an interpretation of the present findings that points to the functional dissociation of the sources of Stroop and PWI effects" (Dell'Acqua et al., 2007, p. 722).

Based on the PRP studies by Dell'Acqua et al. (2007) and Fagot and Pashler (1992), we may arrive at two possible explanations. The first possibility is the one advocated by Dell'Acqua et al., that Stroop and PWI are two different effects. The other possibility is that the interference in Stroop and PWI is distributed over multiple stages. However, the *amount* of interference per stage may differ between Stroop and PWI. This explanation is in line with findings from the Stroop literature that besides response competition, Stroop may be caused by stimulus-related competition as well (e.g., De Houwer, 2003; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005).

We will support this second possible explanation by presenting a cognitive model that can account for reaction time data of both PWI and Stroop experiments. The cognitive model utilizes a single mechanism to account for both data sets; the sole difference is a different speed of processing for the perceptual input (that is, pictures vs. colors). Next, we will demonstrate that the model accounts for the response time patterns of Stroop and PWI tasks under PRP conditions. As these fits are obtained with a single-mechanism model, this argues against the claim that the differential findings for Stroop and PWI favor a dissociation of the sources of Stroop and PWI effects.

The cognitive model that we will describe here is an integrated cognitive model (Gray, 2007b) of the task, implemented in a previously validated cognitive architecture (Taatgen & Anderson, 2008). We will simulate the complete process that is involved in the task, from the presentation of the stimulus up to the participant's response, resulting in quantitative predictions of reaction time data.

The model is implemented in the cognitive architecture ACT-R (Anderson, 2007a). ACT-R assumes that specialized modules process different kinds of information. For instance, a visual module handles visual perception, and a motor module executes motor commands. Other modules that will play a role in the model described below are the declarative module, used for storing and accessing information in declarative memory, the speech module for speech output, the aural module for auditory perception, and a goal module for keeping track of goals and intentions (see Anderson, 2007a; Taatgen, Van Rijn, & Anderson, 2007 for extensive descriptions of the identified modules). A central production rule system integrates the information that is made available by the different modules, and issues new instructions to those modules. The production rule system communicates with the different modules through a set of interfaces called buffers. Behavior in ACT-R emerges from the selection and subsequent execution of production rules that consist of simple conditions-actions pairs. If the information that is present in the buffers matches the conditions of a production rule, that rule may be selected to execute its actions. Production rule actions consist of operations on the buffer contents, such as a request for new information from declarative memory, or a request for pressing a button on a keyboard.

Declarative information in ACT-R is represented in symbolic entities called chunks. Chunks represent simple facts, such as *The capital of The Netherlands is Amsterdam*, or *The object I am attending is green*. Note that where the first example is a typical semantic memory fact, the second example represents a visually observable feature of the world as might be present in the visual buffer. All chunks in an ACT-R model have an activation level that reflects the likelihood that they will be needed in the near future. The activation of a chunk depends on two components: the chunk's history of usage (Anderson & Schooler, 1991) and the current context (Anderson & Milson, 1989). The activation of a chunk is the main determiner of the time it takes to retrieve that chunk from memory. All other things being equal, the higher the activation of a chunk, the faster it will be retrieved.

Because Stroop tasks and PWI tasks usually involve well-known colors and words, and pictures of well-known objects, we will assume that the history component of the activation is approximately equal for all chunks as well as stable over the time span of a single experiment. For the context component of the activation we will adopt a more fine-grained model that will be discussed later. This context component will be the main determiner for the reaction time differences that our models will display.

The model's behavior is determined by the interaction between specified production rules, chunk retrievals and the task setup. In terms of response latency (the usual dependent measure in Stroop and PWI tasks), the model's behavior is the result of an aggregation of the timing of the sub processes, such as the execution time of the production rules and the time associated with module-specific operations, such as declarative retrievals, or button presses (cf., Donders, 1868/1969; Sternberg, 1969). However, although internally most module actions are executed sequentially, the modules themselves operate in parallel. For example, while the visual system is busy with the perceptual processes involved in perceiving a new stimulus, production rules might initiate a request for the retrieval of a fact from declarative memory without disturbing the perceptual process. Critically however, the modules cannot execute multiple operations in parallel. Thus, if two tasks require retrieval from declarative memory, one of the tasks has to wait until the declarative module is finished with the request of the other task. This seriality will be critical in the explanation of PWI and Stroop performance. Although this seriality has not been specifically designed to account for performance in tasks in which scarce resources determine behavior, it has been successfully applied in many different experimental domains (Language Development, Hendriks, Van Rijn, & Valkenier, 2007; e.g., Attentional Blink, Taatgen, Juvina, Schipper, Borst, & Martens, in press; Van Rij, Van Rijn, & Hendriks, submitted; Temporal Cognition, Van Rijn & Taatgen, 2008).

MODEL OVERVIEW

The model presented here, in line with the proposal of Dell'Acqua et al. (2007), assumes that both the Stroop task and PWI consist of three main stages, the perceptual encoding stage, the response selection stage, and the response execution stage. During the perceptual encoding stage, the stimulus features are transferred to the visual buffer. During the response selection stage, a chunk that reflects the syntactic properties of the response (that is, a lemma, Levelt, Roelofs, & Meyer, 1999) is retrieved from declarative memory. In the response execution stage, the model retrieves a motor program associated with the retrieved response, and this motor program is executed.

During each stage, one or more memory retrievals take place. The duration of these retrievals are the main determiners of the response latency. Given that we assumed that all chunks are equally active, all differences in activation are driven by the context activation. If the current state of the system is favorable for a chunk that is requested from memory, the chunk will be retrieved faster than in a situation when the context is less favorable. This context phenomenon determines whether semantic interference or semantic facilitation is observed.

A MORE FINE-GRAINED ACCOUNT OF CONTEXT EFFECTS

The current declarative retrieval module in ACT-R can account for many memory-related phenomena (e.g., Anderson et al., 1998; Taatgen & Anderson, 2002; Van Rijn & Anderson, 2003). While the module accurately predicts the duration of memory retrievals, ACT-R does not provide an account for what happens *during* memory retrievals. This is especially problematic in tasks with multiple stimuli presented at short SOAs, since the first, not-yet-completed retrieval is influenced by a second process. This dependency cannot be explained by default ACT-R (Van Maanen & Van Rijn, 2007b). Another issue is that context effects on retrieval latencies are solely driven by chunks that are available in the buffer of one of the modules. This results in a threshold function: no context effects can be observed of retrieval or perception processes until a chunk is placed in a buffer, at which point it starts spreading activation to contextually related chunks.

To overcome these issues, we proposed an adaptation of the declarative retrieval mechanism in the cognitive architecture ACT-R to account for the time-course of memory retrieval on short time scales (Van Maanen & Van Rijn, 2007b). This retrieval account (RACE/A, for Retrieval by ACcumulating Evidence in an Architecture) predicts what happens during the actual retrieval process. RACE/A is driven by two key assumptions: (1) The activation of one chunk is determined (in part) by the activation of other chunks. (2) The activation of one chunk relative to the activation of other chunks determines the likelihood that it will be retrieved.

The first assumption represents the notion that the relevance of information is context-dependent, even when this context is not yet available at a symbolic level (that is, accessible to the production system). This is for instance reflected in subliminal priming studies in which a related prime decreases the response latency on a target stimulus, even when participants were not consciously aware of the prime (Marcel, 1983). We operationalized this by adopting a spreading activation strategy (Collins & Loftus, 1975) in which increased activation of one chunk increases the activation of related chunks.

$$C_i(t) = \alpha C_i(t-1) + \beta \sum_{j \in k} C_j(t-1) S_{ji}$$

(equation 4.1)

Equation 4.1 implements this assumption. The equation reflects how the activation of a chunk ($C_i(t)$) accumulates during retrieval. The activation at time t depends on the previous activation of that chunk ($C_i(t-1)$), as well as additional spreading activation ($C_j(t-1)S_{ji}$) from other chunks (k). This includes the chunks that are available through perceptual processing. The spreading activation is mediated by the associative strength between two chunks (S_{ji}), such that chunks that are strongly associated exchange more activation than chunks that are loosely associated. α and β are scaling parameters that determine the relative contributions of both components. Because α is set to a value in the range (0,1), it can be interpreted as temporal decay of activation. The accumulated activation thus decays after a retrieval has been attempted.

The second assumption states that the relative activation of a chunk determines the likelihood that that chunk will be retrieved. This assumption reflects the insight that if multiple memory representations are relevant, responding becomes more difficult (Luce, 1986). Following ACT-R, the activation of a chunk determines the likelihood that it will be needed in the near future. However, RACE/A extends the default ACT-R equations to take the activation of competing chunks into account. The activation of competing chunks are accounted for by taking the ratio of activation of the to be retrieved chunk (chunk *i* in Equation 4.2) to the sum of activations of other relevant chunks (chunks j in Equation 2, cf., Luce, 1986; Roelofs, 1992). In the current model, the relevant chunks are all other chunks that match the criteria specified by a retrieval request.

If the ratio specified in Equation 4.2 crosses a threshold (θ, the retrieval ratio), the relative activation of the chunk in the denominator (chunk *i*) warrants the retrieval of that chunk. As soon as a chunk passes this threshold, RACE/A returns that chunk as the result of the retrieval process.

$$\frac{e^{A_i}}{\sum_j e^{A_j}} \geq \theta$$

(equation 4.2)

Based on Equations 1 and 2, we have provided quantitative predictions for variants of both picture-word interference (Van Maanen & Van Rijn, 2007b) and the Stroop task (Van Maanen & Van Rijn, 2007a). The PWI model focused on the effect of SOA differences; the Stroop model fitted a data set in which the distractors were presented subliminally.

Besides a theory of memory retrieval (RACE/A), which is the core of our modeling efforts in this paper, also a theory of perceptual encoding is needed to study the Stroop and PWI effects. Similar to RACE/A, we deviate here from the theory currently implemented in ACT-R, because it is to crude for our purposes here.

PERCEPTUAL ENCODING

Before the processing of task-relevant information can commence, visual or auditory information has to be made available to central cognition. In both PWI and Stroop tasks, all information is presented visually. In ACT-R, the perceptual encoding process results in a chunk entering the visual buffer, and thus becoming available for further processing. By default, it is assumed that this process takes a fixed 85 ms (Anderson, 2007a). Although this is a sensible number when the details of the perceptual processing are less relevant for the task under study, the emphasis on perceptual encoding in Stroop and PWI tasks requires a more detailed account. However, it should be noted that our implementation of a perceptual encoding process should not be considered a complete theoretical account of perceptual processes. Rather, our implementation should be considered a functional description, aimed at differentiating between the effects on the encoding time caused by different stimulus features (cf., Gray, 2007a).

We assume that features of the stimulus become available during the first stages of perceptual encoding (cf., feature integration theory, Treisman & Gelade, 1980). Based on these features, chunks in declarative memory compete for retrieval. This process is implemented using the RACE/A mechanism: The visual features spread activation to chunks that represents concepts that are likely to be needed in the context of those features. Thus, concepts that relate to the visual features of the current stimulus receive activation, while concepts that do not relate do not receive activation. When a concept's activation results in a retrieval ratio larger than the threshold (see Equation 4.2), that concept is made available for non-visual cognition as the result of visual processing. So, instead of a fixed duration for perceptual encoding, we have implemented perceptual encoding as a combination of a feature integration and selection process with a variable duration that depends on the characteristics of the input.

The variation in encoding time between different stimuli originates from two sources. First, as the relative activation of a chunk determines whether it will be encoded, the activation of other chunks in declarative memory partly determines the encoding time. Second, the speed by which activation spreads from features to chunks will be different for different types of stimuli (e.g., a color patch is easier to recognize than a complex line drawing). Therefore, a parameter is introduced that reflects the speed and strength of spreading activation from features to chunks, to account for the different encoding times associated with different types of stimuli (Dell'Acqua, Lotto, & Job, 2000; Rossion & Pourtois, 2004). Although our encoding account is much simpler than existing models of perceptual encoding (e.g., Itti & Koch, 2001; Treisman & Gelade, 1980; Wolfe, 1994) it provides the necessary detail to account for the perceptual processing in the PWI and Stroop tasks.

In the Stroop task, the color feature of the stimulus spreads activation to a concept representing that color, and the text feature of the stimulus spreads activation to the lemma associated with that word. For the PWI task, the features representing the line drawing spread activation to the concept chunk representing the content of the picture, while the word spreads activation to a lemma. As word reading is faster than color naming (M. O. Glaser & Glaser, 1982), more activation spreads from the text feature to the lemma representing the associated word, than from the color feature to the associated color lemma.

In addition, we assume that line drawings are encoded slower than color patches. This assumption is supported by studies that show faster naming for sequences of color patches as opposed to sequences of images (e.g., Denckla & Rudel, 1976; Vukovic, Wilson, & Nash, 2004).

In the model these differences are operationalized by different settings for the parameter that reflects the speed of perceptual processing. In the remainder of this paper, we will demonstrate that the interference dynamics observed in Stroop and PWI PRP tasks might be caused by differences in this perceptual processing speed.

SIMULATION 1: A SINGLE MODEL FOR PWI AND STROOP

Before discussing our model of the Stroop and PWI effects under PRP conditions we will first discuss the model's performance in non-PRP Stroop or PWI trials. Note that although this section focuses on a Stroop trial, exchanging references to color dimensions with references to pictures results in a description of how the model would perform a PWI trial.

As soon as visual features become available to the model activation is spread to the chunks associated with that visual information. The color features spread activation to a conceptual chunk representing that color, whereas the textual features spread activation to a lemma.

When a chunk representing either the lemma or the concept crosses the retrieval ratio threshold, a second stage is initiated. As availability of a lemma is a prerequisite for starting the response execution stage, the model directly enters this stage when the lemma crosses the threshold. In this stage, the model retrieves a word-form and the response is uttered. If the color has to be named, the model waits until a concept is selected on the basis of the visual input, after which a retrieval is initiated to retrieve the lemma that is associated with the concept. However, because word reading is a highly trained and automated process, the word activates its associated lemma even if the model expects a color concept. Because of this activation, and the earlier discussed retrieval ratio, color naming will show interference from the presented word, whereas word reading will show less interference. Note that the model described so far is very similar to models of the Stroop and PWI tasks presented by Roelofs (1992, 2003). In some cases, the visual word representation might have spread so much activation, that the word-based lemma chunk exceeds the color-based lemma chunk in activation. In these situations, an incorrect lemma will be retrieved. To account for this, the model checks if the meaning of the lemma is consistent with the stimulus color (other models of the Stroop task that incorporate a similar strategy include Altmann & Davidson, 2001; Juvina & Taatgen, 2009; cf., Van Rijn & Anderson, 2004). If the correct lemma has been retrieved, the model continues with the response execution. If not, the model tries to retrieve another lemma. The response selection stage is the same for the word reading task: the model retrieves a word-form and utters a response.

SIMULATION RESULTS & DISCUSSION

Figure 4.1A presents the fit of the Stroop model for both the color naming task and the word reading task (RMSE=33ms, R^2=0.93) on the data from Glaser and Glaser (1982, Experiment 1 for SOA=0). The model captures both Stroop interference in the incongruent color naming condition, and facilitation in the congruent color naming condition. In addition, the Stroop asynchrony between color naming and word reading can be observed. Given our hypothesis that the Stroop effect and picture-word interference are manifestations of the same process, the challenge is to demonstrate that both effects can be fitted with the same model. Figure 4.1B presents the model fit on picture-word interference response time data (W. R. Glaser & Düngelhoff, 1984, Experiment 1 for SOA=0). Similar to the fit of the Stroop task, all the behavioral patterns (interference, facilitation, and asynchrony) are captured by the model (RMSE=60ms, R^2=0.85). The only difference between the two simulations underlying Figure 4.1 is a single parameter that controls the speed of processing of the stimulus that is adjusted to reflect the differences in stimuli between the two tasks: The processing speed was set lower for the picture-word stimulus than for the color-word stimulus, to reflect that visual processing of pictures is slower than visual processing of colors.

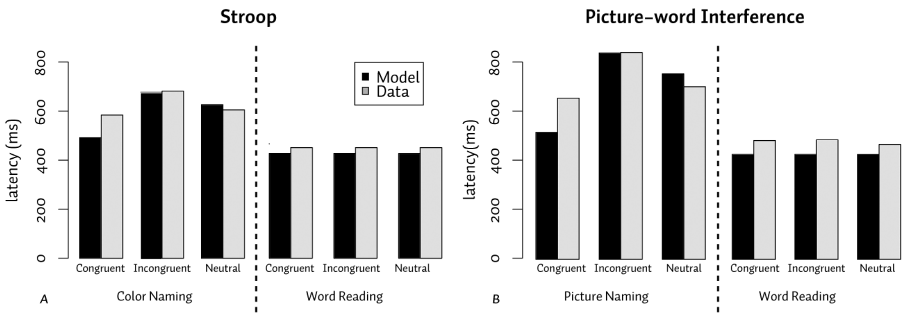


Figure 4.1. Data and Model for the Stroop task (A) and PWI (B). The Stroop data are from (M. O. Glaser & Glaser, 1982, Experiment 1 for SOA=0), the PWI data are from (W. R. Glaser & Düngelhoff, 1984, Experiment 1 for SOA=0).

Following our assumption that semantic interference is an effect of competition during the retrieval of declarative facts, we analyzed the amount of interference by comparing the duration of declarative memory retrievals between conditions. Thus, in our model, interference constitutes the difference in retrieval times. Figure 4.2 presents the difference in memory retrieval time between the model's incongruent and congruent conditions, for both the Stroop task and the PWI task. Following others (e.g., Fagot & Pashler, 1992; Ferreira & Pashler, 2002; Jolicoeur & Dell' Acqua, 1998; McCann & Johnston, 1992; Pashler, 1994; VanSelst & Jolicoeur, 1997) we assume that the bottleneck that Fagot and Pashler (1992) and Dell'Acqua et al. (2007) hypothesize is located after the perceptual encoding stage. Figure 4.2 shows the

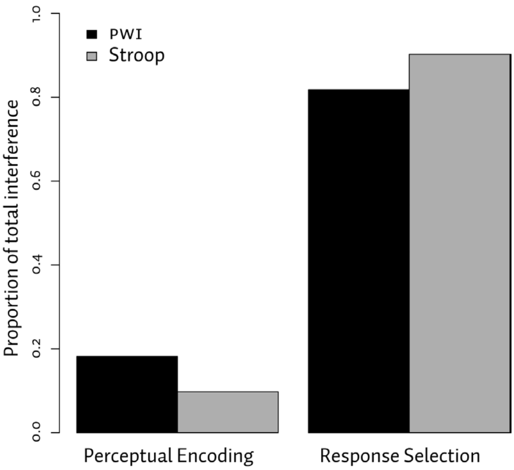


Figure 4.2. The proportion of interference before and after the hypothesized bottleneck. This is calculated as the proportion in each stage of the memory retrieval time for the incongruent trials minus the memory retrieval time for the control trials.

proportion of interference that is located before and after the hypothesized bottleneck. In the model of the Stroop task, 9% of the interference is located before the bottleneck, and 91% is located after the bottleneck. For PWI, these percentages are different. 18% of the retrieval-induced interference is located before the bottleneck, which is twice as much as the early interference in the Stroop task. This is in line with Dell'Acqua et al.'s conclusions, because more interference in the picture-word interference simulation is indeed located early in the processing stream.

Simulation 1 shows how a single cognitive model can account for PWI and the Stroop effect. Although this result has been obtained before (e.g., Cohen, Dunbar, & McClelland, 1990; Roelofs, 2003), the model presented here provides insights in the temporal distribution of the interference patterns. The model assumes that interference is a consequence of competition during memory retrieval, and shows that different processing speeds for different stimulus types (pictures vs. colors) mediates the competition in subsequent stages of the task.

The ability to model both a Stroop task and a PWI task with a single model, supports the view that both tasks are manifestations of the same interference process (e.g., MacLeod, 1991). By studying the memory retrieval times in different processing stages, we showed that different stimulus features may lead to a different temporal distribution of the interference patterns. This result leaves the puzzling observation by Dell'Acqua et al. (2007) that Stroop and PWI behave differently under PRP conditions. In the next section, our model will be extended to account for the PRP experiments by Fagot and Pashler (1992) and Dell'Acqua et al. (2007). We will demonstrate that a difference in stimulus features between colors and pictures may lead to different behavior in the Stroop and PWI tasks under PRP conditions.

SIMULATION 2: A COGNITIVE MODEL OF INTERFERENCE DURING PRP

The cognitive model presented above can account for prototypical Stroop and PWI data sets (M. O. Glaser & Glaser, 1982, Experiment 1 for SOA=0; W. R. Glaser & Düngelhoff, 1984, Experiment 1 for SOA=0). Moreover, analysis of the locus of interference in the model suggested that the magnitude of the interference effects differed between stages of the model, as well as between tasks. To demonstrate that the different loci of interference may appear as different response latencies under PRP conditions, we added PRP conditions to the model. With this extention, we can assess whether the results of Fagot and Pashler (1992) and Dell'Acqua et al. (2007) can indeed be explained with a single mechanism.

MODEL OF FAGOT AND PASHLER (1992, EXPERIMENT 7)

Following the experimental setup by Fagot and Pashler (1992), each trial started with a tone classification task. In this task, the participants were instructed to classify presented tones as either having a low or a high pitch by pressing one of two buttons. This additional task was added to the model using ACT-R's standard auditory perception module. As soon as the auditory system perceived a tone, a retrieval was initiated for a tone-to-button mapping. Finally, the model made a motor response to press the correct button. The model's processing of the Stroop stimulus only commenced *after* the response to the tone was selected, as was required according to the instructions provided to the participants. Note that this does not mean that the button press has to be finished, but that the information that has to be sent to the motor system has become available. As in the Fagot and Pashler experiment, the delays between tone onset and Stroop stimulus were -50ms, 50ms, 150ms, and 450 ms. All other aspects of the model were kept constant, apart from the estimation of a fixed intercept for voice-key responses. To account for differences in naming speed between this experiment and the previously modeled datasets, we estimated the voice-key response parameter at 167 ms (see also Meyer & Kieras, 1997a; Salvucci & Taatgen, 2008).

Simulation Results & Discussion

The model fits of the PRP Stroop experiment (Fagot & Pashler, 1992, Experiment 7) are presented in Figure 4.3. The model shows the same latency effects as Fagot and Pashler observed (RMSE=16ms, R^2=1.00). First, latency in the tone classification task is not affected by the interval between tone and the combined picture-word stimulus. In discussion of their paper, Fagot and Pashler attribute this to the instruction that was given to the participants to always respond to the tone first and to the Stroop stimulus second. We simulated this in the model by requiring a response to be selected for the tone before the model could commence with the response selection for the Stroop task.

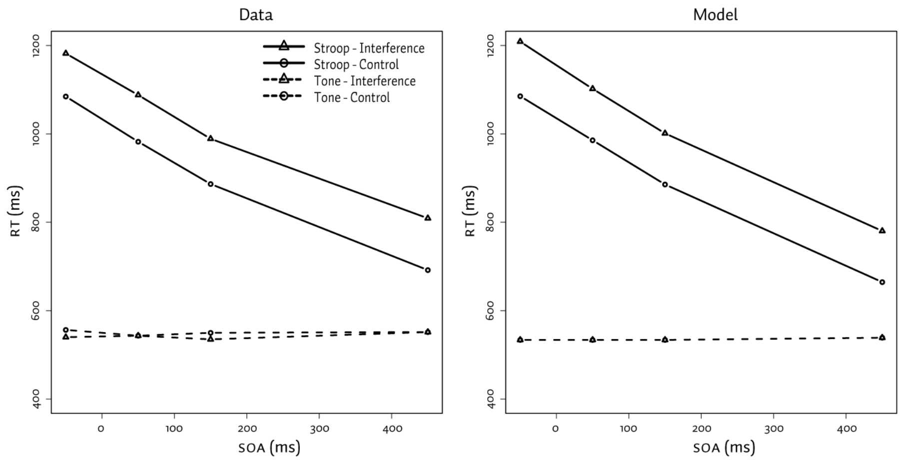


Figure 4.3. Model fit to the data of Fagot and Pashler (1992). Stroop denotes responses on the Stroop task, Tone denotes responses on the tone classification task.

Second, the model shows a decrease in reaction time as a function of an increasing SOA, similar to the typical PRP effect found in the data. The model achieves this because as the interval between the two tasks increases, the probability that both tasks need one of the ACT-R modules at the same time decreases. As only one task can use a module at a certain

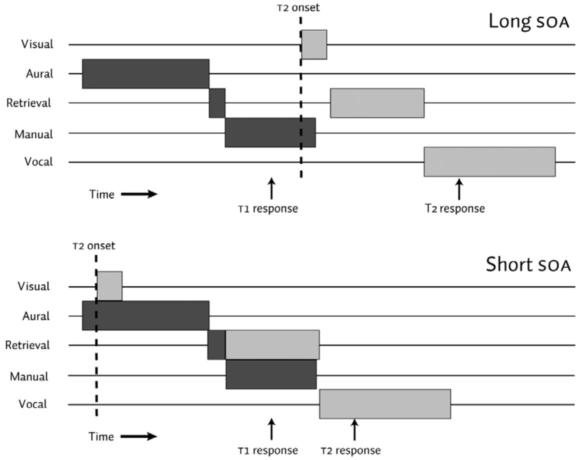


Figure 4.4. Gant diagrams of the model's processing of the tone classification task (dark grey) and the Stroop task (light grey). T2 onset indicates the onset of the Stroop stimulus. Each line denotes activity in a specialized module, indicated on the left. In the short SOA condition, retrieval of the appropriate response in the Stroop task has to wait until the retrieval module is available. The arrows indicate the responses for the two tasks, after which the output modules remain occupied with finalizing the motor action required for the response.

time, one of the tasks has to wait when two parallel requests are made to a module. Also, as the interval between the tasks increases, the delay due to the task instruction regarding the response order is less. That is, if the Stroop stimulus is presented later (that is, when the SOA is longer), the delay for the response is shorter. This is illustrated in Figure 4.4, which presents Gant diagrams of the module-specific processing for both tasks. The dark grey bars represent the tone classification task, and the light grey bars represent the Stroop task. In the upper panel of Figure 4.4, the SOA between the tasks is long, which is reflected by the late activity in the visual module of the model. In the lower panel, the SOA is short. Initially, the Stroop stimulus is perceived, as indicated by the activity in the visual module. However, the model withholds further Stroop processing until the response in the tone classification task has been initiated. Thus, parallel to activity in the manual module, representing the button press in response to the tone classification, the model initiates response selection in the Stroop task.

Finally, the model shows no effect of SOA on the size of the Stroop effect, which has been interpreted as evidence for a late locus of the Stroop effect (Dell'Acqua et al., 2007; Fagot & Pashler, 1992). This is because the interference in the Stroop task is mainly located in the response selection stage (see Figure 4.2), which does not overlap with the tone classification task.

Model of Dell'Acqua et al. (2007)

The model of the Dell'Acqua et al. experiment is a direct combination of the PWI model described under Simulation 1 and the PRP aspects described above. To match Dell'Acqua et al.'s experiment, we changed the SOAs to 100ms, 350ms, and 1000 ms and we extended the tone classification task to three options to mirror the exact experimental setup of Dell'Acqua et al. The single free parameter, the voice-key intercept, was estimated at 212ms.

Simulation Results & Discussion

The model again found no effect of SOA on tone classification latency and the PRP effect (manifested as decreased response times on the PWI stimulus as SOA increases). In addition, this time the models also shows an increase in interference with increased SOA (Figure 4.5, RMSE=61ms, R^2=1.00), which is in line with the result from Dell'Acqua et al. (2007). Because the interference is primarily expressed during the perceptual encoding stage, postponement of the response on the PWI stimulus (which happens for short SOAs) results in a decrease (or disappearance) of the interference effect. This is depicted in the Gant diagram in Figure 4.6. The upper panel shows the presentation of a neutral word-picture pair, resulting in fast

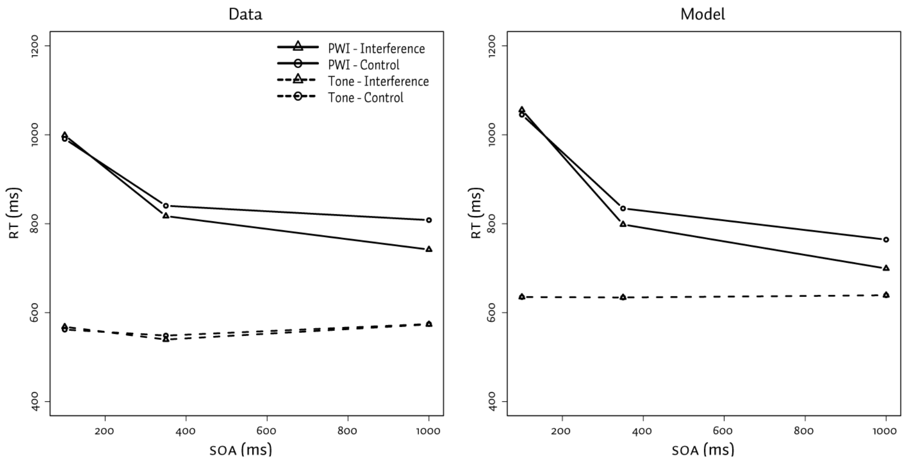


Figure 4.5. Model fit to the data of Dell'Acqua **et al.** (2007). PWI denotes responses on the PWI task, Tone denotes responses on the tone classification task.

perceptual encoding. However, although all information is available for further processing, production rules associated with the PWI task will not be selected because of the instruction to first finish the tone classification task. In the lower panel of Figure 4.6, an incongruent word-picture pair is presented, which results in a longer perceptual encoding stage. However, as the PWI production rules have to wait for the tone task, this increased processing time does not affect the response latency. Of course, if the SOA is long, the interference effect is still apparent in the PWI response latency because the increased latency in the perceptual system is not absorbed in the delayed response (Figure 4.4).

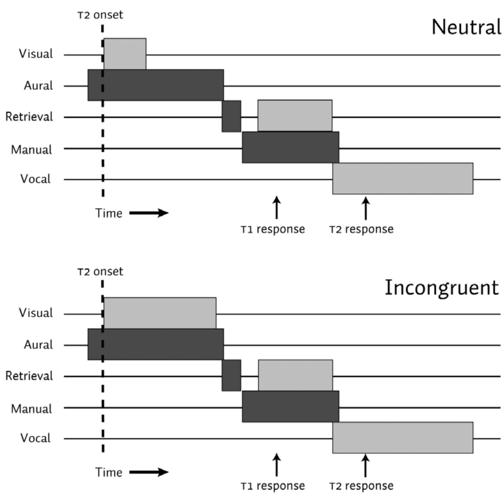


Figure 4.6. Gant diagrams of the model's processing of the tone classification task (dark grey) and the PWI task (light grey). Each line denotes activity in a specialized module, indicated on the left. The processing times of the PWI task are the same, irrespective of the duration of the perceptual encoding stage, manifested as activity in the visual module. The arrows indicate the responses for the two tasks, after which the output modules remain occupied with finalizing the motor action required for the response.

DISCUSSION

In this study, we reconciled the empirical findings of Dell'Acqua et al. (2007) with the vast body of evidence towards a one-mechanism theory (as reviewed by MacLeod, 1991). Using computational cognitive modeling, we demonstrated that it is possible to account for both Stroop and picture-word interference effects in a single mechanism model. Besides this point, the model contributes to other discussions as well.

The first discussion our modeling resutls contribute to relates to theoretical accounts of the PRP effect. The model implemented the PRP effects by assuming an adaptive executive control structure (Meyer & Kieras, 1997b) that controls the order in which the two responses are made. After the perceptual encoding stage of the second task (in this case, encoding of the Stroop or the PWI stimulus), the control structure locks out execution of the second task until the response of the first task has been initiated. This assumption reflects the typical task instructions given in PRP tasks, in which participants are either instructed to respond to the first task before the second task or the importance of the first task is stressed. In the PRP-Stroop and PRP-PWI experiments conducted by Fagot and Pashler (1992) and Dell'Acqua et al. (2007) the correct response order was stressed as well.

Although the model diverges in this respect from the standard central bottleneck model of three stages (input, bottleneck, output, e.g, Ferreira & Pashler, 2002; McCann & Johnston, 1992; Pashler, 1994; Welford, 1967, 1980), it does not refute the logic of the experimental PRP design. The executive control structure creates a cognitive slack time as well in which the initial interference can be absorbed, similar to the central bottleneck theory. However, the models presented here do support the view that such a bottleneck is not a necessary assumption for the PRP effect (Meyer & Kieras, 1997b).

Because the model implements the executive control theory of the PRP effect, we are in a unique position to predict behavior in a PRP tasks in which response order is not stressed. If the model's executive control structure is relaxed, that is, if the model is not constraint in the response order, the model's behavior does not show a PRP effect. Indeed, dual-task experiments have been conducted in which response order was not a factor, and these do not display the PRP effect (e.g., Greenwald & Shulman, 1973; Hazeltine, Teague, & Ivry, 2002; Schumacher, Lauber, Glass, Zurbriggen, Gmeindl, Kieras, & Meyer, 1999; Schumacher, Seymour, Glass, Fencsik, Lauber, Kieras, & Meyer, 2001; but see Levy & Pashler, 2001; Ruthruff, Pashler, & Klaassen, 2001 for opposing evidence). One caveat that should be made however is that the participants in these studies were trained to perform the task concurrently. Therefore, it might also be that practice, not response order was the main cause of the decrease of the PRP effect. This explanation would also be in line with experimental and model findings that show that the PRP effect also decreases with practice in a PWI variant of the task (Van Maanen, Van Rijn, & Taatgen, submitted).

Another discussion related to the picture-word interference effect is concerned with the locus of the interference. The early locus that was reported by Dell'Acqua et al. (2007), might suggest that theories that interpret PWI as a lexical selection effect (e.g., Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992) are wrong. This view is supported by evidence that lexical selection is subject to central processing and therefore does not take place before the central bottleneck in the mental processing stream (Ferreira & Pashler, 2002). By contrast, our model assumes that interference is distributed over many stages, and that depending on the task and on task instructions lexical selection is either a key process, or appears as one of the many sub processes. This is in line with the observation that the Stroop asynchrony is specific to the verbal response modality. If manual responses are required, less emphasis is placed on the lexical selection, and Stroop asynchrony is often absent (MacLeod, 1991).

A third discussion our modeling efforts contribute to is the discussion between single-stage and multiple-stage accounts of semantic interference. The hypothesis that semantic interference might be distributed over different processing stages is not new (e.g., McClelland, 1979). Multiple studies (e.g., De Houwer, 2003; N. Janssen, Schirm, Mahon, & Caramazza, 2008; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005; Van Veen & Carter, 2005) show that the locus of interference is not fixed for a particular phenomenon. Following the assumptions underlying our model, we postulate that semantic interference is associated with a particular sub *process*: the retrieval of declarative knowledge from memory instead of with the outcome of that process during a specific stage of the task. This assumption is in line with studies that suggest a dissociation between stimulus-related interference and response-related interference in the Stroop task (e.g., De Houwer, 2003; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005).

CONCLUSION

The most important point that our computational modeling efforts make, is that the observed difference between the Stroop effect and PWI effect under PRP conditions should not necessarily be interpreted as different cognitive mechanisms. Instead, interference may be driven by one single mechanism, if we assume that different types of stimuli are encoded at different speeds. With this assumption, a single computational model could account for the data that Dell'Acqua et al. (2007) and Fagot and Pashler (1992) observed, while doing justice to the many theoretical accounts that assume that PWI and Stroop are two manifestations of the same mechanism.

Introduction to Part II

In Part I, we studied how the context of a stimulus influences the retrieval processes that are triggered by that stimulus. For this purpose, we studied small-scale behavioral effects and provided explanations for the observed effects by extending the declarative memory system of the ACT-R architecture of cognition. The most important conclusion that can be drawn from Part I is that memory retrieval is not a static, ballistic process, but rather a dynamic process that is adaptive to current environmental demands.

In Part II of this thesis, we will apply this insight in a larger context, and study how prior knowledge and the current environment interact in predicting which concepts will be retrieved from memory. The cognitive models that we will develop in this Part, will be used for recommending relevant items, first in the cultural heritage domain (Chapter 5), and then in the domain of scientific literature search (Chapter 7). This way, we will explore whether these so-called recommender systems may be based on principles from cognitive science. That is, the main concern in Part II of this thesis is whether in principle it is possible to develop useful recommender systems that incorporate models of memory. Part I of this thesis will act as the theoretical basis of these models, in the sense that many concepts developed or discussed in Part I will be applied in the context of cognition-based recommender systems in Part II.

Chapter 5 will discuss how a memory model can be used to develop an artwork recommender system that selects artworks for presentation that are interesting to an individual museum visitor (the Virtual Museum Guide or VMG). The VMG combines the perceived interests from its users with its knowledge on the museum's collection to provide a personalized museum tour through the online collection of the Rijksmuseum Amsterdam (www.rijksmuseum.nl).

Chapter 5 will also discuss the results of a user study that we performed with the VMG. In this study, the visitors that were presented a tour constructed by the VMG did not rate the tour as better suited to their interests as visitors who saw non-individualized tours. This unexpected result may be due to two assumptions that we made in developing the VMG. Chapters 6 and 7 report work that assesses the effects of these assumptions.

The first assumption underlying the work in Chapter 5 is that binary feedback ("Interesting" vs. "Not interesting") given to the artworks presented by the VMG is sufficiently detailed to develop an accurate user model of the visitor that reflects his or her art interest. Given that no effect of personalization was found, it might be that this measure was too coarse. In Chapter 6 we tested whether eye gaze is a better measure of interest in artworks than binary feedback. In Chapter 6, we will present a study in the cultural heritage domain, in which the visitor's eye gaze is used to determine which information on the observed artwork is presented next.

The second assumption we implicitly made in developing the VMG, is that the interest of museum visitors is mainly related to the to the cultural-historical value of the artwork. This was at least partly due to the annotations of the Rijksmuseum that focus mainly on the cultural-historical values. As we based our recommendations on these annotations, we modelled the VMG's knowledge on the museum's collection as a semantic network (Collins & Loftus, 1975; Quillian, 1968) of cultural-historic and art-related concepts. Thus, we implicitly assumed that visitors would have a similar interest in artworks that had similar cultural-historical values. This means that we ignored other aspects of the artworks that also may have influenced visitor interest. For example, the use of certain colors, painting or crafting techniques, or a particular arrangement of figures in the scene may also contribute to the appeal a work of art may have to a visitor. These aspects relate to the expressiveness of an artwork (Arnheim, 1954/1974), which is not easily described. For this reason, we turned to the domain scientific literature search (Chapter 7), in which the relation between interest and semantic similarity is clearer. That is, when searching for scientific literature, it is highly likely that a scientist is interested in those papers that share certain keywords with the scientists own published work.

In Chapter 7, we will discuss a recommender for scientific papers that is based on formal models of declarative memory, the Personal Publication Assistant (Publication PA). The chapter discusses how the history of usage of words in published abstracts of researchers can predict which abstracts an individual researcher currently finds of particular interest. We will present an experiment that shows that participants preferred abstracts that were selected for them by the Publication PA to abstracts that the Publication PA deemed uninteresting for them.

In Chapter 8, we will study the performance of a memory-based recommender system for scientific literature search, by comparing it to a set of often-used decision-making algorithms. The reason for this comparison is that, although we refer to scientific literature search as a problem of *information selection*, it can also be perceived as a problem in the broader class of decision-making problems. That is, for every abstract, the recommender system has to decide whether to recommend it. To compare our system with a broad array of alternative techniques, we chose different types of alternative algorithms. Some of the algorithms in the comparison are chosen because they provide a performance benchmark in the decision-making literature (e.g., take-the-best or multiple regression, Gigerenzer, Todd, & the ABC Group, 1999). Others are included because they share features with the Publication PA. The results of Chapter 8 show that the Publication PA outperforms the other competitors that were trained on the same data set for some users, but is outperformed for other users.

Based on the results obtained in Chapters 7 and 8, it seems that it is possible to develop a useful recommender system based on a memory model. However, it is important to consider the domain for which the recommender system is being developed. In particular, it is important to assess whether the cues that are going to be stored in the declarative memory model represent the relevance of the items that may be selected. In the domain of literature search, this is the case, resulting in a positive evaluation of the memory model. However, in the cultural heritage domain, this seems not to be the case (Chapter 5). Here, the recommendations of the VMG did not align with the interests of the museum visitors, possibly because the cues, cultural-historic concepts, did not represent what museum visitors find important when perceiving a work of art.

Artwork selection for a Virtual Museum Guide

This chapter is an extended version of Van Maanen, L. (2007). Mediating expert knowledge and visitor interest in art work recommendation, LWA 2007. Halle (Saale), Germany.

INTRODUCTION

With the advent of online information presentation, cultural heritage institutions are starting to make their collections available online. Many museums already have websites displaying digital reproductions of parts of their collection. Some of these online repositories are annotated, making it possible to search for specific artworks: For example, the website of the Amsterdam Rijksmuseum in The Netherlands is driven by an ontology on art and artists.

With the online presentation of cultural heritage content, new issues arise. While one of the advantages of digitalization and online presentation is the greater accessibility of cultural heritage (e.g., because of better search capabilities, Van Ossenbruggen, Amin, Hardman, Hildebrand, Van Assem, Omelayenko, Schreiber, Tordai, De Boer, Wielinga, Wielemaker, De Niet, Taekema, Van Orsouw, & Teesing, 2007), one of the drawbacks is that there is less control over what is presented to an individual visitor. Cultural heritage institutions have as one of their aims to educate people on history and culture, which becomes harder to realize once the contents of their collection is accessible from anywhere; They can no longer cater the individual interests of museum visitors while maintaining coherence in the presented information. Besides the decreased control that cultural heritage institutions experience, finding interesting artworks in an online museum poses a problem. Just like in a real museum, most online museum visitors are not aware of their specific interests or of the exact contents of the museum's collection (Bell, 2002). Instead, they only have a general impression of what they want to see and what is available. This makes it difficult to adjust the presentation of the artworks to the visitors' personal interests.

Consider the example of a professional, educated museum guide, touring a party of interested visitors through a museum. The guide can (and has to) select information on the artworks from her extensive knowledge that relates to the personal interests of the party, and can choose which artwork to present next from the collection on display. To reproduce a similar personal experience in an online setting, personal interests as well as relationships between artworks have to be known. A successful recommender system for the cultural heritage domain should incorporate both issues mentioned above: On the one hand, it should take care of the educational role of a cultural heritage institution, and on the other hand it should provide an enjoyable and personalized experience.

OVERVIEW

In this chapter we will present an online recommender system that presents artworks from the Amsterdam Rijksmuseum collection. In our approach we will try to model the way a human museum guide will behave while touring a visitor through a museum. In order to achieve this, we will ground the structure of the recommender system in cognitive theories (Anderson, 2007a).

To stress the analogy with a museum guide touring a group of visitors through a museum, we termed the system the Virtual Museum Guide (VMG). The VMG combines the relationships that artworks have to each other with the personal interests of the visitor to arrive at suitable art recommendations. We will first give an overview of the most important aspects of the system, and then discuss each aspect in more detail.

In the system we will present here, the artworks presented online are accompanied by sets of keywords that describe the interesting aspects of the artwork. As these keywords are provided by the museum's art experts, expert knowledge on the artworks and their interrelations are contained therein. We have applied statistical inference tools from natural language research (Landauer, Foltz, & Laham, 1998) to infer how the artworks relate to each other (details will be provided in the section on the Knowledge Base below). This way, all artworks are related to each other with an association value indicating the relevance of one artwork for another. This structure can be thought of as a *semantic network* or *spreading activation network* (Collins & Loftus, 1975; see also Niessen, Van Maanen, & Andringa, 2008; Quillian, 1968).

For the system presented here we opted for the use of an explicit interest indicator using an *Interesting* and a *Not interesting* button. Each time a user indicates interest in an artwork by clicking one of the two interest-buttons, this feedback is stored as a declarative chunk in the VMG's memory. All chunks representing a visitor's feedback form a user profile of that particular visitor (the Visitor Model). A new artwork will be selected by computing the most relevant and interesting artwork, given the current state of the Visitor Model. First, the visitor's interest in the already visited artworks will be assessed. Second, a spreading activation algorithm (described in more detail below) is used to compute a combined measure of interest and relevance.

THE RATIONAL ANALYSIS OF MEMORY

The Virtual Museum Guide's memory is based on a formal theory of human declarative memory, referred to as a rational analysis of memory (Anderson & Schooler, 1991), which we will apply for the development of a recommender system. The key idea of the rational analysis of memory is that human memory is optimally adapted to deal with information that has been presented in the past (Anderson & Milson, 1989; Anderson & Schooler, 1991). Anderson and Schooler (1991) demonstrated that for each declarative fact stored in memory, the probability that that piece of information will be needed in the immediate future reflects the history of usage of that piece of information. That is, information that has been presented recently is more likely to be needed again than items that have been presented in the more distant past. Also, information that has been presented more frequently is more likely to be needed again. Often, the probability that information will be needed in the immediate future is represented by a quantity called *activation*. The declarative memory representation consists of small pieces of declarative knowledge, called chunks, which together represent a person's long-term memory.

The two environmental observations (recency and frequency) have crystallized (Anderson et al., 2004) into the following equation:

$$B_i = \ln\left(\sum_{j=1}^{n} \frac{1}{\sqrt{t_j}}\right) \qquad \text{(equation 5.1)}$$

B_i represents the base-level activation of a chunk (indicated by the index i). Equation 5.1 captures the effect of frequency of presentation by summing over multiple presentations, and the effect of recency of presentation by dividing by the square root of each presentation time lag (represented by t_j), that is, the time since each presentation of a chunk. This equation has been used in numerous studies predicting memory retrieval effects, both for theoretical purposes (e.g., Anderson et al., 1998; Van Maanen & Van Rijn, 2007b) and for application-based research (e.g., Pirolli, 2005; Van Maanen et al., in press).

Besides the frequency and the recency with which chunks are encountered, also the contexts in which they are encountered add to their activation. The context effects between chunks are represented by an association value that reflects the likelihood that two chunks have co-occurred in the past (Anderson & Milson, 1989). The association between chunks predicts how relevant each chunk is in the context of another chunk.

We used a rational analysis of memory to develop a model of a museum guide. The VMG will remember (and forget) the interests of the visitor in a way that is similar to human behavior. That is, the VMG will compute which artworks are relevant in the context of its recollection of a visitor's previous feedback. We incorporated the association between artworks by computing the *semantic similarity* between two artworks, based on the likelihood of co-occurrence of certain keywords that relate to the artworks.

VIRTUAL MUSEUM GUIDE

Figure 5.1 presents a flowchart of the Virtual Museum Guide. We start out with extraction of a Resource Description Framework (RDF) specification of each artwork from the online ARIA (Amsterdam Rijksmuseum InterActief) repository, which can be inspected at http://media.cwi.nl/sesame/[11]. The RDF specification is transformed to an associative network structure, called the Knowledge Base. The Knowledge Base contains the knowledge the VMG has on the artworks and their interrelations. Besides the knowledge on the artworks, the VMG maintains a user profile, called the Visitor Model. The Visitor Model incorporates which artworks the visitor has already visited, as well as the interests the visitor had in those. Based on both knowledge sources, the VMG selects a suitable artwork and displays it for the visitor, together with some background information on the artwork.

11. To inspect the RDF repository, select Topia's RDF Aria for Sesame in the drop-down menu and select one of the read actions. More information on how to query this repository can be found on openRDF.org.

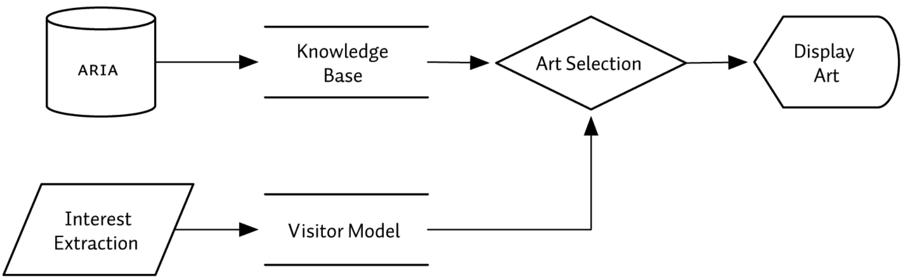


Figure 5.1. A flowchart of the Virtual Museum Guide.

KNOWLEDGE BASE

To be able to recommend a sequence of paintings based on their similarity, we chose to represent the artworks in the Knowledge Base by an associative network structure (Collins & Loftus, 1975; Quillian, 1968). The association values in the network indicate similarity: The stronger the association between two artworks, the more similar they are. The associative values in the network are based on the frequency statistics of the keywords that occur in the RDF specifications of the artworks. The general idea is that two paintings of which the keywords greatly overlap might be considered similar to one another.

To determine the similarity of artworks, we applied Latent Semantic Analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997; Landauer, Foltz, &

Laham, 1998) on the scaled frequency vectors representing the artworks. For this we used the standard TF-IDF weighting scheme (Salton & McGill, 1983), which scales the frequency of terms in a document by the number of documents in which the terms occur. In the VMG's knowledge base, this means that the frequency of keywords in all RDF specifications is taken into account.

At this point, it is important to note that LSA is more than just a correlation of frequency counts (Deerwester et al., 1990; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). Instead, LSA depends on a mathematical analysis (singular value decomposition) that is capable of a higher-order inference. For example, let us assume that the specification of Rembrandt's *The Night Watch* contains the key word *claire-obscure,* and Gerard van Honthorst's *The Merry Fiddler* contains the key word *caravaggists*. The human museum guide knows that these two artworks are similar, because of the two terms are used in similar cultural-historic contexts.[12]

If *caravaggists* and *claire-obscure* co-occur in similar, yet other, RDF specifications, LSA is capable of making a similar inference as the human museum guide. For example, the RDF specification of Dirck van Baburen's *Prometheus Being Chained by Vulcan* might mention the keywords *caravaggists, light, dark,* and *contrast,* and the RDF specification of Rembrandt's *Ecce Homo* might contain the keywords *light, dark, contrast,* and *claire-obscure*. Because both *claire-obscure* and *caravaggists* co-occur with the keywords *light, dark,* and *contrast,* LSA infers that they are related. In a sense, LSA estimates the likelihood that the word *claire-obscure* would occur in the specification of *The Merry Fiddler,* and the likelihood that *caravaggists* would occur in the specification of *The Night Watch*. For a more detailed, but still non-technical introduction to Latent Semantic Analysis, the reader is referred to Landauer, Foltz, and Laham (1998).

12. The Caravaggists are a group of Dutch painters that were strongly influenced by the Italian painter Caravaggio. Caravaggio was the first to study the use of shades, and light and dark contrasts. The Caravaggists in turn influenced Rembrandt's usage of light-dark contrasts.

After LSA-values have been computed for all keywords and for all RDF specifications, each RDF specification can be represented by a vector of LSA-values for that specification. The similarity between two RDF specifications is computed by calculating the cosine between the vectors (Salton, Wong, & Yang, 1975). The cosine between two vectors represents their angle, which indicates how much the RDF specifications differ.

VISITOR MODEL

To select relevant artworks, the long-term, factual knowledge that is stored in the Knowledge Base, is combined with VMG's knowledge on visitor interests, stored in the Visitor Model. The rationale of the Visitor Model is that a museum guide perceives how the visitor feels about a particular artwork, but forgets these interests over time, similar to declarative chunks. Therefore, the perceived interest (PI_i) in a particular artwork *i* may be represented by

$$PI_i = \ln\left(\frac{fb_i}{\sqrt{t_i}}\right) \qquad \text{(equation 5.2)}$$

in which t_i is the time stamp of the presentation of an artwork, and fb_i represents the feedback that the visitor provided when he or she was presented with that artwork. If the artwork appeals to the visitor, $fb_i = 1$; if the visitor is not interested in the artwork, $fb_i = -1$. Because visitors encounter each artwork only once, the summation that is present in Equation 5.1 has been left out.

Figure 5.2 presents an example of the dynamics of the perceived interest in two different artworks. When the visitor sees Artwork A, he or she indicates interest. This is represented by the positive PI-value. However, over time, the VMG forgets the attitude that the visitor had towards the artwork, indicated by the decay of PI. In this example, the visitor disliked Artwork B, which is indicated by the negative PI-value. This PI-value also decays, but increases to indicate that the amount of negativity decreases.

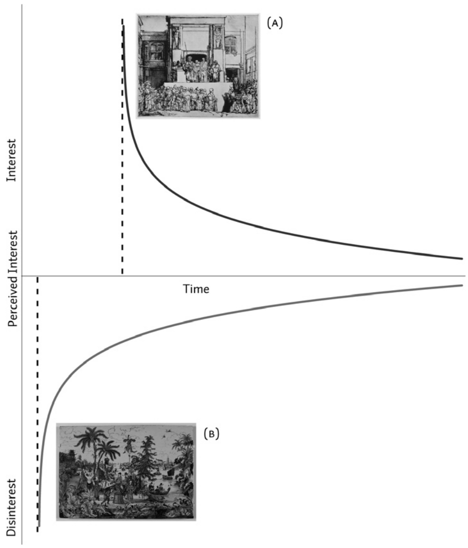


Figure 5.2. An example of the dynamics of the perceived interests. The visitor has expressed interest in artwork A and disinterest in artwork B. The perceived interest decays over time, decreasing the influence it has on the art selection.

ART SELECTION

The selection of artworks depends on a weighted scheme of visitor interest and similarity between artworks. Thus, for the selection of a new artwork for presentation, the spreading activation from already visited art is computed. Art that is rated as uninteresting spreads negative activation (because $fb_i = -1$); art that is rated as interesting spreads positive activation (because $fb_i = 1$). In addition, the spreading activation is scaled according to the similarity between artworks. Thus, artworks that are highly similar spread relatively more activation towards each other. Because of the inclusion of the recency component in Equation 5.2, the influence of recently presented artwork is higher than the influence of artwork presented longer ago. These considerations result in the following equation (Equation 5.3), in which R_i represents the relevance of a certain artwork *i*, PI_j represents the perceived interest in already presented artworks (*j*), and S_{ji} represents the similarity between artworks *i* and *j*.

$$R_i = \sum_j PI_j S_{ji}$$

(equation 5.3)

This equation represents a match between the Visitor Model, represented by the perceived interests, and each artwork that has not yet been presented. The artwork with the highest relevance will be selected next for presentation.

Because PI_j can be either a positive value or a negative value (depending on the visitor's feedback), artworks that were considered uninteresting decrease the relevance of related artworks, while artworks that were considered interesting increase the relevance of related work. Thus, the relevance of an artwork will be high if a visitor expressed interest in related artwork, and did not expressed disinterest in related artwork. Similarly, the relevance will be low (that is, negative), if a visitor only expressed disinterest in related artworks.
After an artwork has been selected for presentation, a web page will be generated that contains a digital reproduction of the selected artwork and some information on the artwork (Figure 5.3). These snippets of information are taken from the Rijksmuseum database, to ensure its correctness and relevance.

EXPERIMENT

In this section, we test whether the combination of an activation-based visitor model and an associative network-like Knowledge Base is useful for art recommendations. To this

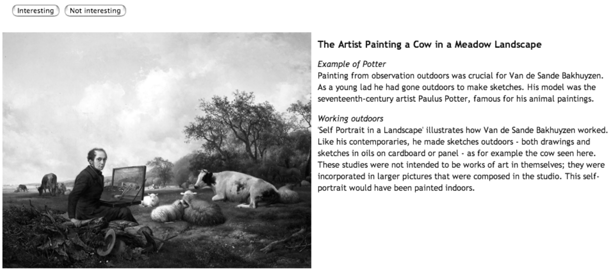


Figure 5.3. The interface of the Virtual Museum Guide. This interface has been used in the training block of the Experiment. In the experimental block, the feedback options were extended to six.

end, we have presented people with artworks that were either selected by the VMG or not, and had them indicate how much they appreciated the selection. The idea is that after a training phase, the VMG should be capable of inferring the participants' interests from the feedback they had already provided. Therefore, artworks that are selected by the VMG should be better aligned with the participants' interests than when the artworks are randomly selected.

To study this hypothesis, we developed three conditions. The first two conditions used the cognitive model outlined above (they were termed the Knowledge condition and the VMG condition, respectively); in the third condition the artworks were selected randomly (the Random condition). In the Knowledge condition, the model incorporates the factual knowledge of the museum guide (the Knowledge base), but did not take the visitor feedback into account. In the VMG condition, the model weights perceived interests and the similarity between artworks.

The assumption is that user satisfaction in this study will correlate with the feedback that the participants will give on the presented artwork. If a participant is satisfied with a certain art selection algorithm, he or she will give more positive feedback on the artworks than negative feedback. By analyzing the feedback per condition, we can probe the user satisfaction with a particular selection algorithm. We thus hypothesize that the participants will provide more positive feedback in the VMG condition than in the Knowledge condition, and even less in the Random condition.

In addition, the participants had to indicate their agreement with a set of statements for each condition to determine their attitude towards the different art selection algorithms. A second hypothesis is that participants have a more positive attitude towards the VMG condition than to the Knowledge condition, and again the least positive attitude towards the Random condition.

PARTICIPANTS

Twenty-five undergraduates (19 female) from the University of Groningen participated for course credit. The participants' age range from 19 to 25 (mean of 21.5). All had normal or corrected-to-normal vision. All were proficient speakers of English.

DESIGN & PROCEDURE

The participants started with a training block of twenty art presentations of historical paintings. Of these, five presentations included landscape paintings, five presented still lifes, five were portraits of historical figures, and five presented genre pieces[13]. The set of training art presentations was the same for all participants, but the order was randomized. The experimental block consisted of three sequences of art presentations, each consisting of ten items. The sequences differ as to the art selection algorithm used. The order of the selection algorithms was different between participants.

13. Genre pieces are paintings that depict scenes from everyday life. The most famous example of this style is The Milkmaid by Johannes Vermeer.

In the Random condition, ten art presentations were randomly selected from the complete ARIA database. In the Knowledge condition, ten semantically related art presentations were selected. In the VMG condition, the ten art presentations were selected based on the semantic similarity *and* the feedback that participants gave in the training block (that is, the VMG was used to select the artworks).

Each trial consisted of the presentation of an artwork with the accompanying background information, taken from the ARIA database (Figure 5.3). The display also contained feedback buttons that the participants could press to indicate his or her attitude towards the artwork presentation. In the training block, there were two feedback buttons (labeled *Interesting* and *Not interesting*). In the experimental block, there were six feedback buttons, ranging from *Extremely interesting* to *Extremely uninteresting*.

Each sequence of ten artwork presentations ended with a small questionnaire consisting of six statements. Three statements related to the associative nature of the sequence of artworks, while the other three statements related to the personalization aspects. The statements were adapted from a questionnaire on usability aspects of an artwork recommender (Cramer, Evers, Ramlal, van Someren, Rutledge, Stash, Aroyo, & Wielinga, 2008). The participants could indicate their agreement with the statements on 6-point Likert-type scales, ranging from *Very strongly agree* to *Very strongly disagree* (0-5). The statements are provided in Table 5.1.

The participants were tested in pairs, and were allowed to take as much time as needed to read the information and study the artwork.

	Type	**Statement**
1.	Associative	The artworks that this museum guide selects relate to each other.
2.	Associative	I think that this museum guide has no consistent story line in mind when selecting artworks.
3.	Associative	I think the artworks that this museum guide selects form a coherent sequence.
4.	Personalization	I think this museum guide does not understand why I like certain artworks I rated as interesting.
5.	Personalization	I think that the artworks that this museum guide selects correspond to my art interests.
6.	Personalization	The artworks that this museum guide selects do not interest me.

Table 5.1

RESULTS

One participant was excluded for an excessive negative attitude (all responses were in the two most negative feedback options, of which 90% was in the single most negative feedback option). The feedback per condition of the remaining participants is presented in Figure 5.4. An analysis of variance did not show any effect of our condition manipulation (m_{VMG}=2.49; $m_{Knowledge}$=2.40; m_{Random}=2.48; $F(2,23) < 1$). The same general result was obtained when analyzing the statements. Figure 5.5 presents the scores on the Likert scales for each statement. For this analysis, the scores on Statements 2, 4, and 6 are inverted. Therefore, higher scores reflect a more positive attitude towards the experimental manipulations than lower scores. For analysis, we aggregated the values of the three Associative items into one value, as well as the values of the three Personalization items. An analysis of variance with

Condition as factor showed no significant effect on the Likert scores for the Associative items (m_{VMG}=2.51; $m_{Knowledge}$=2.28; m_{Random}=2.50; $F(2,23) = 1.65$; $p = 0.20$), and no significant effect on the Likert scores for the Personalization items (m_{VMG}=2.38; $m_{Knowledge}$=2.14; m_{Random}=2.53; $F(2,23) = 2.36$; $p = 0.11$).

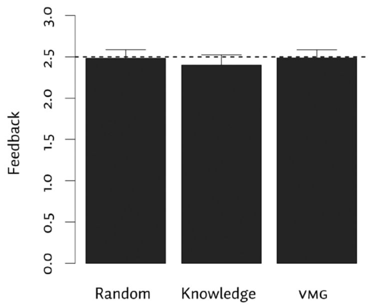


Figure 5.4. Mean feedback per condition. Error bars denote standard errors.

DISCUSSION

Following our assumption that feedback reflects user appreciation, we did not find that the participants appreciated the VMG more than a random selection, or a selection in which only factual knowledge of the museum collection is taken into account. The first reason for the lack of appreciation from the VMG users might have to do with the way the information is conveyed. In the experiment, the virtual museum tour consisted of a sequence of HTML pages containing an image of the artwork and some textual information on the artwork and the artist. One reason why the results of our study are not as expected might have to do with explainability (Cramer et al., 2008). The study by Cramer et al. suggest that users of art recommenders prefer explanations of *why* a certain recommendation is given. Since this extra layer of explanation is not included in the VMG experiments, this might account for the negative feedback that the participants of the experiment gave.

This issue is illustrated by Figure 5.6. In this example, the visitor has given positive feedback to *The Night Watch*. The visitor has provided negative feedback to the chalk drawing *Resurrection* by Lucas van Leyden. If the visitor is now presented with *Ecce Homo*, a chalk drawing by Rembrandt, this may seem as an incorrect recommendation, because of the similarity between the disliked artwork (*Resurrection*) and *Ecce Homo*. A "reason-giving" extension the VMG that explains that this recommendation follows from the visitor's perceived interest in Rembrandt's work may increase the visitor's appreciation. One option of providing this extension may be by reasoning over the sequence of artworks (cf. Bocconi, 2006), the recommender system might offer meaningful and connective extra information on the overlap

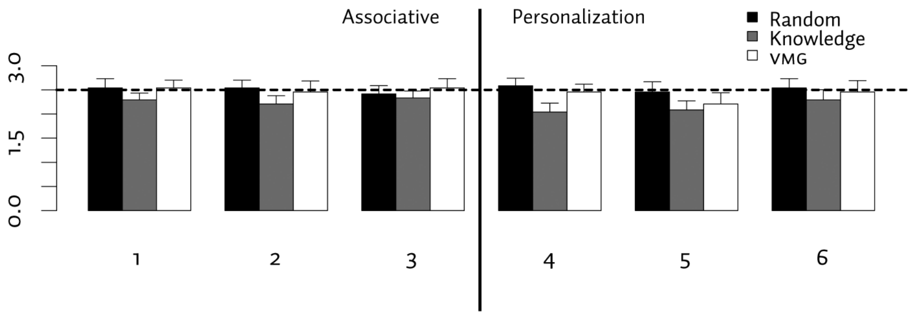


Figure 5.5. Scores on the Likert scales for the six statements. Error bars denote standard errors. The dashed horizontal line indicates chance levels.

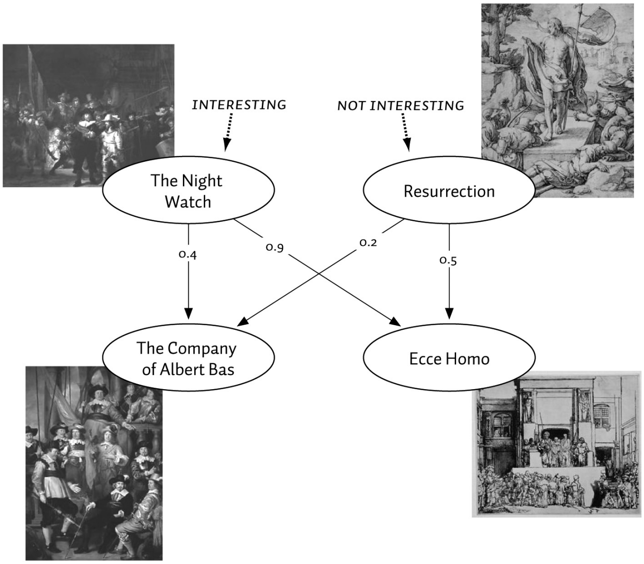


Figure 5.6. An example of part of an art associative network. This image illustrates how the VMG selects relevant artworks in the presence of artworks that are related to the already presented works.

between consecutive artworks (e.g., Falkovych, Cena, & Nack, 2006; Stock, Zancanaro, Busetta, Callaway, Kruger, Kruppa, Kuflik, Not, & Rocchi, 2007). One recommender system in the cultural heritage domain that seems to have implemented this approach is PEACH (Stock et al., 2007), in which personalization is included by having the user select one of four guides at the start of the tour, each with his or her own focus on the works presented.

DISCUSSION AND CONCLUSION

RELATED WORK

The key features of the VMG are the combination of the spreading activation network structure of the knowledge base combined with the decaying level of visitor interest. Also, the generation of the knowledge base using Latent Semantic Analysis is an important aspect, as well as the dynamic generation of web content.

Although most of these features have been applied in previous information presentation tools for the museum domain, the combination we apply is, to our knowledge, unique. Also, most other applications focus on the presentation aspects of dynamically generated content, especially in the context of a real, non-virtual, museum, where the mobility of the visitors poses specific challenges for the presentation of information (Hatala & Wakkary, 2005; for a review see Raptis, Tselios, & Avouris, 2005; e.g., Stock et al., 2007). A third obvious difference between related work and our approach is that while most applications focus on the personalized presentation of *background information* with an artifact, personalization in the VMG involves the selection of the museum artifacts themselves. In this section, we will discuss two systems that seem to be most similar to ours in the key features we have identified for the VMG. That is, both systems - ec(h)o (Hatala & Wakkary, 2005) and PEACH (Stock et al., 2007) - are constructed around a conceptual network, in which selection of concepts is mediated by expressed visitor interests.

Similar to VMG, PEACH (Stock et al., 2007) also adopts an activation-based network. Since PEACH's main output modality is video, the nodes in the network represent video segments, and the edges represent semantic relations between these video segments. Interest expressed in one video segment propagates as activation through the network to all related other segments, and new information will be presented based on the activation values of all video segments. This seems to be a similar approach as the VMG deploys, although the level of semantic relatedness in PEACH is less fine-grained, since the relations in PEACH are hand-coded. In the VMG on the other hand, the Latent Semantic Analysis performed on the keywords that represent the artworks ensure that also unexpected – yet relevant – relations may be present.

PEACH also differs from the VMG in the temporal aspects of the relevance feedback. Visitor's expressed interest in a video segment in PEACH does not extend to another artifact, but only applies to the current artwork. Therefore, decay of visitor interest values is unnecessary. Since the VMG is intended for the dynamic selection of artworks, visitor interest must extend to other artworks.

Just like the VMG, ec(h)o (Hatala & Wakkary, 2005) uses a conceptual ontology as a knowledge base. In ec(h)o, the ontology is based on the Conceptual Reference Model (Crofts, Doerr, & Gill, 2003), which is specifically developed for cultural heritage concepts. Selection of information is subsequently established by reasoning over the relationships in the ontology. ec(h)o also has a decay mechanism to ensure that interests that are more recent are more important than older ones. The mechanism implemented in ec(h)o is however not time-based (as is the decay mechanism of the VMG), but rather the interest values of concepts are normalized such that the highest value stays under a certain upper bound. An advantage of that approach could be that a longer visit to an artwork does not result in 'forgetting' of interests, which is a side effect of the way interest decay is modeled in the VMG.

The ec(h)o system differs from the VMG and PEACH in the way relevance feedback can be expressed. Were VMG and PEACH adopt an explicit strategy in which interest as well as disinterest can be expressed, ec(h)o presents the user with three small audio snippets, from which the visitor can choose. The assumption is that the visitor chooses the audio fragment that is the most interesting to him or her. As a result of this design choice, visitors cannot express disinterest. Moreover, they have to base their decision on a small snippet of the actual information, and could well change their minds after they hear all the information. In this sense, ec(h)o does not really incorporate a relevance feedback mechanism.

CONCLUSIONS

This chapter describes an recommender system for sequences of artwork presentations, in which the decision on which artwork to present next depends (1) the semantic similarity between the artwork descriptions, and (2) the feedback that users of the system provide at each artwork presentation. The recommender system, the Virtual Museum Guide or VMG, uses principles from cognitive science to provide recommendations that are similar to actual museum guide behavior. In developing the VMG, we ensured that both the visitor's interests are considered (maintained by the Visitor Model), as well as the relationships between artworks (stored in the Knowledge Base). In the context of a museum, both aspects are important. Because of the educational role of museums, recommending artwork is more than mapping visitor interest on the museum's collection. The museum needs to ensure that the resulting sequence of artworks is coherent and transfers (part of) the museum's message. It seems that the Virtual Museum Guide ensures both aspects in artwork recommendation.

An experiment that tested whether users would assess the VMG's recommendations more positively than randomly selected recommendations did not show an beneficial effect of our current implementation. Two aspects that may relate to this will be further scrutinized in this thesis. In Chapter 6, we will study if a fine-grained representation of interest may be useful in recommendation of art. While the VMG used explicit feedback buttons, the way feedback may be provided can be very diverse, ranging from the simple button presses used here to more unobtrusive methods, including the time spent observing the artwork (e.g, Claypool, Brown, Le, & Waseda, 2001) or eye gaze analysis. This last option will be explored in Chapter 6.

Another assumption that we implicitly made in developing the VMG, is that the interest that museum visitors expose in art is mainly related to the to the cultural-historical value of the artwork. Thus, we implicitly assumed that visitors would have a similar interest in artworks that had similar cultural-historical values. This means that we ignored other aspects of the artworks that also may have influenced visitor interest. For example, the use of certain colors, painting or crafting techniques, or a particular arrangement of figures in the scene may also contribute to the appeal a work of art may have to a visitor. These aspects relate to the expressiveness of an artwork (Arnheim, 1954/1974), which is not easily described. For this reason, we turn in Chapter 7 to the domain of scientific literature search, in which the relation between interest and semantic similarity is clearer.

Gaze-based personalization for the presentation of cultural heritage

INTRODUCTION

THE NEED FOR PERSONALIZATION IN CULTURAL HERITAGE INFORMATION PRESENTATION

Consider visiting a cultural heritage museum accompanied by a professional, educated museum guide. The guide has much more information to share about each object in the exhibit than the visitor can, or is willing to, grasp in the time available. Thus, the guide has to make a selection from his extensive knowledge of the museum's collection that applies to the artwork currently being attended. At the same time, the visitor expects the guide to provide an interesting and appealing tour. Thus, a good museum guide provides an entertaining tour, in which he adapts the information on each exhibited work of art to the perceived interests of the visitors. A system that provides visitors with a personalized tour should ideally incorporate this, what we could call, interest-based entertainment aspect.

On the other hand, cultural heritage institutions also serve an explicit educational role in modern day society. Often, museums are funded by public means, and have as objective to educate visitors on a particular artist, era, or style. Therefore, an ideal museum guide balances the interests of the visitors, who expect an entertaining visit, and the interests of the museum, that wants to educate the visitors (Bell, 2002).

Besides the obvious advantages of personalization from an entertainment point of view, personalization of a learning experience also offers advantages from an educational stance. For instance, providing extra information that matches the interests of the visitor will extend or even deepen the knowledge visitors gain during the museum visit (e.g., Hsi & Fait, 2005; Stock et al., 2007). At the same time, this might also lengthen the time visitors are willing to spend at a certain exhibit. This is particularly important since the average time spent at each object on display is estimated at about only 30s (Beer, 1987; Cone & Kendall, 1978), which is hardly enough to communicate the bare facts of an exhibited artwork, let alone provide interesting extra information. Therefore, also personalization of information presentation within the scope of a single exhibit seems worthwhile.

One way for a museum guide to probe the visitor's interests is to follow his or her gaze on the work of art. A good museum guide will notice the visitors' gaze, and will adapt his story accordingly. For instance, a long fixation of one of the visitors to a house in the background of a painting might trigger the museum guide to tell more about the purpose of that house for the painting, or about the architecture from the depicted period.

In this paper, we will exploit this feature of human attention to develop a personalized storyteller for cultural heritage presentations, which will be referred to as Gaze-based Personalization for Art or GPA system. First, we will review work that establishes the relation between eye gaze and interest. Next, we will introduce the GPA system, which will be evaluated in a laboratory experiment to demonstrate that adapting stories told at cultural heritage presentations to the personal interests of visitors improves their experience. We will finish with reviewing related work, and by suggesting extensions and possible further applications of our system.

EYE GAZE AS A WINDOW TO ATTENTION

Eye tracking is a technique that has been available for over 40 years. Many aspects of the eye can be measured, including position, pupil dilation, saccadic movements, and fixation duration (Toet, 2006). Essentially, by tracking (one of) the eyes with a camera that is fixated with respect to the position of the head, eye movements can be singled out from head movements. New computational techniques enable eye tracking without fixating the camera position with respect to the head. Instead, the head position is estimated using pattern analysis of whole-head video sequences, and the eye movements are computed using the head position estimates (e.g., Babcock & Pelz, 2004; Boening, Bartl, Dera, Bardins, Schneider, & Brandt, 2006; Li, Babcock, & Parkhurst, 2006).

EYE GAZE IS AN INDICATOR OF OVERT ATTENTION

The direction of gaze and attention correlate to a high degree. Therefore, eye gaze can be regarded as an indicator of attention (Henderson, 2003). Although it has been known for a long time that fixations and attention may deviate (Posner, 1980), during free viewing of natural scenes they are very likely to align (Findlay & Gilchrist, 2005). In this sense, eye movements can be regarded as a behavioral indicator of the spatial allocation of attention.

In addition to the correlation between eye gaze and attention, eye gaze and informativeness are also related (Henderson & Hollingworth, 1999). Several studies (Antes, 1974; Loftus & Mackworth, 1978; Mackworth & Morandi, 1967; Yarbus, 1967) established the relationship between fixations on different regions of a picture and the informativeness of these regions. An interesting aspect of this relationship is that the fixation patterns are directly related to the goal that the participants have. For example, a request to memorize a scene results in different fixation patterns than viewing the painting with the goal of determining the wealth of the people depicted (Yarbus, 1967). Mackworth and Morandi (1967) and Antes (1974) based their measure of the informativeness of a region on ratings that users gave to the different regions. Thus, they established a relationship between eye gaze and users' *idea* of informativeness. Under the assumption that when perceiving an artwork in a museum, a visitor's goal is to find those regions on a painting that are most interesting to him or her, these studies suggest that eye gaze is an indicator of interest that jmuseum visitors might have in specific regions of realistic paintings.

Various researchers (e.g., Chen & Zelinsky, 2006; R. M. Cooper, 1974; Huettig & Altmann, 2005; Yee & Sedivy, 2006) have shown the relationship between semantic content and eye gaze. In an experiment that involved eye fixations and spoken words, Cooper (1974) showed that people fixate more on regions of interest (ROIs) on the screen that they hear in a snippet of spoken text than on ROIs that are not mentioned in the spoken text. In addition, ROIs that bear a semantic relation to words mentioned in the spoken text are fixated more than ROIs that are unrelated to the text. These results demonstrate that eye gaze is also mediated by the semantic content of the display. Again, if we assume that one of the goals of museum visitors is to find interesting aspects on each work of art on display, then a prolonged gaze on a certain object in the painting may be the result of the overlap between the semantic content and the interest of the visitor.

EYE GAZE AS A POINTING DEVICE

In many Human-Computer Interaction applications, gaze has been deployed as an explicit pointing device, much like a computer mouse. Typically, users may interact with an interface using stares or blinks to indicate that an action needs to be performed (e.g., Hornof & Cavender, 2005; Jacob, 1991). Although the user interaction is mediated by gaze in our application as well, it explicitly differs from these applications because we do not intend users to control the presentation of information with their gaze, but rather analyze the free viewing behavior to determine the most interesting regions of an artwork.

The above reviewed studies suggest that it is possible to develop a system that personalizes background information on the basis of eye gaze when presenting a work of art. In what follows, we will discuss how we implemented this idea, and we will present an experiment in which that implementation is put to the test.

GAZE-BASED PERSONALIZATION FOR ART

The Gaze-based Personalization for Art (GPA) system that we propose uses the eyes' gaze as an indicator of interest, and presents information on paintings based on the perceived interests of the user. The information about the artworks is presented using natural speech. Because the selection of information depends on the direction of gaze, which differs between individuals, the story that is told at each artwork also differs between individuals.

The sum of the duration of the various fixations on different regions of interest (ROIs) on a painting determines the choices made by the GPA system. This summation will be referred to as gaze duration.

ANNOTATING ARTWORKS

Before interesting audio snippets can be selected, each painting has to be annotated. That is, the spatial dimensions of every ROI on a painting that might be of interest to a virtual museum visitor of the system have to be identified and related to an interesting piece of information. In GPA, the spatial dimensions of the ROIs are represented in a *content map* that indicates which pixels on the screen belong to which ROI (Figure 6.1). Regions of interest (ROIs) on the content map can be anything: a figure or a group of figures, objects, buildings, animals, et cetera. In addition, ROIs may consist of multiple components: Several smaller areas on the display can combine to one ROI. An example of this is if there are multiple animals on the artwork, and the audio snippet explains something about the use of animals in paintings of a certain genre.

Figure 6.1. Content map of the painting "The Sacrifice of Iphigenia" by Jan Steen. In this content map, ten regions of interest (ROIs) are identified on which information can be presented.

Because ROIs may differ in size, bigger ROIs would have a higher probability of being attended than smaller ones, if there would be no control of eye movements. However, the control of eye movements is accurate enough to ensure that most of the fixations are on the intended ROIs. We therefore assumed that the number of fixations that land on an unintended ROI would not significantly influence ROI selection.

DETERMINATION OF INTEREST

The choice of which audio snippet to present depends on the recorded fixations. The recorded fixations are represented by a *fixation map* (Velichkovsky, Pomplun, & Rieser, 1996; Wooding, 2002), which is a two-dimensional array in which the gaze duration on a specific region of the screen is represented. Depending on the need for spatial resolution of the fixations, the dimensions of the array could vary between the pixel dimensions of the display – providing maximum spatial resolution – or one or more orders of magnitude smaller than that. Because in the context of the GPA system we want to identify which ROIs on a painting are being fixated, we are only interested in whether fixations are within the boundaries of these ROIs. Since in our study the average ROI size is relatively large (43 pixels in diameter, see also Figure 6.1), we set the size of a patch on the fixation map at 10 x 10 pixels. The pixel dimensions of the display we used in the interest-aware system were 1024 x 768, and therefore the fixation map consisted of 102 x 77 patches of 10 x 10 pixels (with the patches on the right side of the display 8 pixels wide, and the patches on the bottom of the display 14 pixels high). Because the number of fixations is typically very large, we followed Wooding's (2002) suggestion that each fixation could be represented by the fixation location only, instead of a Gaussian distribution reflecting noise in the eye gaze recording. It is unlikely that this simplification will result in different information selection, because the ROIs we defined were much larger than the variance in the Gaussians.

Over time, the durations of new fixations are added to the already accumulated values of the fixation map. In the interest-aware system, the updates to the fixation map are made after successive presentations of audio snippets. Thus, the selection of information is based on a user's gaze from the initial presentation of a painting up until the moment of information selection.

REGION OF INTEREST SELECTION

The interest that people show in the ROIs on the painting is estimated with the following equation:

$$I_i = \sum \mathbf{D}\mathbf{C}_\mathrm{i} \text{ with } \mathbf{C}_\mathrm{i} = \begin{cases} 1 \text{ if } \mathbf{c} = i \text{ with } \mathbf{c} \in \mathbf{C} \\ 0 \text{ otherwise} \end{cases}$$

(equation 6.1)

I_i indicates the interest in ROI *i*, **D** indicates the array representing the fixation map, and $\mathbf{C}_\mathrm{i}$ indicates the array representing which cells belong to ROI *i*. $\mathbf{C}_\mathrm{i}$ can be derived from the original content map by substituting all cell values (**c** in Equation 6.1) by 1 if and only if they belong to ROI *i*, and by zero otherwise.

By combining the fixation map and the content map, we arrive at a set of gaze durations representing individual interest for each ROI on the screen. Based on these values the selection of information can be performed by selecting the ROI with the highest interest that has not been selected before, and presenting the associated audio snippet.

In developing the GPA system, we experimented with different algorithms that estimated the probability that a fixation on a certain ROI was the result of intentionally looking at that

ROI, instead of by other factors, such as saliency differences between various regions on the painting (Itti & Koch, 2001; Itti, Koch, & Niebur, 1998; Koch & Ullman, 1985), or saccadic noise (Kowler & Blaser, 1995). Many studies show that, besides attention-mediated processes, eye gaze is also controlled by the physical properties of whatever a person is watching (e.g., Henderson, 2003; Kootstra, Nederveen, & De Boer, 2008; Theeuwes, 1992). The typical distinction is between top-down control of eye-movements, in which some cognitive operations are performed, and bottom-up control, which refers to eye movements caused by features of the visual field. Bottom-up control is usually linked to *saliency*, a concept representing the differences in physical properties of the visual field, such as color, contrast, orientation, or motion. A study by Theeuwes (1992) demonstrates that if a highly salient distractor is presented in a visual search task, participants make an eye movement to this distractor before fixating on the search target. Therefore, given that the visual acuity of a region on a painting is high enough, it will cause the eyes to fixate there. However, when freely viewing natural scenes (such as classical paintings), the influence of saliency on the fixation *duration* is typically much less than the influence of top-down modulating factors (such as informativeness of the region, or interestingness, Henderson & Hollingworth, 1999). Thus, it is interesting to see whether in the context of artworks, the saliency of an artwork contributes to the direction of gaze, and whether saliency should be considered when developing gaze-based personalization systems.

PRESENTING INFORMATION USING SPEECH

When the most interesting ROI is selected, the associated audio snippet can be presented. Obviously, when a subsequent ROI needs to be selected, all already presented ROIs are excluded from selection. Care was taken in the construction of the audio snippets that they only referred to a single ROI, thus diminishing the chance that gazes to other ROIs could be caused by reference to the semantic content of those ROIs (R. M. Cooper, 1974; Huettig & Altmann, 2005; Yee & Sedivy, 2006).

We chose to present the users of the interest-aware system with spoken text. Using another modality than vision for information presentation is useful in the context of attention-aware systems, because the eyes are not distracted by extra visual information that otherwise might appear on the screen (cf., Starker & Bolt, 1990).

VALIDATION OF GAZE-BASED PERSONALIZATION FOR ART

We performed a laboratory experiment to study whether the GPA system outlined in the previous section would prolong user interest when perceiving art. In this study, participants were asked to look freely at a series of paintings while a voice-over gave information on these paintings, and to indicate for each painting when they lost interest in that particular painting. We hypothesize that when participants receive information on ROIs they implicitly expressed interested in with their gaze, they would stay interested longer and indicate their loss of interest later.

EXPERIMENTAL MANIPULATIONS

To contrast the behavior of participants interacting with the actual GPA system, we designed three information-selection conditions. In the first condition, henceforth referred to as the *maximum* condition, the participants received information via audio snippets for ROIs in which they showed maximum interest. The maximum condition is contrasted with two control conditions: The *random* condition and the *minimum* condition. In the random condition, participants were presented with the information in a random order, not depending on their gaze durations; in the minimum condition, participants received information pertaining to the ROIs for which they expressed the least interest (that is, the ROIs with the lowest gaze durations). Our main hypothesis is that viewing time on a painting is mediated by these information-selection conditions: Longest for the maximum condition, in between for the random condition, and shortest for the minimum condition. Note that the random condition can be seen as a standard online guide in that the order in which information was presented to the participants was not modulated by knowledge of the participant's perceived interests.

A secondary hypothesis will be that in the maximum condition users will look longer to the ROIs on which they receive information than in the random and minimum conditions. This effect is hypothesized to result from the mapping between the users' interest and the presented information. The rationale of this hypothesis is as follows: Given the assumptions that eye gaze follows interest, and that eye gaze is influenced by the semantic content of presented audio, the gaze durations on selected ROIs in the maximum condition would benefit from both factors, whereas the gaze durations on selected ROIs in the random and minimum condition would only benefit from the latter factor. Due to a random presentation order of the information in the random condition, we expect the gaze duration on presented ROI in the random condition to be in between that in the maximum and minimum conditions. This reflects the probability that a participant is presented with an audio snippet that relates to a highly attended ROI.

PARTICIPANTS

Thirty undergraduate students from the University of Groningen participated in the study. The mean age of the participants was 22.6 years. All participants had normal or corrected-to-normal vision and normal hearing and received course credits for their participation. All were native speakers of Dutch.

APPARATUS

The paintings were displayed on a 19″ CRT monitor. We measured eye gaze with the Eyelink I eye tracker from SR Research. The Eyelink I is a head-mounted eye tracker with a spatial resolution of 0.01°, a sampling rate of 250Hz, and an average gaze position error of 0.5°-1.0°. Note that the average ROI size is 1.2° (an average diameter of 1.7 cm at approximately 80 cm distance). The Eyelink I delivers pupil position and head marker position in real-time (based on four LED-arrays attached to the screen), and can thus account for both eye movements and (small) head movements. Although the Eyelink I is a head-mounted system, the strain on the participants was acceptable. We therefore do not expect differences in viewing behavior as compared to natural viewing behavior.

STIMULI

Paintings

To stay as close as possible to a museum setting, whilst still in a controlled laboratory environment, we used digital reproductions of paintings from the Rijksmuseum collection, with a resolution of 1024 x 786 pixels. The selection of paintings was such that they contained enough identifiable elements to present sufficient audio snippets. The paintings either depicted a biblical or mythological scene, or depicted a scene from everyday live (Genre painting). Most selected artworks were created during the Dutch Golden Age (roughly the 17^{th} century). Figure 6.2 presents examples of the paintings used in the experiment.

(a) (b) (c) (d)

Figure 6.2. Four examples of paintings used in the experiment. (a) "The Sacrifice of Iphigeneia" by Jan Steen (1671) (b) "Prince's Day" by Jan Steen (c. 1665) (c) "The Fall of Man" by Cornelis van Haarlem (1592) (d) "Aeneas at the Court of Latinus" by Ferdinand Bol (c. 1661-63).

Audio snippets

Each painting was accompanied by a sequence of audio snippets. The first, introductory audio snippet had an average duration of 20s (±1.5s). It contained the name of the artist, the year the painting was produced, a comment on technique or a small anecdote regarding the painting as a whole or the artist, and finally the title of the painting. The introductory audio snippets were the same for every participant. We made sure that no direct references to ROIs on the painting were made, except when this occurred because of the painting's title. After the introductory audio snippet a maximum of 10 other audio snippets were presented during 10s each (±1.5s), separated by one-second intervals. The short intervals between the audio snippets ensured that a sequence of audio snippets sounded naturally. The texts of the audio snippets were based on the explanations that accompanied the paintings as found on the website of the Rijksmuseum (www.rijksmuseum.nl). We adapted the texts so they would be self-contained and would fit in a timeframe of 10s. As an example, Appendix A presents the text of the audio snippets for the painting "The Sacrifice of Iphigeneia".

DESIGN AND PROCEDURE

Every participant was tested individually. First, the eye tracker was set up and calibrated, and the participants were instructed on the remainder of the experiment. Second, participants were presented with 26 paintings while hearing a sequence of audio snippets accompanied each painting. Each presentation of a painting constituted one trial. Each trial was preceded by a drift correction screen. The first painting was a practice trial, and was always the same for all participants, as was the sequence of audio snippets presented during this trial.

The participants were instructed that they could indicate loss of interest in the painting or the information by pressing the RETURN key, which would result in a new painting being presented. Pressing RETURN signaled the participants' response.

After the practice trial, 25 paintings were presented in pseudo-random order to each participant. Initially, we also were interested in the effects of saliency on the selection of audio-snippets. However, since there were no effects from the saliency manipulation, we collapsed those conditions with the default conditions. Therefore, ten paintings were presented in the maximum condition, ten were presented in the minimum condition, and five were presented in the random condition. To counterbalance possible effects of fatigue or loss of interest across participants, each painting was presented as often in the first half of the presentation sequences as in the second half of the presentation sequences. Moreover, every painting was presented equally often in the maximum as in the minimum condition.

If a participant pressed the RETURN key, the voice-over finished the current audio snippet, and then moved on to the next trial. If participants listened to all audio snippets of a trial, the system also progressed to the next trial. The set of audio snippets for a painting was the same in every condition; only the order in which they were selected differed, depending on the condition and the gaze of the participant.

RESULTS

As the main dependent variable is the listening duration, two participants were excluded from further analyses as they listened to only a single audio snippet in more than 50% of the trials, and three participants were removed because they listened to all audio snippets in more than half of the cases. This leaves 25 participants for further analysis, of which the average listening duration in the initial trail was 69s (SD=42s) and in the experimental trials 76s (SD=31s).

Although the main analysis of interest is the effects of different audio-selection conditions on listening duration, we will first turn to the effects of saliency. The saliency condition was included as the assumption was that saliency could influence eye movements independent of actual interest. However, no differences were found related to the saliency condition ($F(1,24)<1$) or to the interaction of saliency with audio-selection condition ($F(1,24)<1$). We have therefore collapsed the data over Saliency for all subsequent analyses.

In contrast to our hypothesis, there was no effect of Audio-selection on listening duration (means: $m_{maximum}$=76.5s, $m_{minimum}$=74.3s, m_{random}=76.8s, $F(2,24) = 1.01$; $p=0.37$). Additional exploratory analyses including different personal characteristics (e.g., age, gender, reported interest in art, etc) or performance measures (e.g., ratio of gazes on ROIs) as co-variates did not change this outcome.

Based on the assumption that interest is reflected by increased gaze durations on the regions of interest, the second hypothesis was that participants look longer to items on the artwork they are interested in than to items they are not interested in. To test this hypothesis, we studied the gaze duration on ROIs *during* the presentation of audio snippets.

Gaze durations *before* the presentation of an audio snippet were used to select audio snippets in the maximum (and minimum) condition. Thus, the second hypothesis is that *during* the presentation of audio snippets, gaze durations on the ROIs associated with the presented audio-snippet should be longer in the maximum condition than in the random or minimum conditions. Furthermore, to rule out the possibility that the participants had not seen the ROI under consideration, we excluded all instances in which the critical ROI was not fixated before the presentation of an audio snippet (this led to one participant being excluded that had not fixated the critical ROIs in one condition).

Figure 6.3a presents average gaze durations on the presented ROIs during the presentation of the associated audio snippets. The gaze durations between conditions differ significantly ($F(2,23) = 110$; $p<0.001$). This may be a first indication that participants will remain interest in ROIs that they gaze at irrespective of the presented audio information. If gaze was primarily determined by the content of the audio snippet (R. M. Cooper, 1974), we would have expected no difference between the bars in Figure 6.3a.

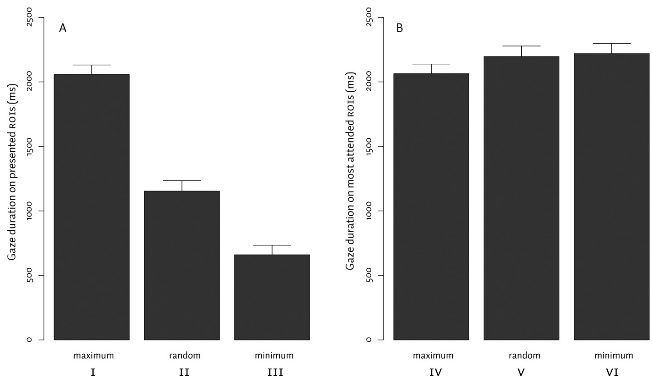


Figure 6.3. (a) Average gaze durations (during presentation of an audio-snippet) on the ROIs associated with the audio-snippets that are currently being presented; (b) Average gaze durations (during presentation of an audio-snippet) on the ROIs that were most attended before the selection of the audio-snippet. Error bars denote standard errors.

A second indication that gaze duration reflects interest is presented in Figure 6.3b. If prior interest guides the eye movements, then the ROIs that the participants previously gazed at should be gazed at again, irrespective of the presentation of an audio snippet. However, if the audio snippet determines the direction of gaze, then we would expect a difference between the gaze durations on the most attended ROIs *during* presentation of the audio snippet. This is because for the random and minimum condition, the presented audio snippet is *not* about the most attended ROIs. Thus, gaze would be expected to deviate from these regions. However, an analysis of variance did not show a difference in gaze duration in Bars IV, V, and VI (means: $m_{maximum}$=2064ms, $m_{minimum}$=2219ms, m_{random}=2167ms, $F(2,23) < 1$). We did not find evidence that the content of the audio snippet determines which ROI will be gazed at more, suggesting that interest is a stronger determinant of eye gaze than the content of an audio snippet.

DISCUSSION

In this experiment, we demonstrated the role of eye gaze in interest detection. Regions on a painting that are selected because they are being attended are gazed upon to a larger extent than regions that are selected because they were not or hardly attended (Figure 6.3a). Interestingly, this observation does not result in longer listening durations (the total time spent viewing each artwork). This may be in line with earlier observations that museum visitors have a typical dwell-time they are willing to spend at each exhibited art work (Beer, 1987; Cone & Kendall, 1978). Thus, assuming that the time people are willing to invest in each art work on display is roughly constant, as well as limited, the need for a smart ordering of the information that the museum wants to convey becomes prevalent.

One complicating factor in the interpretation of the results of our study is that eye gaze is implicated in multiple behavioral mechanisms. Thus, other factors could have influenced eye gaze. In the "Regions of Interest Selection" Section, we already mentioned the possible influence of saliency. Areas on the painting that stand out on a certain dimension such as color or contrast are more likely to be attended. However, when viewing paintings, the influence of saliency is much less than the influence of interest (Henderson & Hollingworth, 1999). Moreover, we counterbalanced the artworks in such a way that each artwork occurred in every condition equally often.

In addition, it has been shown that gaze is also affected by the semantic content of an aurally presented message (R. M. Cooper, 1974). This could have affected the results of our study as well. However, Cooper found that fixation durations were *increased* on ROIs that were mentioned in the message. If the eye gaze of the participants in our study was mainly driven by the content of the audio snippets, we would therefore find a difference in gazes on the most attended ROIs in the maximum and the minimum condition. The fact that we do not find this difference (e.g., Figure 6.3b) can be seen as evidence that in our study, the gaze durations are not driven by the semantic content of the audio snippet. One interpretation of this difference between Cooper's study and ours is that in his study, the ROIs were well-defined and known to the participants in advance (because of the clear arrangement of items on the display), whereas in our study, the demarcations of ROIs are less clear, due to the natural scenes depicted. Moreover, the participants were unaware of the existence of ROIs, and may thus be less likely to acknowledge auditory references to the depicted items by a quick fixation.

DISCUSSION AND CONCLUSION

In this paper, we proposed a personalized system for information presentation at art exhibitions, termed Gaze-Based Personalization for Art. The system uses point of gaze to infer a visitor's interest, following prior studies that suggest a correlation between eye gaze and attention. We determined regions of interest (ROIs) for the artworks, which together form content maps of the artworks. By combining the content maps with dynamically updated fixation maps, we can compute the how much each ROI is fixated, and select the most fixated ROIs for information presentation.

In an experiment, we contrasted the GPA system with two control systems that had a random selection mechanism and a negative selection mechanism, respectively. The results show that presented information on the most attended ROIs increases the fixation duration on those ROIs, although the total time spend examining an artwork does not seem to be influenced by personalized information presentation.

Our results suggest that if visual perception of the artwork is important when presenting information, gaze-based ordering of the to-be-presented information is useful. For instance, when the information relates to a specific use of colors (e.g., the use of dark and light colors in the paintings by Rembrandt van Rijn), it is important that the museum visitor attends the regions of the painting on which that technique is exposed. The experiment discussed above suggests that if the visitors are not attending those regions in the first place, they are not likely to attend them when the presented information is about the use of light and dark in Rembrandt's paintings. Therefore, our work may help cultural heritage institutions to adapt the order of the information that they want to convey to optimize the knowledge transfer to the visitor.

RELATED WORK

Personalization of cultural heritage presentation

The idea of personalizing certain aspects of a cultural heritage experience has been studied extensively. Most applications focus on the presentation aspects (e.g., Falkovych, Cena, & Nack, 2006; Hatala & Wakkary, 2005; for a review see Raptis, Tselios, & Avouris, 2005; Sparacino, 2002; Stock et al., 2007). Some work is directed at personalized sequences of the objects at display (e.g., Fink & Kobsa, 2002; Van Maanen, 2007). All of these approaches provide personalization of museum content by adapting which information is being presented to the individual user. To our knowledge, the current study represents the only attempt at using eye-gaze as an informative devive for personalization of cultural heritage content.

Gaze-based information selection

Another relevant concept that has been studied before is the selection of information using eye gaze. Previous work in the gaze-based selection of information has focused on explicit information retrieval systems (e.g., Oyekoya & Stentiford, 2007; Puolamäki, Salojärvi, Savia, Simola, & Kaski, 2005). In the GPA system, the selection of information is implicit, because museum visitors are not intentionally looking at certain regions of the artwork in order to receive information about these regions.

There exists however two systems that share many features with the GPA system. These are iTourist (Qvarfordt & Zhai, 2005) and the gaze-responsive system by Starker and Bolt (1990). Both systems track the user's point of gaze, and present new information that is selected on the basis of where the user was looking. In iTourist, the information is related to tourist information on a virtual city map. iTourist presented spoken as well as visual information on points of interest on the map. The system by Starker and Bolt presents spoken information on objects on a small planet, similar to iTourist and GPA. These systems differ from GPA in that users of both applications are aware of the manipulatory role of their gaze, due to the particular layout of the display (a city map with marked tourist highlights and a rotating planet with isolated objects). By contrast, users of the GPA system are not aware of the manipulatory role of their gaze. Participants in our experiment were not informed on the reason for measuring their eye movements. In addition, the regions of interest on an artwork are less well-defined than the regions of interest on a city map or on a constructed 3D world, which makes it harder to intentionally fixate a certain region. Therefore, it seems to make more sense to use eye gaze in the cultural heritage domain in a diagnostic fashion rather than an manipulatory or intentional fashion (Duchowski, 2002).

MORE NATURAL SETTINGS

Although we tested the GPA system in a controlled laboratory environment, the approach seems suitable for environments that are more natural as well. For instance, the online presentation of sequences of artwork may be augmented with gaze-based interest awareness (Van Maanen, Janssen, & Van Rijn, 2006). With the advent of online information presentation, cultural heritage institutions are starting to make their collections available online. Many museums have websites displaying digital reproductions of part of their collection. Moreover, recent advances in eye tracking technology as well as increased quality of standard webcams have brought online non-intrusive gaze tracking to the desktop (e.g., the COGAIN initiative, www.cogain.org, or Hansen, Hansen, & Johansen, 2001). Using webcams for eye tracking may enable the use of gaze-based interest awareness to adapt the presentation of information on the artworks available at museum websites.

In addition, more advanced eye tracking devices have been developed that allow for more free movement of the user (e.g., Babcock & Pelz, 2004; Boening et al., 2006; Li, Babcock, & Parkhurst, 2006). With these devices, it becomes possible to freely wander through a museum, while your point of gaze is being tracked. This allows for adaptive information presentation in the real museum, for instance using headphones.

CONCLUSION

Eye gaze may provide useful insights in people's interest, which can be used in cultural heritage applications. Detecting personal interests of museum visitors enables personalized presentation of the exhibit's information. This may increase the enjoyment that visitors have when attending an exhibition, but may also improve their learning experience, because the ordering of the presented information may be such that the information aligns with the visitor's prior knowledge and interest. The GPA thus balances a museum's educational role and a visitor's personal interests, just like a good real-life museum guide.

APPENDIX

Example of audio snippet s of "The Sacrifice of Iphigeneia". 0 indicates introductory snippet, which was presented first. 1-10 indicates the remaining snippets, the order of which may be determined by the participants gaze and the condition.

0. Het volgende schilderij is gemaakt door Jan Steen en dateert uit de periode rond 1671. Jan Steen werd vooral bekend om zijn "genrestukken" met vrolijke gezelschappen. Maar ook met portretten en schilderijen over mythische verhalen. Dit werk is gemaakt met olieverf op doek. De titel van het schilderij is "Het offer van Iphigeneia".
1. Boven in de rook zit Artemis, de Godin van de jacht. Ze is herkenbaar aan de maansikkel op haar hoofd en de pijl en boog. Artemis werd ook vereenzelvigd met de maangodin Selene.
2. Jan Steen hield zich niet aan de regels van de historische schilderkunst. Om het verhaal voor zijn tijdgenoten herkenbaar te maken, gaf hij de meeste personen geen oud Griekse, maar 17de eeuwse kleding.
3. Een jongetje loopt bedroefd weg op het schilderij. Het is Amor, de god van de vleselijke liefde. Hij is herkenbaar aan zijn pijl en boog waarmee hij mensen verliefd kon maken.
4. Bij dit geschilderde offer vallen geen doden. Als door een wonder verandert het menselijke slachtoffer, zonder dat iemand het merkt, in een hert. Hierdoor komt alleen het hert om het leven.
5. De schilder beeldt het moment af vlak voordat een offer wordt gebracht. De beul staat al klaar om zijn slachtoffer te doden, het mes glinstert in zijn hand en hij kijkt bloeddorstig naar het offer.
6. De geknielde vrouw is met veel zorg afgebeeld, met details als de glanzende kleding en de voeten die vies zijn van het lopen. Het licht van het vuur rond haar afgewende hoofd geeft haar zelfs iets mysterieus.
7. In het midden ziet u Iphigeneia. Omdat het oorlog is zal zij worden geofferd aan de Goden om hen gunstig te stemmen. Het offeren van mensen, dieren en voorwerpen was zeer gebruikelijk in de Griekse oudheid.
8. Rechts bovenin is een vrouw aan het bidden. De gevouwen handen zijn een christelijk gebaar wat niet hoort bij de Griekse oudheid. Door dit wel te gebruiken is het verhaal herkenbaarder voor de tijdgenoten van de schilder.
9. De figuren rechts hebben een exotisch tintje vanwege hun verschillende hoofddeksels. De tulband, de lauwerkrans, een soort bisschopsmijter en een variatie op een Romeinse helm zijn te zien.
10. De afgebeelde koning is Agamemnon uit Griekenland. Hij wil Troje belegeren, maar telkens als hij wil uitvaren is het windstil. Om de boze godin Artemis gunstig te stemmen moet hij zijn dochter offeren.

English translation

0. The following painting is from Jan Steen en dates from around 1671. Jan Steen became especially well known for his genre paintings with fun-loving, cheerful groups of people. But also with portrait paintings as well as paintings of mythological scenes. This piece is oil on canvas. The title of this painting is "The Sacrifice of Iphigenia".
1. In the upper part of the painting sits Artemis, the Goddess of hunt. She can be

recognized by the crescent moon and her bow and arrow. Artemis was considered the same Goddess as the moon Goddess Selene.

2. Jan Steen did not keep strictly to the rules of history painting. To make the story recognizable for his contemporaries, he has not dressed all the people in 'Iphigenia' in classical Greek costumes, but seventeenth-century clothes.
3. A little boy leaves the scene crying. It is Amor, the God of erotic love. He can be recognized by his bow and arrow that he used to make people fall in love.
4. No one dies in this sacrifice. In a miracle, the human victim is switched with a deer. Consequently, only the deer dies.
5. The painter has pictured the moment just before Iphigenia is to be sacrificed. The executioner is about to kill her, the knife shines in his hand while he is looking cruelly at his victim.
6. The kneeling woman has been depicted with great care, with realistic details such as the shine of her silk clothes and her bare feet that have become a little dirty from walking. The light from the wood fire that shines around her turned head gives this woman a mysterious air.
7. In the center stands Iphigeneia. Because of a war, she will be sacrificed to pacify the Gods. Human and animal sacrifices as well as sacrifices of goods were very common in ancient Greece.
8. In the upper right corner, a woman is praying. Her hands are folded in a Christian manner, which is anachronistic for a scene of ancient Greece. By using this, the story has become recognizable for the painter's contemporaries.
9. The people on the right all have an exotic appearance due to their different headpieces. You can see a turban, a laurel wreath, a miter, and something similar to a roman-style helmet.
10. The depicted king is Agamemnon of Greece. He is determined to lay siege to the city of Troy. Yet, every time he tried to embark his fleet, a calm descended. To sooth the angry Goddess Artemis, he must sacrifice his daughter.

Abstract Recommendations by a Cognitive Model

This chapter is in press as Van Maanen, L., Van Rijn, H., Van Grootel, M., Kemna, S., Klomp, M., & Scholtens, E. (in press). Personal Publication Assistant: Abstract recommendation by a cognitive model. Cognitive Systems Research.

INTRODUCTION

In cognitive science, there has been a long tradition to perceive human behavior as a form of information processing. Within this tradition, human cognitive processes are seen as operating on similar principles or algorithms as computer programs, since both cognition and computer programs have or have been developed to process information. This view has lead to the birth of Artificial Intelligence as an independent research field (McCarthy, Minsky, Rochester, & Lebiere, 1955), but has also guided the development of cognitive theories (e.g., Anderson & Milson, 1998; Marr, 1982; Newell, 1990). Even today, the apparent functional overlap between artificial computational systems and the human information-processing system is still influential in cognitive theorizing (e.g., Griffiths, Steyvers, & Firl, 2007).

Many cognitive theorists believe that human beings optimize their behavior to successfully cope with the environment (e.g., Anderson, 1990; Marr, 1982; Oaksford & Chater, 1998). This means that, through evolution and learning, human behavior has adapted to be the most suitable behavior in any given circumstance or environment. This is a capacity also desirable in artificial systems design, especially when these systems have to operate on an unknown or dynamic environment. Therefore, computer scientists and artificial intelligence researchers have studied how computer systems can optimize their behavior as well (e.g., Goldberg & Holland, 1988; Kohonen, 2001).

A domain that has not benefited that much from this cross-fertilization is the problem of selecting relevant information, either for oneself or for others. The research field that studies how to disclose relevant information is known as Information Retrieval (Salton & McGill, 1983). A typical field in which the problem of selecting relevant information arises is the scientific community. For example, the number of scientific publications in the relatively small ISI subject category *Information Science & Library Science* was 2054 in 2006.[14] This means that researchers working in this area have to read (or at least scan through) over two thousand papers a year to keep up with the current developments. However, this number is, if anything, an underestimation of the total number of potentially relevant papers, as this number only holds if the researcher is interested in a single subject area. In practice, most researchers work on the intersection of multiple domains, increasing the number of potentially relevant papers enormously. In general, because of the continuous increase of storage capacity for digital media, and the increased availability of digital or digitized media sources, companies, institutions, and individual people are being confronted with an increase in the amount of information that potentially is relevant to their purposes.

14. Source: ISI Web of Knowledge, retrieved 19-12-2007.

In this paper, we will describe a system that partly solves this problem for the scientific domain: Our system selects relevant scientific papers from a large collection of scientific abstracts. Instead of working from a pure computer science perspective, we will present a system that is based on constraints from cognitive theories. In particular, we chose to follow the rational analysis approach (Anderson, 1990; Oaksford & Chater, 1998), as incorporated in the ACT-R architecture of cognition (Anderson, 2007a). The rational analysis approach states that human memory is optimally adapted to fit the needs of the environment we live in, based on the interactions of the cognitive agent with the environment in the past. This approach has been successfully applied to predict various aspects of human behavior (e.g., as reviewed by Chater & Oaksford, 1999).

We will begin with an analysis of how users behave when engaging in the selection of information. Next, we will discuss how the ACT-R cognitive architecture incorporates rational analysis, and how this can be applied to information selection. We will continue with an outline of an application based on the resulting model, the Personal Publication Assistant (or Publication PA for short), and how this application behaves under different conditions, as well as a user study that will demonstrate the applicability of our approach in a real world setting. In the last section, we will discuss in what way the Publication PA deviates from other approaches towards the task of matching papers to researchers, or vice versa.

INFORMATION SELECTION

An example of the problem addressed in this paper is the selection of relevant information when attending a large, multi-track scientific conference. Often, an attendee finds him or herself overwhelmed by the amount of presentations that can be attended. With so little time to find the talks that are really interesting, changes are that one ends up in the wrong track, listening to presentations that hardly kindle ones interest, while in another track relevant work is being discussed. Although this might bring unforeseen beauty, often a better selection of relevant work would be preferable. There are solutions to this problem. For example, giving the attendees the proceedings well in advance so they have more preparation time. However, this solution is often not viable due to practical constraints. A better solution might be to provide an automatic recommendation based on the personal interests of the conference attendees, which is the approach that will be discussed in this paper.

To build a successful recommendation system, it is important to know how the selection process takes place in unsupported settings. The information selection process starts when a researcher registers at a conference and receives a copy of the conference proceedings. Based on informal analyses, the next step is to perform a quick scan of all titles, author names, or abstracts for words or names that are familiar. If an entry contains enough interesting words, it is selected for further and more careful reading. Obviously, the assumption that is made implicitly, is that individual words in the abstract accurately reflects the contents of the paper or presentation. Ries et al. (Ries, Su, Peterson, Sievert, Patrick, Moxley, & Ries, 2001) have shown that this assumption holds for abstracts and papers, at least in the medical domain. In order to determine if a word qualifies as interesting in the context of the conference, the researcher might assess whether she has used the word in her own research in the past. One could say that the researcher tries to discover the degree of familiarity she has with an abstract, and if that degree of familiarity is high enough, she selects that presentation as potentially worthwhile to visit.

To assist a researcher in the information selection task, we propose a model of the recognition aspects of the task. That is, we propose a model that makes a preselection from the available information based on a notion of familiarity adapted to the individual researcher. To achieve this, we will develop models of the declarative memory systems of individual researchers (henceforth refered to as user models) and of the process of recognizing words. Each user model can be seen as a representation of an individual researcher's interests, as it incorporates the frequency, recency, and context of the words used by the researcher to describe her research. In previous research (Anderson & Milson, 1989; Anderson & Schooler, 1991), a formal model has been developed of how the retrieval of declarative facts from memory can be described. In the next section, we will give a detailed overview of that model, but we will highlight the two most important aspects here. One key idea is that declarative memory is optimally adapted to serve the needs of the cognitive agent (Anderson, 1990; Oaksford & Chater, 1998). The other is that most facts in declarative memory are initially formed by perception (Anderson & Schooler, 1991). Combined, this means that the adaptive nature of declarative memory is essentially a reflection of the perceptions of the cognitive agent. As a consequence, this means that looking for structure in the environment can derive the structure of declarative memory.

RATIONAL ANALYSIS OF MEMORY

Anderson and Schooler (1991) showed that the probability that a memory will be needed in the near future depends on the pattern of prior exposures to the piece of information stored by that memory. For example, the probability that someone will contact you by email today depends on the frequency and recency of her emails to you in the past (Anderson & Schooler, 1991). Likewise, the probability that you will need some declarative fact from memory right now depends on the frequency and recency of the prior usage of that fact. Both relations are captured by Equation 7.1, in which *B* stands for the base-level activation (reflecting the probability), t_i stands for the time since exposure to event *i*, and *d* represents the speed with which the influence of each exposure decays. The summation is over all (*n*) previous encounters of the events (*i*).

$$B = \ln\left(\sum_{i=1}^{n} t_i^{-d}\right)$$

(equation 7.1)

Besides frequency and recency of usage of declarative facts, the context in which these facts occur also plays a role in the activation of these facts. This activation component will be called the spreading activation (Quillian, 1968), and represents the likelihood that one declarative fact will be needed if another one is currently being used. These likelihoods depend on the pattern of prior exposures with the declarative facts, as represented by the relatedness measure R_{ji} between two facts j and i (Anderson & Lebiere, 1998; Anderson & Milson, 1989):

$$R_{ji} = \frac{F(W_j \,\&\, W_i)F(N)}{F(W_j)F(W_i)}$$

(equation 7.2)

where $F(W_j)$ and $F(W_i)$ are the frequencies of respectively fact *j* and *i*, $F(N)$ the total number of exposures and finally $F(W_j \,\&\, W_i)$ is the number of co-occurrences of the facts *j* and *i*. Equation 7.2 is sometimes refered to as associative strength (Anderson & Lebiere, 1998; Anderson & Milson, 1989), to indicate that the relatedness between two facts is determined by the environment. The model of declarative memory outlined here has been successfully deployed in predicting behavior in a variety of memory related cognitive tasks (e.g., Anderson et al., 1998; Anderson & Schooler, 1991; Van Rijn & Anderson, 2003).

IMPLEMENTATION OF THE PERSONAL PUBLICATION ASSISTANT

The Personal Publication Assistant is a personalization tool based on a personalized rational analysis of memory. Therefore the user models underlying the recommendations are constructed on an individual basis. In these models, each word that occurs in one of the abstracts of the user is represented by a combination of base-level activation (adapted from Petrov, 2006) and spreading activation from the other words in the model (Anderson &

Lebiere, 1998). These activation values can be calculated using the statistical properties of the words in the published abstracts of an individual researcher:

- The year in which it appears for the first time in one of the user's abstracts,
- The year in which it most recently appears in one of the user's abstracts,
- The frequency of appearance,
- The frequency of co-occurrence with another word.

Based on these properties, we create an individual representation of a researcher's interests using the rational analysis described above. The Publication PA applies these individual user models to predict the relevance of words that occur in other scientific abstracts, by calculating how familiar these abstracts are. In the next sections, we will describe in more detail how the Publication PA calculates the base-level and spreading activation values, which words from the abstracts are taken into consideration, and how the system comes to a selection of the relevant information.

THE RELEVANCE OF INDIVIDUAL WORDS IN THE USER MODEL

With the equations that are provided by the rational analysis approach to declarative memory, we can calculate the base-level activation of a word based on its occurrences in publications of the user. The base-level activation can be seen as a measure of interest, with the most interesting words having the highest base-level activation.

For this application, an optimized version (Petrov, 2006) of the base-level equation discussed earlier (Equation 7.1) was used. In this equation (Equation 7.3), the decay parameter is fixed at .5 (and is reflected in Equation 7.3 as the square root operators) and a history parameter (*h*) is added:

$$B = \ln\left(\frac{1}{\sqrt{t_1 + h}} + \frac{2n-2}{\sqrt{t_n} + \sqrt{t_1 + h}}\right) \text{with } h > 0 \qquad \text{(equation 7.3)}$$

The first component of this equation reflects the most recent encounter of that word: the longer ago the word was encountered, the smaller the contribution is. The second component reflects the frequency of usage of the word. This optimized version of the base-level activation equation assumes that the encounters of the word are evenly spaced over time between the first encounter and the last encounter of the word. In the default equation, the base-level activation is a product of both recency and frequency. However, in a recommendation system, it might be useful to be able to change the balance between both factors. For example, a researcher might still be interested in work relating to older work, even though a recent project has resulted in a set of papers on a new topic. To enable this, we added the history parameter. The history parameter influences the effect of recency. Informally, a higher value for *h* spreads the publications over a longer time frame, decreasing the relative activation of a word that only recently came up in analyzed texts. In Experiment 1 we will demonstrate that the *h* parameter is an important parameter when recommending papers with the Publication PA.

THE INFLUENCE OF CONTEXT ON WORD RELEVANCE

Apart from the frequency and recency of usage of a word, the context in which a word occurs is also important. For instance, using the word *model* in your paper on user models should not elicit conference talks on fashion models. So, context words - like in this example *user* or *rational* - are important in determining the activation of words such as *model* or *analysis*. The context in which a word has occurred in previous abstracts is incorporated in the model by spreading activation (Equation 7.2), which reflects the personalized probability that a word will be needed in connection with another word.

Recommendations occur by mediating the base-level activation of a word with the spreading activation of other words:

$$A_i = B_i + \sum_j WR_{ji}$$

(equation 7.4)

In Equation 7.4, the base-level activation of the word *i* in a specific abstract is increased with the sum of all weighted connections with the words also found in that abstract. The connections are weighted because otherwise the ratio between the base-level activation and the spreading activation would be dependent on the number of associations. For this application the base-level activation of the connecting word (*j*) is used as the weight (*W*), to scale down with the spreading activation from words that have a low base-level activation. This would be the case when the word *i* co-occurred often in the past with a word *j* that is present in the current abstract but which is not often used anymore (i.e., has a low base-level activation). This would cause the spreading activation to be high while the connection is less relevant at the current time, negatively influencing the selection of relevant papers.

Filtering of non-content words

The relatedness measure R_{ji} has shown to be a robust method of boosting the base-level activation as a function of the connectedness. That is, if two words always occur in tandem, the activation of the second word will be boosted when the first word is encountered. At the same time, a word that occurs in combination with many other words does spread less activation. In normal word usage, words as *the* and *is* spread only a small amount of activation because of this. In normal word usage, this effect makes sure non-content words do not influence base-level activations of other words too much. However, a problem might arise when the formulation of sentences in scientific abstracts differs from normal word usage. Because of spatial constraints, word usage in scientific abstracts might differ from normal written English. This might result in a lower frequency of function words, increasing their spreading activation (Equation 7.2), with a possibly negative influence on the eventual recommendations. To counter the unwanted influence of normally high-frequent words, these words are filtered from the data using a lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). An analysis of the frequency distribution of words in both scientific abstracts and normal written English will demonstrate that filtering out high-frequent words will not interfere with how well an abstract represents the contents of a paper.

Analysis

To compare word usage in scientific abstracts with word usage in normal lexical content, the abstracts of all publications that appeared in the *Cognitive Science Journal* between 2004 and 2006 were used. Numeric symbols and punctuation were removed from the abstracts, resulting in a list of the words that were used in the abstracts. For each word, the frequency in all the abstracts was contrasted with an estimate of the normal frequency in written English, taken from the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). If a word was not found in the database because of spelling mistakes or terminology, the CELEX frequency was assumed 0, and the frequencies of multiple occurrences of a word were summed because in CELEX the frequencies of homonyms are counted separately. The CELEX frequencies were scaled to the total number of words of the abstracts to make them comparable.

Results

In Figure 7.1, the ratio between the CELEX word frequencies and the abstract word frequencies is plotted. We used a logarithmic scale for easier presentation. Figure 7.1 visualizes that the usage of words in scientific abstracts differs from the distribution of words used in normal written text. If the distributions were similar, then the dashed horizontal line would have represented the ratio. However, it becomes clear that a large part of the words used in the abstracts occur less often in normal written English; those are the words with a frequency ratio below one. Only a small part of the words occurs more often in normal written English. Thus, 2190 of the words used in the abstracts of the *Cognitive Science Journal* between 2004 and 2006 occur more frequently in scientific abstracts than in normal written English, while only 412 words occur more often in normal written English. However, those 412 words account for a large portion of the total amount of word occurrences found in the CELEX database (440,000 of the total of 740,000 occurrences of these words), while the 2190 words that are less frequent in normal written English generate less occurrences than the 412 high frequent words (300,000 of 740,000 word occurrences). This difference is caused by abstracts containing jargon and the tendency to use as little function words as possible, whereas in normal language these words are used very frequently. Thus, removing the words from the scientific abstracts that are most frequent in normal written English will not remove any of the important content words, as only words above the dashed line in Figure 7.1 are deleted, while words below the dashed line in Figure 7.1 are the words that are relevant to the Publication PA.

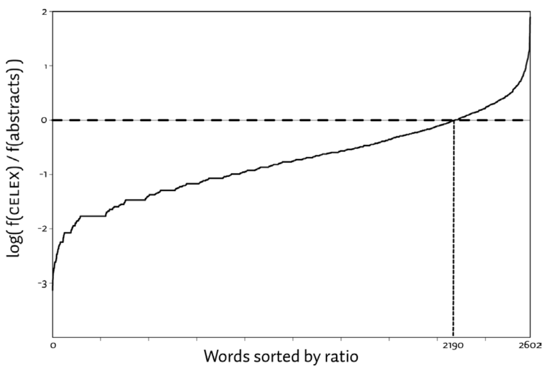


Figure 7.1. Log ratio of word usage frequencies in scientific abstracts and normal written English (CELEX), sorted by increasing frequency ratio. The dashed line indicates when words are used more often in scientific abstracts than in normal usage.

SELECTION OF RELEVANT ABSTRACTS

The final part in the recommendation is finding the amount of activation for each paper and presenting the user with a ranking or selection. In general, abstracts in which many words have a high activation, have a high degree of familiarity to the researcher, and are thus interesting enough to select. To compare the relevance of papers with each other, every abstract has to be represented by a single value. One solution would be to sum the activations of all the words in the conference abstract. However, simply summing activation values would result in a bias towards longer abstracts. To counteract this bias, we chose to average the activation of the words that occur in both the abstract and the user model. This means that the effect of abstract length is neutralized, while still taking all activation values of the words in the abstracts into account.

EXPERIMENTS

To validate the Publication PA, we first analyzed what the influence of the *h* parameter is. Second, we performed a user study with a sample of researchers from the field of cognitive science, asking them to rate how much a recommended abstract aligned with their interests.

EXPERIMENT 1: HISTORY PARAMETER ANALYSIS

Methods

We analyzed the behavior of the Publication PA with four different values for the history parameter: $h = 0.0001$; $h = 0.1$; $h = 10$; $h = 1000$. The parameter values were chosen to maximize a potential effect. The only other parameter in the system (the decay parameter *d*) was left at the default value of 0.5.

As a test set, we took the abstracts of the publications of professor John R. Anderson, for as far as indexed by PsycINFO[15]. When visually inspecting his publication record, it shows some stable interests over time, but also some changes in interest. As a cognitive modeler, almost all of Anderson's publications deal with cognition and the cognitive architecture he developed, ACT-R. However, a change in focus can be observed. From the start of his career, Anderson's interests seem to be related to learning and memory (as witnessed by for instance Anderson & Bower, 1972, 1973), whereas more recently he seems to have developed an interest in functional brain imaging techniques (e.g., Anderson, 2007b; Anderson, Albert, & Fincham, 2005). These trends should also be visible if we apply different parameter values to the *h* parameter and construct different user models.

15. http://psycinfo.apa.org/

Results

To compare the user models that were constructed with the various values for the *h* parameter, we ordered the words in the user models according to their activation values. Thus, the ordering represented the estimated importance of a word for a person's interest. Figure 7.2 presents the rank order values of various words that are exemplary of the trends found in Anderson's publication record. Small values of *h* indicate that the relative influence of more recent publications increases; this effect is reflected by the decreasing rank (and thus increasing importance) of the words *functional* and *imaging* for decreasing values of *h*. These words do all relate to the recent research interests. On the other hand, the words *memory* and *experiments* show the opposite trend. This reflects a shift of interest from proto-typical memory-related research in which multiple experiments are presented per paper. Also, the ranks of some words stay constant with changing *h* values. ACT-R and *cognitive* are words that appear in both recent and past abstracts of professor Anderson, indicating a stable interest in these concepts.

This qualitative inspection of the results leads us to believe that the history parameter plays an important role in the selection of relevant abstracts, because it determines the ranking of the activation values. What the optimal setting for this parameter should be might be determined in a large user study in which we ask participants to rate the relevance of abstracts that are selected using various values for the history parameter (as has been done for this analysis). However, given the personal nature of interest, it seems better to leave the optimal setting to the user, for example, by presenting the user with the possibility to set this parameter in the user interface. To evaluate the performance of the Publication PA independent of the relative importance of word usage history, we decided to run the user study with *h* set to 10.

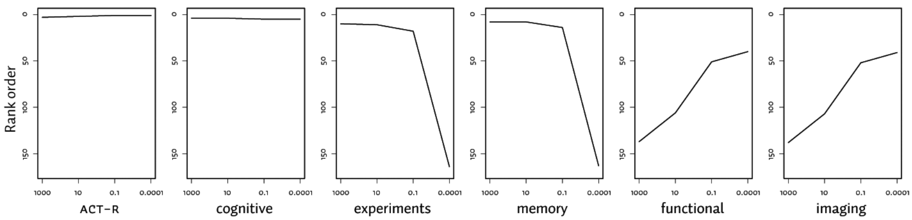


Figure 7.2. Rank order of example words that occur in the user models created for one researcher with varying history parameter values. The history parameter values we tested are indicated on the x-axis. Low rank order values indicate that a word is important for determining the researchers interests. If h decreases, the relative influence of words that were used in the past also decreases, and the influence of recently used words increases.

EXPERIMENT 2: USER STUDY

We performed a user study to evaluate the recommendations provided by our abstract recommender system. We asked 10 researchers (2 full professors, 2 associate professors, 5 assistant professors, and 1 post-doc) from various subfields of cognitive science and from various countries how much they are interested in a paper after reading the abstract.

Methods

For each of the researchers, we constructed user models based on the abstracts of their published work insofar it was indexed by PsycINFO. Next, we ordered all abstracts from the last three volumes (2004-2006) of the *Cognitive Science Journal* according to their relevance for an individual researcher, based on the researcher's published abstracts.

From the ordered list of abstracts, we presented the top five abstracts, the bottom five abstracts (that is, the least relevant abstracts), and five abstracts from the middle of the list to each researcher. The presentation order of these 15 abstracts was randomized, to eliminate any effects from expectations about the presentation order. We asked the researchers to indicate with a grade between 0 and 9 how much they are interested in the papers, based on the abstracts. We adopted this scale from similar work done by Dumais and Nielsen (1992) in order to be able to make a comparison between their approach and ours. Following Dumais and Nielsen (1992), we characterized the meaning of the rates as follows:

- 8-9: right up my alley
- 6-7: good match
- 4-5: somewhat relevant
- 2-3: I'm following it, sort of
- 0-1: how did I get this one?

Results

To analyze the performance of the Publication PA, we applied two measures of relevance:

- Mean rated relevance,
- Precision.

The precision and mean rated relevance were applied to each of the three groups (top 5, middle 5, bottom 5). Because it is not feasible for the participants to rate all available abstracts from the *Cognitive Science Journal* between 2004 and 2006 (129 abstracts), we did not calculate the rate of recall, as is often used in these kinds of applications (Salton & McGill, 1983). However, the recall rate is implicitly accounted for in the measures we did apply.

Mean rated relevance

We analyzed the relevance rates given to the abstracts for each group. Figure 7.3 shows the means of the rates per group. Welch t-tests between the groups reveal that the rates given for the top 5 abstracts differ significantly from the other two groups ($t=4.20$, $df=86.54$, $p<0.001$ for the top 5 vs. the bottom 5 and $t=3.64$, $df=94.06$, $p<0.001$ for the top 5 vs. the middle 5). The rates for the bottom five abstracts did not differ significantly from the rates for the middle five abstracts. This is in line with the observation that in multidisciplinary journals such as Cognitive Science, the relevance rate does not decrease linearly, but instead that only a small part of the published papers is relevant for a researcher, and the rest is not.

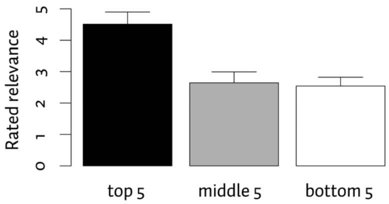


Figure 7.3. Mean ratings per group. Error bars denote standard error. The mean ratings that the participants provided for the top 5 recommended abstracts is significantly higher than for the other two groups.

If the Publication PA would not be able to suggest relevant papers, this would mean that in all three groups the number of highly rated papers would be equal on average. However, if this were the case, we would not be able to observe significant differences in the mean rated relevancies between the top 5 recommended papers and the other two groups. The fact that we do find this difference indicates that the system is able to provide a meaningful rank order in which the higher rated papers will be ranked higher.

Precision

Precision of retrieval is usually defined as the number of relevant documents that is retrieved relative to the total number of documents retrieved (Salton & McGill, 1983). Following the meanings of the anchor points of the scale we provided to the participants, relevance should be taken as rated with 4 or higher. Using Equation 7.5, the precision of the Publication PA in the top 5 recommended abstracts is p = 0.58.

$$p = \frac{\left|\{\text{rating} > X\} \cap \{\text{retrieved abstracts}\}\right|}{\left|\{\text{retrieved abstracts}\}\right|} \quad \text{(equation 7.5)}$$

Because this notion of relevance may be considered arbitrary, we also calculated the precision of the Publication PA with different assumptions on relevance. For example, we calculated precision under the assumption that only abstracts rated 8 or higher where relevant, or that all abstracts rated 2 or higher were relevant. In Figure 7.4, the results of this analysis are presented. The figure shows that, although precision declines with a more stringent notion of relevance, the precision in the top 5 recommended abstracts is always higher than in the other two groups.

DISCUSSION AND CONCLUSION

With our experiments, we demonstrated both the flexibility of the Publication PA and its applicability. With only one parameter, we could change the recommendations of the system in such a way that the relative influence of older papers changed, resulting in different recommendations.

With the *h* parameter at a fixed value, we demonstrated that the Publication PA can provide meaningful recommendations for individual users. Two observations from this experiment should be further discussed.

From both the precision measure and the mean ratings, it becomes clear that there is no real difference between the group of abstracts from the bottom of the order list of abstracts from *Cognitive Science Journal* (2004-2006) and the 'middle' group. This shows that from a large collection of papers, only a very small subset is relevant for a particular user, underlining the need for filtering mechanisms or recommender systems.

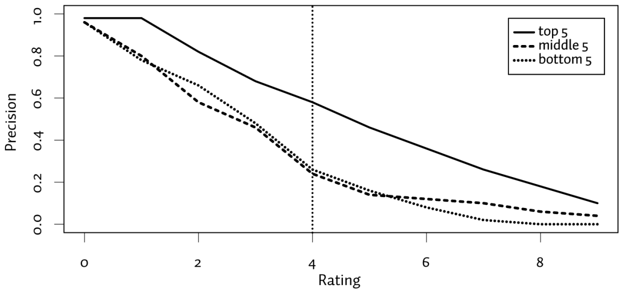


Figure 7.4. Precision of the Publication PA for different interpretations of relevance. The dotted vertical line indicates the point at which relevance is interpreted as somewhat relevant or better (rating 4 on the scale provided to the participants). The figure shows that for all interpretations of relevance, the precision for the top 5 recommended papers is higher than for the other two groups.

Figure 7.4 shows that the mean rated relevance for the top 5 recommended abstracts is 4.5. This qualifies as *somewhat relevant*, but not *right up my alley*. We contribute this to the nature of the data set we used to recommend abstracts from. *Cognitive Science Journal* is a highly multidisciplinary journal, accepting papers from a wide range of research areas (as witnessed for instance by the set of keywords authors can use when submitting, published on the website of the Cognitive Science Society[16]). As a result, papers addressing very specific topics, that may be *right up my alley*, will be presented to other, more specialized, journals. Thus, the ratings provided by our participants might be a bit lower than expected, because abstracts that would be rated as *right up my alley* were probably underrepresented in the data set.

16. http://cognitivescience-society.org/journal_csj_submission_keywords.html

RELATED WORK

The problem of matching researchers and papers has been addressed before, in the context of systems that use Latent Semantic Indexing (LSI) (Dumais, 2003; Dumais & Nielsen, 1992; Foltz & Dumais, 1992). Our approach deviates from these earlier attempts in a number of ways. LSI assumes that the similarity of two documents is reflected by the similar word frequency distributions that are manifest in these documents (Deerwester et al., 1990; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). However, instead of taking the raw frequency statistics into account, LSI performs a mathematical analysis (singular value decomposition) that is capable of higher-order inference. That is, LSI calculates the probability of each of the words occurring in a document, given multiple documents.

Instead of LSI, the measure of semantic relatedness that we apply, associative strength (Anderson & Lebiere, 1998; Anderson & Milson, 1989), is equivalent to Point-wise Mutual Information (PMI), given a reasonably large data set (Farahat, Pirolli, & Markova, 2004). PMI is also based on the statistical properties of the documents, but, in contrast to LSI, PMI is a direct measure of the likelihood that one word will occur, given the presence of another. As a measure of *semantic similarity*, PMI has been shown to perform equal to or better than LSI (Turney, 2001). We expect therefore that when PMI will also be a better representation of semantic relatedness than LSI.

Besides the method of calculating the semantic relatedness, also the corpus of text on which it is performed differs. Dumais and colleagues (Dumais & Nielsen, 1992; Foltz & Dumais,

1992) used a fixed semantic space for all users of their system. Recently, however, it has been shown that the choice of corpus greatly influences the semantic distance, even when applying the same measure of semantic relatedness (Lindsey, Veksler, Grintsvayg, & Gray, 2007). By contrast, we constructed personalized semantic spaces for individual users. That is, the associations between words in the semantic space reflect the semantic relatedness as apparent from the statistical properties of word usage in the abstracts of one user. This obviously will result in more individualized recommendations, because only the associations between words that a single researcher would also make, are present. Also, the problem of corpus selection does not arise, because the corpus used is already the best possible representation of a researcher's interest, namely her own publication record.

When it comes to the performance of the Publication PA as compared to the approach taken by Dumais and Nielsen (1992), the Publication PA seems to perform equally well. Dumais and Nielsen (1992) report a precision of $p = 0.51$, slightly lower than our value of $p = 0.58$. However, the computation of precision differs between the two approaches. In general, comparison is difficult because of the different nature of the data sets used. While the abstracts from the *Cognitive Science Journal* are very multi-disciplinary and thus very diverse, the abstracts used by Dumais and Nielsen (1992) are from a very specialized conference (ACM *Hypertext'*91). This difference in diversity of topics included in the data sets could explain the difference in mean rated relevance between the Publication PA (4.5) and the system by Dumais and Nielsen (5.75). In the Dumais and Nielsen experiment, both the abstracts and the researchers they are being assigned to, are specialized in hypertext. Therefore, the mean relevance of the data set for the researchers is already higher than in our experiment.

To a certain extent, our work bears resemblance to the work of Pirolli and colleagues towards Information Foraging (Fu & Pirolli, 2007; Pirolli, 2005; Pirolli & Card, 1999; Pirolli & Fu, 2003). They provided a rational analysis of how users search for relevant information, and applied this to information search on the World Wide Web. This way, they were able to model web navigation aspects of a typical user. In Information Foraging theory, the likelihood that a certain document or webpage is relevant is based on the base-level activation of the words in that document and the spreading activation from the words in that document to the words in the search query. Similarly, the Publication PA computes the relevance of a paper based on the base-level activation of the words in the abstract and the spreading activation from the words in that abstract to the words in the user model. One of the important components of the Publication PA is the construction of the user model, which ensures that only words that are relevant for an individual researcher are considered in computing the relevance of an abstract.

However, the Information Foraging models differ in that they capture the information search behavior of a typical human being serving the web, whereas the Publication PA is a personalization tool, and is intended to model the information needs of an individual researcher. As outlined above, the semantic relatedness estimates applied by the Publication PA are therefore personalized for each individual researcher, resulting in different behavior of the model for each researcher.

CONCLUSION

In this paper, we proposed a method for the personalization of information selection, based on rational analysis and cognitive architectures. We developed an application, the Personal Publication Assistant (Publication PA), for the recommendation of relevant scientific abstracts to researchers, based on their publication record to date. In two experiments, we analyzed the behavior of the Publication PA and found that it is a flexible and adaptable system, as well as an adaptive system. From Experiment 1 we concluded that users of the Publication PA can adapt the nature of the recommendations to their own personal wishes, using only one parameter. In a final version of the interface, this parameter could be controlled by a slider bar. Some researchers might be only interested in their current topic, for instance because they have just switched research topics. They can choose a low value for this *h* parameter. Researchers that would rather want to follow what is being published in research fields they previously published in may choose a high value of the *h* parameter.

Experiment 2 demonstrated that the Publication PA can select relevant papers for individual researchers. Papers that were recommended by the system were rated higher by the participants than papers that were not recommended.

The techniques applied in the Publication PA might also be applied to develop recommender systems in other domains in which personalized information retrieval is desirable. The domain should be primarily characterized by textual information sources, such as conference or journal papers, and the users should also be characterized by textual testimonials of their interests. Two examples of the wider applicability of the method of information selection that we proposed here are the problem of assigning manuscripts submitted to a conference to reviewers, and the problem of selecting relevant press bulletins from the stream of bulletins provided by press agencies world wide. We will discuss both these examples and hint at an implementation of our technique.

The assignment of manuscripts submitted to a conference to reviewers is a problem very similar to the selection of relevant abstracts for a reviewer. Even though reviewers can often indicate their areas of expertise, it is hard for conference program chairs to match every submission to the most qualified reviewers. Since the area of expertise of a reviewer is reflected in his or her publication record, user profiles that reflect the areas of expertise could be generated based on the publication record. By matching the profiles against each submitted abstract, the best-suited reviewer for each abstract will be associated with the highest relevance score. This way, conference chairs can easily assign submitted manuscripts to reviewers without having to rely on the reviewer's own opinion of his or her expertise, or without having to burden them with long questionnaires about their fields of research.

Press agencies produce many bulletins a day, often over 12.500 bulletins a year.[17] A reporter trying to read the most important press bulletins for his or her interests has to make a selection from this vast amount of information. Although press agencies often tag their bulletins or assign them to a certain category, it is easy to miss the one that is important. By creating profiles of reporters based on the news articles they have written over the years, an application similar to the Publication PA could make a meaningful selection for them.

17. Source: ANP Press support

A cognitive model of information selection can thus guide the development of a recommender system, because it provides insights in which features from the pieces of information are relevant for the selection process. The analysis suggests that the features that people that are engaged in retrieving relevant information use are the history of usage of words, and the co-occurrence of words. By incorporating these features in the same way as a cognitive model of human memory does, we have created a successful Publication PA, that for example can decrease the work load of individual researchers attending a conference by creating a preselection in the conference proceedings.

A Comparison between Decision Making and Memory Models for Literature Selection

This chapter is an extended version of Van Maanen, L. & Marewski, J.N. (2009). Recommender Systems for Literature Selection: A Competition between Decision Making and Memory Models. In N. A. Taatgen & H. Van Rijn (Eds.), Proceedings of the 31st Annual Meeting of the Cognitive Science Society.

LITERATURE SELECTION

In 2006, the number of scientific publications in the relatively small ISI subject category *Information Science & Library Science* was 2054. In other words, researchers working in this area had to scan through over 2,000 papers a year to keep up with the current developments. However, this number is, if anything, an underestimation of the total number of potentially relevant papers, as this number only holds if a researcher is interested in a single subject area. In practice, most researchers work on the intersection of multiple domains, increasing the number of potentially relevant papers enormously. Not only professionals in the scientific domain are confronted with masses of potentially relevant information. Also, government or business employees often need to decide which of numerous reports, leaflets, and bulletins to read, and which to ignore - a challenge that is aggravated by the continuously increasing amount of information that is available online. For instance, many press agencies produce over 12.500 bulletins a year. Reporters trying to read the most important ones have to make selections, and although the agencies often tag their bulletins, the sheer mass of information makes that it is easy to miss important ones.

In this paper, we will focus on one solution to this problem: recommender systems. Typically, corresponding decision aids automatically come up with a pre-selection of information that is worth further consideration, saving institutions, firms, and people parts of the time and effort otherwise required to separate the relevant from the irrelevant. In particular, here we will evaluate six models that can solve the problem of information selection for the scientific domain.

All models select relevant scientific papers from a large collection of scientific abstracts. They include (I) the *Publication Assistant* (Van Maanen et al., in press), a recommender system that was recently developed to assist scientists in identifying relevant articles. We will compare the performance of this system to that of (II-IV) three simple decision heuristics, including a unit-weight linear model (see Dawes, 1979; Dawes & Corrigan, 1974; Gigerenzer & Goldstein, 1996), and two lexicographic rules, called *take-the-best* (Gigerenzer & Goldstein, 1996), and *naiveLex*. We will also pit all models against (V-VI) two more complex linear weighted additive models, one being *Franklin's rule* (Gigerenzer & Goldstein, 1999) and the other multiple regression (Slovic & Lichtenstein, 1971).

While we do not aim to model the cognitive processes that are actually going on when scientists make literature choices, except for multiple regression all models tested here are grounded in cognitive theories. The Publication Assistant is a memory model that is based on the *rational analysis* framework (Anderson, 1990; Oaksford & Chater, 1998), as incorporated in the ACT-R *cognitive architecture* (Anderson, 2007a). The heuristics are models of decision making that are grounded in the *fast and frugal heuristics framework* (Gigerenzer, Todd, &

the ABC Group, 1999). The linear weighted additive model, Franklin's rule, is also a model of decision making (Gigerenzer & Goldstein, 1999). All models are common in the memory and judgment and/or decision making literature.

In what follows, we will give an outline of the Publication Assistant. Next, we will introduce the five alternative models. In an experiment, we will then evaluate the models' performance in predicting scientists' literature preferences.

THE PUBLICATION ASSISTANT: A MEMORY MODEL

An example of the problem addressed in this paper is the selection of relevant talks when attending a large, multi-track scientific conference such as the Annual Cognitive Science Conference. The information selection process starts when a researcher registers and receives a copy of the conference program. For instance, a strategy often employed by many conference attendees is to scan talk titles, author names, or abstracts for words or names that sound familiar. If an entry contains enough interesting words, it is selected for more careful reading, and the corresponding talk might be attended. In order to determine if a word qualifies as interesting in the context of the conference, a researcher might assess whether she has used the word in her own research in the past. The assumption is that the words used by someone in the context of their own research reflect their scientific interests. The Publication Assistant is a literature selection tool that could be run over a (digitized) conference program prior to attending the conference. The model recommends talks a given scientist might find useful to attend, saving that researcher the time and effort required to scan the conference program on his own. To this end, the model searches through the scientist's own work, examining in how far words that appear in conference abstracts also occur in the scientist's work. Specifically, the model bases its recommendations on the following properties an abstract's words:

Recency of occurrence in the scientist's own work

- The year in which a word from a conference abstract appears for the *first* time in the abstracts the scientist has published in the past,
- The year in which a word from a conference abstract appears for the *last* time in the abstracts the scientist has published in the past,

Frequency of occurrence in the scientist's own work

- The frequency of appearance of a word from a conference abstract in the abstracts the scientist has published in the past,
- The frequency of co-occurrence of a word from the conference abstract with another word in the abstracts the scientist has published in the past.

Based on these properties, the model creates an individual representation of a researcher's interests. The Publication Assistant applies these *user models* to predict the relevance of words that occur in other scientific abstracts, essentially estimating how familiar the contents of these abstracts would be to the scientist. In the next section, we will describe in more detail how the Publication Assistant estimates familiarity.

MODEL EQUATIONS

The Publication Assistant works like a model of the contents of a researcher's memory. Its equations are based on Anderson and Schooler's (1991) rational analysis of memory. According to their analysis, the probability that a *fact* (e.g., a word) stored in memory will be needed to achieve a processing goal can be predicted from the organism's pattern of prior exposure to the corresponding piece of information. For example, the probability that a fact about a scientific topic is of relevance to a researcher may depend on the frequency and recency of his writings about it in the past. Frequency and recency, in turn, feed into a memory currency called *base-level activation*, which influences a researcher's familiarity with the fact. These relations are captured by Equation 8.1, in which *B* stands for the base-level activation of a fact *i*, t_i stands for the time that has passed since the last exposure to that fact, and *d* represents the speed with which the influence of each exposure decays away. The summation takes place over all *n* previous encounters with the fact.

$$B = \ln\left(\sum_{i=1}^{n} t_i^{-d}\right) \quad \text{(equation 8.1)}$$

Besides frequency and recency of encounters with facts, the context in which these facts occur also plays a role in the activation of the facts. This *spreading activation* (Quillian, 1968) component represents the likelihood that a fact will be needed if another one is currently being used. These likelihoods depend on the pattern of prior exposures with the facts, as represented by the relatedness measure R_{ji} between two facts *j* and *i* (Anderson & Lebiere, 1998; Anderson & Milson, 1989):

$$R_{ji} = \frac{F(W_j \,\&\, W_i)F(N)}{F(W_j)F(W_i)} \quad \text{(equation 8.2)}$$

where $F(W_j)$ and $F(W_i)$ are the respective frequencies of facts *j* and *i*, $F(N)$ is the total number of exposures, and $F(W_j \,\&\, W_i)$ is the number of co-occurrences of the facts *j* and *i*.

With the equations that are provided by the rational analysis of memory, one can calculate the base-level activation of a word based on its occurrences in publications of the user. However, rather than using Equation 8.1 directly, the Publication Assistant uses Petrov's (2006) version of it. In Equation 8.3, the decay parameter is fixed at .5 and a history factor *h* is added, which represents a free parameter:

$$B = \ln\left(\frac{1}{\sqrt{t_1 + h}} + \frac{2n-2}{\sqrt{t_n} + \sqrt{t_1 + h}}\right) \text{with } h > 0 \quad \text{(equation 8.3)}$$

To stick to the example of selecting abstracts from a conference program, the Publication Assistant makes recommendations by combining the base-level activation of a word (*i*) with the weighted base-level activation of related words (*j*) in the abstract (Pirolli & Card, 1999):

$$A_i = B_i + \sum_j B_j R_{ji} \quad \text{(equation 8.4)}$$

To compare the relevance of abstracts with each other, each one is represented by the average activation of the words that occur in it. In a comparison of two abstracts, the Publication Assistant then recommends the more activated one. Abstracts in which many words have high base-level and spreading activation values have a high match with the researchers own word usage, and thus may be more interesting.[18] The Publication Assistant's recommendations are thus based on the structure of the environment of a particular researcher. In particular, the structure of word usage in previously published abstracts. The only parameter that may be varied is the history parameter h, which represents the relative importance of recency versus frequency in determining activation. In the research reported here, we kept *h* constant at the same value reported in Van Maanen et al. (in press).

18. Van Maanen **et al.** (in press) found that the frequency of words in scientific abstracts differs from normal word usage in written English. To counter the unwanted influence of normally high-frequent words (e.g, "the"), van Maanen **et al.** built a filter for these words when they developed the Publication Assistant. Here, we run all analyses using that filter. As they showed, the filtering does not interfere with how well an abstract represents the contents of a paper.

ALTERNATIVE MODELS: DECISION STRATEGIES

To evaluate the performance of the Publication Assistant in predicting scientists' literature preferences, we compared it to five alternative models. While the Publication Assistant essentially mimics a model of memory, these alternative models have originally been proposed as decision strategies in the judgment and decision making literature.

In particular, we focus on a class of models that have been prominent in the fast and frugal heuristics framework. According to this framework, humans (and other organisms) often make decisions under the constraints of limited information processing capacity, knowledge, and time - be they about the relevance of scientific articles, or the likely performance of stocks, or the nutritional value of food. Such decisions can nevertheless be made successfully because humans can rely on a large repertoire of simple decision strategies, called heuristics. These rules of thumb can perform well even under the above-mentioned constraints. They do so by exploiting the structure of information in the environment in which a decision maker acts and by building on the ways evolved cognitive capacities work, such as the speed with which the human memory system retrieves information.

One of the heuristics tested here, the unit-weight linear model, is particularly simple, requiring no free parameters to be fitted. Related models have proved to be almost as successful (or even better) in predicting unknown events and quantities as multiple regressions (see Dawes & Corrigan, 1974; Dawes, 1979). Just as the unit-weight linear model, also naiveLex dispenses with all free parameters. If these two particularly simple heuristics predicted scientist's literature preferences successfully, then they would simplify the selection of abstracts more than the Publication Assistant does. In order to be considered a useful tool, the Publication Assistant should thus be able to outperform these models. Take-the-best is a little more complex, requiring one free parameter to be fitted for each individual scientist. Take-the-best and related models have been found to be, on average, more accurate than multiple regression in predicting various economic, demographic, and environmental, variables (e.g., Czerlinski, Gigerenzer, & Goldstein, 1999). Finally, the most complex models tested here, Franklin's rule and multiple regression, require for each individual researcher as many free parameters to be fitted as there are words in the abstracts under consideration. While these two models are prominent in the judgment and decision making literature, due to their large complexity they are not considered heuristic decision strategies in the fast and frugal heuristics framework. Rather, they are often used as benchmark to evaluate the performance of heuristics in model comparisons (Czerlinski, Gigerenzer, & Goldstein, 1999; Gigerenzer & Goldstein, 1996).

LEXICOGRAPHIC HEURISTICS: TAKE-THE-BEST, NAIVELEX

The first model to be considered here is take-the-best. To make literature recommendations, take-the-best uses attributes of articles as *cues*. In our context, cues are the words that occur in an abstract. If such a word also occurs in a scientist's own publication, then it is considered a *positive* cue, suggesting that an abstract is of interest to that scientist. Take-the-best considers all cues sequentially (i.e., one at a time; hence lexicographic) in the order of their *validity*. The validity of a cue is the probability that an alternative A (e.g., an article) has a higher value on a criterion (e.g., relevance for a researcher) than alternative B, given that alternative A has a positive value on that cue and alternative B does not. In a comparison of two abstracts, take-the-best bases a decision on the first cue that *discriminates* between the abstracts, that is, on the first cue for which one abstract has a positive value and the other one does not. The heuristic is defined in terms of three rules:

(1) Look up cues in the order of their validity.

(2) Stop when the first cue is found that discriminates between the abstracts.

(3) Choose the abstract that this cue favors.

The second lexicographic model, here called naiveLex, is identical to take-the-best, except that it does not estimate the validity order of cues. Rather, cues are simply considered in the order of the frequency of occurrence of the corresponding words in each researcher's published abstracts. This aspect of the model is similar to the Publication Assistant, in which the word frequency is also taken into account (but weighted with recency).

A UNIT-WEIGHT-LINEAR HEURISTIC

Lexicographic heuristics such as take-the-best can avoid going through all words (i.e., cues) from an abstract, which can save effort, time, and computations once the order of cues is known. Unit-weight linear heuristics, in contrast, integrate all cues into a judgment by adding them. These models can nevertheless simplify the task by weighing each cue equally (hence unit-weight). In a comparison of two abstracts, it reads as follows:

(1) For each abstract, compute the sum of positive cues.

(2) Decide for the abstract that is favored by a larger sum.

WEIGHTED-ADDITIVE MODELS: FRANKLIN'S RULE AND MULTIPLE REGRESSION

Franklin's rule (Gigerenzer & Goldstein, 1999) is similar to the unit-weight linear heuristic, but instead weights all the cues by their validities prior to summation. (The cue validities are identical to those relied on by take-the-best.) Multiple regression, in turn, estimates the weights of the cues by minimizing the error in the calibration set using maximum likelihood estimation. In a comparison of two abstracts, Franklin's rule and multiple regression read as follows:

(1) For each abstract, compute the weighted sum of cues.

(2) Decide for the abstract that is favored by a larger sum.

EXPERIMENT

To compare the Publication Assistant to the alternative models' capability of predicting actual scientist's literature preferences, we re-analyzed data from a study by Van Maanen et al. (in press, Experiment 2). They had asked researchers from the field of cognitive science to rate how much they were interested in a paper after reading the abstract. In this study, Van Maanen et al. had found that the Publication Assistant could fit researcher's interests reasonably well; however, they did not compare its performance to that of alternative models.

METHODS

Participants

Ten researchers (2 full professors, 2 associate professors, 5 assistant professors, and 1 post-doc) from various subfields of cognitive science and from various countries were asked to participate.

Procedure

For each of the researchers, Van Maanen et al. (in press) constructed user models of the Publication Assistant based on the abstracts of their published work insofar it was indexed by PsycINFO. They then ordered all abstracts from the last three volumes (2004-2006) of the

Cognitive Science Journal according to the predicted relevance for the researcher, based on the researcher's published abstracts.

From the ordered list of abstracts, they presented each researcher the top five abstracts, the bottom five abstracts, and five abstracts from the middle of the list. For each researcher, the presentation order of these 15 abstracts was randomized. Each researcher indicated with a grade between 0 and 9 how much he or she was interested in the papers, based on the abstracts.

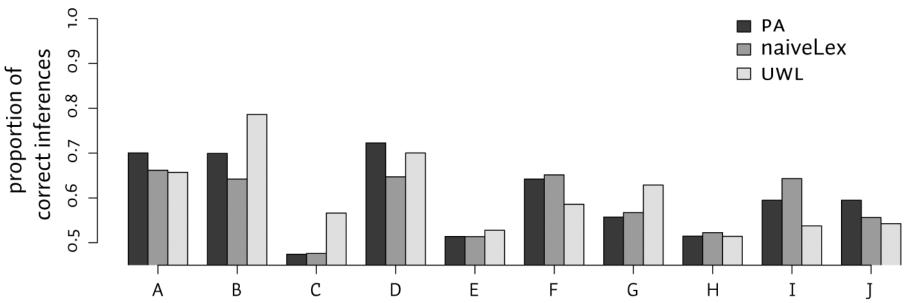


Figure 8.1 The performance of the non-calibrated models. A-J represent individual participants. PA: Publication Assistant; UWL: unit-weight linear heuristic.

Analyses

To compare the performance of the Publication Assistant to that of the five alternative models in predicting each researcher's ratings, we ran a cross-validation. To this end, we constructed paired comparisons of all 15 abstracts for each participant individually (210 pairs). We divided each participant's abstracts pairs randomly into two parts. The first part represented the *calibration set* in which we calculated for each participant that person's optimal values for the free parameters in take-the-best, Franklin's rule, and multiple regression, respectively. That is, we identified the parameter value at which each model would correctly predict the largest proportion of literature preferences. Take-the-best, Franklin's rule, and multiple regression will therefore be referred to as the calibrated models.

We used these optimal values to compute the proportion of preferences consistent with each model in the other half, the *validation set*, where the models' generalizability is evaluated. For each partition, we also computed the proportion of preferences consistent with the three not-calibrated models (the Publication Assistant, naiveLex, and the unit-weight linear heuristic. The free parameter of the Publication Assistant, *h*, we set to 10. In fitting the very same participants as we do here, van Maanen et al. (in press), had found this value to work reasonably well. The other two models were parameter free in this respect.

We ran these analyses for a subset of possible sizes of the calibration and validation sets; that is, we first computed the proportion of each model's correct predictions for a calibration set size of 1 and a test set size of 209, then for a calibration set size of 11, and a test set size of 199, and so on. The larger the size of the calibration sets, the larger is the sample of paired comparisons from which the parameterized decision models can estimate an individual researcher's interests, that is, the more "experience" these models can accumulate before making their predictions. This procedure was repeated enough times to average out noise due to the random selection of calibration sets.

RESULTS

When comparing the Publication Assistant with the other non-calibrated models (naiveLex and the unit-weight linear model), we found that the three models performed differently for different participants (Figure 8.1). The Publication Assistant made the most correct inferences for three participants (A, D, and J), while unit-weight linear heuristic scored outperformed the other two non-calibrated competitors on four occasions (B, C, E, and H). NaiveLex scored best for three participants (F, G, and I). Overall, the performance of the models did not differ much (mean_{PA}=0.60, $\text{mean}_{\text{naïeveLex}}$=0.59, mean_{UWL}=0.60).

For each of the 10 participants, Figure 8.2 shows the proportion of correctly predicted preferences for the three calibrated models as a function of the size of the calibration set. As one would expect, for all participants the accuracy of the predictions of the parameterized models increases with the size of the calibration set. Of the calibrated models, Franklin's rule was consequently outperformed by the take-the-best heuristic and the multiple regression model, which performed equally well, but differed among participants. Take-the-best was the best predictor for participants B, C, E, G, I, and J, while the regression model performed best for participants A, D, F, and H. Overall, take-the-best performed best (mean_{TTB}=0.84, mean_{MR}=0.81, $\text{mean}_{\text{Franklin}}$=0.71).

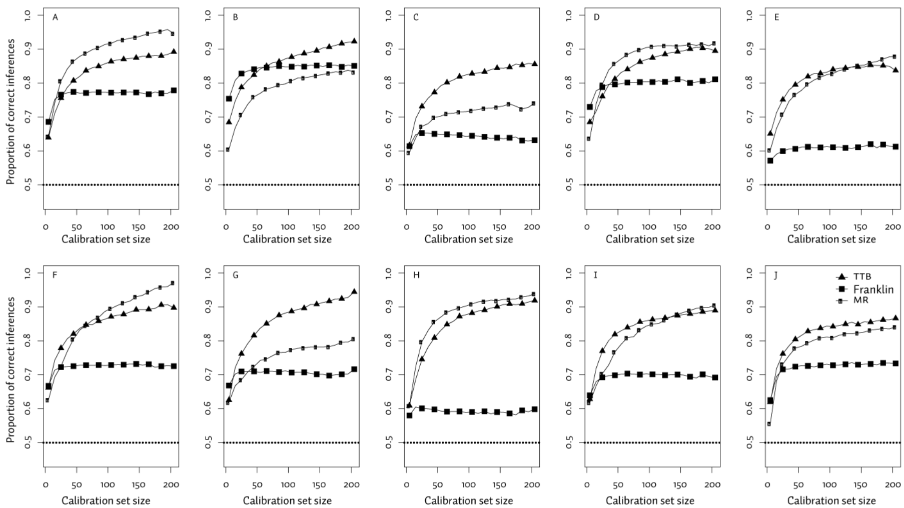


Figure 8.2. The calibrated models' individual predictions of literature selections. Each panel represents one participant. TTB: take-the-best; Franklin: Franklin's rule; MR: multiple regression.

DISCUSSION

We examined the ability of six models to predict scientists' literature preferences: (I) the Publication Assistant, a recommender system that is based on a rational analysis of memory and the ACT-R architecture; (II-IV) three simple heuristics, including take-the-best, a naive lexicographic model, and a unit-weight linear model, and (V-VI) two complex weighted-additive models, Franklin's rule and multiple regression.

For some participants and calibration set sizes, the regression model outperformed take-the-best. One reason why take-the-best did not fare as well as multiple regression on every occasion might be that the structure of information available in the abstracts was not well suited for this simple heuristic (Martignon & Hoffrage, 2002). For instance, take-the-best essentially bets on a noncompensatory information structure, always preferring the most valid discriminating cue to all others. In the domain of literature selection, such information structures might not be prevalent. To give an example, the words "Memory" and "Retrieval" might be equally good predictors of some cognitive scientist's research interests.

One result was that the performance of the non-calibrated models differed between participants. However, it should be realized that naiveLex and the Publication Assistant only differ with respect to the use of the recency component. Both models use the frequency of words in published abstracts in the same way. Therefore, the difference in the participants in which naiveLex is the better recommender may be attributed to the importance that these participants contribute to recency. That is, the Publication Assistant overestimates the importance of more frequent words in the published abstracts. Thus, recommendations of the Publication Assistant could improve if we would allow the *h* parameter to be fit individually.
The fact that the non-calibrated models performed differently between participants is in agreement with other findings in the judgment and decision making literature, where large individual variation in people's use of decision strategies are commonly observed (Bröder & Gaissmaier, 2007; Mata, Schooler, & Rieskamp, 2007; Pachur, Bröder, & Marewski, 2008). In this respect, it is somewhat surprising that take-the-best consistently outperforms all other competitors. Based on the literature, we would have expected that the most useful approach designing recommender systems would have been to build different systems for different users, depending on which model predicts the respective scientist's preferences best.

WHY WAS THE PUBLICATION ASSISTANT OUTPERFORMED BY THE CALIBRATED MODELS?

Take-the-best, Franklin's rule, and the regression model learned about the scientists' interests directly from the paired comparisons between abstracts that were included in the calibration sets. The Publication Assistant, in turn, was trained on a participant's published abstracts (Van Maanen et al. in press), under the assumption that word frequencies in those abstract would reflect the participants' interests. While this way of training the model better reflects real-life situations of information selection, in which people's appraisal of items (such as abstracts) is often unknown, it might have been detrimental for the model's performance.
In addition, our results complement findings by Lee, Loughlin, and Lundberg (2002), who, in a study on literature search, examined the performance of a simple heuristic in identifying articles that are relevant to a given topic of interest (e.g., eyewitness testimony). Their analyses show that a researcher going by a variant of take-the-best would have had to search through fewer articles in order to find the relevant ones than a person behaving in accordance with a weighted-additive model.

CONCLUSION

In this paper, we evaluated the ability of cognitive models of memory and decision making to serve as literature recommender systems. Three of the models were trained on the participants' published abstracts (the non-calibrated models), while three other models were allowed to train on a calibration set that contained abstracts that were also part of the paired comparisons in the validation set. This second type of training generally yields better results, but at a cost of a less realistic training situation. The non-calibrated models showed large individual variability, suggesting that for successful recommendation, the best predicting model should be determined first. Future work will show if this result is generalizable to other domains of literature recommendation and information search.

To conclude, in today's world of mass media, the choice which information to attend to, and which to ignore becomes an ever more important challenge for professionals. Automatic recommender system might help to cope with these demands of the information age - savings in time and effort that can eventually be invested elsewhere. We hope that comparisons between different approaches, such as the ones tested here, help along that way.

Conclusion

RESULTS FROM PART I

In Part I of this thesis, a new model for retrievals from declarative memory was proposed. The model, called Retrieval by Accumulating Evidence in an Architecture (RACE/A), accounts for semantic interference effects within the constraints of the cognitive architecture ACT-R (Anderson, 2007a). By adding a sequential sampling model of declarative memory as a component to ACT-R, it becomes possible to accurately predict human behavior in tasks that involve competitive processes in memory retrieval, such as the Stroop effect (Chapter 4), picture-word interference (Chapters 2, 3.2 and 4), lexical decision (Chapter 3.1), or subliminal perception (Chapter 3.3). In addition, RACE/A demonstrates the usefulness of the architectural approach towards modeling cognition. Without RACE/A, it would have been harder to model to explain some of the effects discussed in this thesis.

One example is the interaction between stimulus repetition and semantic interference (as discussed in Chapter 2), which could not be accounted for with a single sequential sampling model. RACE/A computes the long-term activation dynamics (estimated by the history of usage of the chunks) as well as the short-term activation dynamics (estimated by a sequential sampling process). The long-term activation dynamics effectively act as a starting point of the evidence accumulation process. Because the long-term activation dynamics is determined by the ACT-R architecture, the short-term activation dynamics start at a principled of starting point. In this respect, RACE/A differs notably from the standard sequential sampling approach in which the starting point of evidence accumulation is treated as a free parameter, which is often interpreted after model fitting has taken place.

Another example is the mechanism that we proposed that may explain the differences (and similiarities) between the Stroop task and the picture-word interference (PWI) task (Chapter 4). It has often been assumed that the effects observed in these tasks are caused by the same principle (MacLeod, 1991). A recent study (Dell'Acqua et al., 2007) questioned this long-standing position by presenting data from a dual-task experiment, in which the effects from a picture-word interference manipulation differed from those from a Stroop manipulation. In two experiments, we replicated and confirmed these observations, but also demonstrated that these observations could still be caused by the same underlying principle (Chapter 4). The theory put forward in this chapter is that interference in both Stroop tasks and PWI tasks is caused by competition among multiple potential memory representations. Because at multiple stages during the tasks memory retrievals take place, the interference effects can manifest at multiple stages as well. This theory could not have been modeled using traditional sequential sampling models, because the amount of competition at every stage critically depends on the duration of the stages. In Chapter 4, we showed that the amount of interference at every stage could be manipulated by adapting one single parameter, which we used to account for the difference between Stroop and PWI.

The contribution of RACE/A to the theory of declarative memory retrieval lies in the possibility to model the retrieval process on a very small time-scale. This proved to be especially useful when explaining the temporal dynamics of memory retrievals. For instance, asynchronously presented stimuli can influence the time course of memory retrievals (e.g., W. R. Glaser & Düngelhoff, 1984), which can be modeled using RACE/A (Chapter 2). Moreover, the integration of RACE/A in a cognitive architecture restricts the freedom a modeler has when trying to fit data sets. This is because RACE/A models – besides aligning with the specific theory behind RACE/A – also have to adhere to the ACT-R central assumptions (R. P. Cooper, 2007). In particular, besides the retrieval latency equation, RACE/A does not change any of the central architectural assumptions. With respect to the retrieval latency equation, we showed (Chapter 2) that in the absence of competition during memory retrieval, RACE/A makes similar predictions as the default ACT-R retrieval latency equation. Thus, previous ACT-R models of tasks in which interference does not play a critical role are not invalidated by our extension. The combination of sequential sampling and cognitive architecture only increases the set of phenomena that can be accounted for by the architecture, while at the same time remaining true to the architecture's central assumptions.

RESULTS FROM PART II

In Part I, we studied how the context in which a stimulus is presented influences the retrieval processes that are triggered by that stimulus. For this purpose, we studied small-scale behavioral effects and tried to explain them within the ACT-R architecture of cognition. In Part II, we have focused on the effects of a more general context on the retrieval process. With the use of cognitive models we have studied how individual contexts predict which concepts will be retrieved from memory. Based on this, we have developed information retrieval algorithms in two domains. The applications that have been developed for the domains of artwork selection and scientific literature search, the Virtual Museum Guide (VMG, Chapter 5) and the Personal Publication Assistant (Publication PA, Chapter 7) incorporate a model of the user's memory structure, and use that to predict user preferences. That is, based on an analysis of the statistical properties of the textual environment that a user interacts with, a prediction is made about which aspects of this textual environment are more likely to be recalled in the near future. These aspects are hypothesized to be the most relevant facts for that user given the current environment and time. The Publication PA used this prediction to select scientific papers for individual researchers. The VMG used this prediction to select a sequence of artworks for individual (online) museum visitors.

In Chapter 5, we developed a software environment in which users can interact with the collection of the Amsterdam Rijksmuseum. An online version of the system was available for four months in 2007, allowing users to study the Rijksmuseum's collection. The system can be used to test the applicability of cognitive theories for user modeling in the cultural heritage domain. Particularly, we tested an activation-based user model for the cultural heritage domain (the VMG) that is capable of infering visitors' interests by incorporating a model of a museum guide's memory.

The aim of this user model was to present an interesting tour to visitors, given the VMG's extensive knowledge on the museum's collection, and the perceived interests from the visitors. However, the study revealed that participants were just as satisfied with an online museum guide that incorporated the participants' feedback as with a simpler algorithm that did not take their feedback into account. One reason why the VMG did not perform better might be related to the way the participants could provide feedback during the training phase. The participants could only indicate with a button press that they liked or disliked the artwork. That is, there was no option to indicate why an artwork was liked or disliked, or which aspect of an artwork led to the decision between positive or negative feedback. In Chapter 6, we studied - in a slightly different setting - if a feedback mechanism in which point of gaze was used would lead to more appreciation of personalized information presentation. Although we did find hints that interest and gaze are related, we did not find a sufficient strong effect that could be used to include a gaze-based feedback mechanism in a artwork recommender system.

Another reason why it was difficult to provide good recommendations with the VMG relates to the representation of interest in the items that are being selected. One important aspect of the cultural-heritage domain is that interesting aspects of art are not necessarily text-based. Although in the context of an art recommender for the cultural heritage domain descriptions of the cultural-historic value of an artwork may contribute a great deal to the interest a user has in that particular artwork, it may not be the only aspect. The use of certain colors, painting or crafting techniques, or a particular arrangement of figures in the scene may also contribute to the appeal a work of art may have to a user. These aspects are often not included in the descriptions of the artworks that we used in developing the VMG. In addition, these relate to the expressiveness of an artwork (Arnheim, 1954/1974), which is not easily described.

Thus, the nature of art makes it hard to determine user preferences based on cultural-historic descriptions of the art only. On the other hand, in scientific literature (Chapter 7) the interesting aspects of a paper or abstract are names, concepts, and keywords, all of which are inherently textual. Therefore, representing the interests of users of a scientific paper recommender as a network of textual features is natural. In addition, the mapping between the textual properties of the training set (in this case, the already published papers of an author) and the user's interests is more straightforward than in the cultural heritage domain. Since the main output channel for science is written text, the expression of interest in words is natural, contrary to the cultural heritage domain. Consequently, the ramp-up problem for users (Konstan, Riedl, Borchers, & Herlocker, 1998) applies to a lesser extent, since the user's preferences are learned faster. Informal analyses suggested that about five or six recent publications are enough to capture a scientist's current interest.

Another reason for the difference between the cultural heritage and scientific literature domains relates to the need to explain the reasons behind certain recommendations. A study by Cramer et al. (2008) suggest that users of art recommenders prefer explanations of *why* a certain recommendation is given over unexplained recommendations. The issue of providing reasons for recommendations plays a less prominent role in the selection of relevant scientific publications. The selection of publications does not involve a sequence of selections, but rather one isolated selection. Therefore, providing reasons for successive recommendations is not an issue. By contrast, the selection of artworks for the purpose of an automated museum tour is by definition a sequential process, because a tour necessarily consists of multiple artworks.

In the domain of scientific paper selection, we demonstrated that our approach to predicting interests works reasonably well (Chapter 7). A selection of articles that the Publication PA considered relevant for a certain researcher was rated higher on a relevance scale by that researcher than a selection of articles that the Publication PA considered irrelevant. Also, in a competition between various selection algorithms (Chapter 8), the Publication PA performed equally well as the other competitors that were were trained on the same calibration set as the Publication PA (the profiles of the users). For some users, the Publication PA outperformed the other competitors, while for other users, the Publication PA is outperformed. The competitors that were cross-validated on a subset of the possible paired comparisons of the abstracts in the data set performed better than the Publication PA. However, this only shows that if more and more suitable training data is available, performance goes up.

CONCLUSION

This thesis discussed new formal models of memory retrieval, and at the same time discusses how these kinds of models can be deployed for information selection problems. We showed that activation-based theories of declarative memory retrieval, such as a Rational Analysis of Memory can be used in information retrieval applications. This result depends on the similarities between declarative memory retrieval and information retrieval, which both can be characterized by the history of usage of certain items or relevant words. This result complements the results from Part I, in which we presented a new, activation-based, theory of declarative memory retrieval. The theory accounts for a wide range of interference related phenomena that relate to the history of usage of certain items, and the context in which they are retrieved.

The complementing findings from Part I and Part II show that theoretical cognitive science (cognitive modeling) and computer science (cognitive engineering) may be a fruitful combination in which theoretical development as well as application-based research may go hand in hand.

References

Altmann, E. M., & Davidson, D. J. (2001). An integrative approach to Stroop: Combining a language model and a unified cognitive theory, *Proceedings of the 23rd annual meeting of the Cognitive Science Society*: Hillsdale, 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. (1989). A rational analysis of human memory. In H. Roediger & F. Craik (Eds.), *Varieties of memory and consciousness* (pp. 195-210). Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (2007a). *How can the human mind occur in the physical universe?* New York: Oxford UP.

Anderson, J. R. (2007b). Using brain imaging to guide the development of a cognitive architecture. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 49-62). New York, NY: Oxford UP.

Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing problem solving in real time: fMRI analysis of the subject-paced tower of hanoi. *Journal of Cognitive Neuroscience, 17*(8), 1261-1274.

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., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*(4), 341-380.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free-recall. *Psychological Review, 79*(2), 97-123.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington: Winston and Sons.

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: Lawrence Erlbaum.

Anderson, J. R., Matessa, M., & Lebiere, C. (1998). The visual interface. In J. R. Anderson & C. Lebiere (Eds.), *The atomic components of thought*. Mahwah, NJ: Lawrence 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.

Antes, J. R. (1974). Time course of picture viewing. *Journal of Experimental Psychology, 103*(1), 62-70.

Arnheim, R. (1954/1974). *Art and visual perception: A psychology of the creative eye*. Berkeley and Los Angeles: University of California Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects fo subjects and items. *Journal of Memory and Language*.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database (cd-rom).

Babcock, J. S., & Pelz, J. B. (2004). Building a lightweight eyetracking headgear, *Proceedings of the 2006 Symposium on Eye tracking research and applications* (pp. 109-114). San Antonio, TX: ACM.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology-Learning Memory and Cognition, 14*(4), 579-589.

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.

Beer, V. (1987). Great expectations: Do museums know what visitors are doing? *Curator, 30*(3), 206-215.

Bell, G. (2002). *Making sense of museums: The museum as 'cultural ecology'*: Intel Labs.

Bocconi, S. (2006). *Vox populi: Generating video documentaries from semantically annotated media repositories.* Unpublished PhD Thesis, Technical university Eindhoven, Eindhoven.

Boening, G., Bartl, K., Dera, T., Bardins, S., Schneider, E., & Brandt, T. (2006). Mobile eye tracking as a basis for real-time control of a gaze driven head-munted video camera, *Proceedings of the 2006 Symposium on Eye tracking research and applications* (pp. 56). San Diego, CA: ACM.

Borst, J. P., & Van Rijn, H. (2006). Memory decay problems in a level-repetition switch task model. In D. Fum, F. Del Missier & A. Stocco (Eds.), *Proceedings of the International Conference on Cognitive Modeling* (pp. 351-352). Trieste, Italy: Edizioni Goliardiche.

Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattributte decisions. *Psychonomic Bulletin & Review, 14*(5), 895-900.

Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review, 112*(1), 117-128.

Brown, S., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153-178.

Brusilovsky, P., & Tasso, C. (2004). Special issue on user modeling for web information retrieval preface. *User Modeling and User-Adapted Interaction, 14*(2-3), 147-157.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience, 21*, 47-74.

Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences, 3*(2), 57-65.

Chen, X., & Zelinsky, G. J. (2006). Real-world visual search is dominated by top-down guidance. *Vision Research, 46*(24), 4118-4133.

Claypool, M., Brown, D., Le, P., & Waseda, M. (2001). Inferring user interest. *IEEE Internet Computing, 5*(6), 32-39.

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.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.

Cone, C., & Kendall, K. (1978). Space, time, and family interactions: Visitor behavior at the science museum of Minnesota. *Curator, 21*(3), 245-258.

Cooper, R. M. (1974). Control of eye fixation by meaning of spoken language: New methodology for real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*(1), 84-107.

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.

Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., et al. (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction, 18*(5), 455-496.

Crofts, R., Doerr, M., & Gill, T. (2003). The CIDOC conceptual reference model: A standard for communicating cultural contents. *Cultivate Interactive, 9*.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd & the ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 97-118). New York: Oxford UP.

Dawes, R. M. (1979). Robust beauty of improper linear-models in decision-making. *American Psychologist, 34*(7), 571-582.

Dawes, R. M., & Corrigan, B. (1974). Linear-models in decision-making. *Psychological Bulletin, 81*(2), 95-106.

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition, 31*(3), 353-359.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391-407.

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 & Computers, 32*(4), 588-615.

Denckla, M. B., & Rudel, R. G. (1976). Rapid automatized naming (ran) - dyslexia differentiated from other learning-disabilities. *Neuropsychologia, 14*(4), 471-479.

Donders, F. C. (1868/1969). On the speed of mental processes. In *Attention and performance* (Vol. 2, pp. 412-431). Amsterdam.

Duchowski, A. (2002). A breadth-first survey of eye tracking applications. *Behavior Research Methods Instruments & Computers, 34*(4), 455-470.

Dumais, S. T. (2003). Data-driven approaches to information access. *Cognitive Science, 27*(3), 491-524.

Dumais, S. T., & Nielsen, J. (1992). Automating the assignment of submitted manuscripts to reviewers, *15th Annual International SIGIR*. Denmark: ACM.

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E. J. (submitted). A diffusion model decomposition of the practice effect.

Dyer, F. N. (1973). The Stroop phenomenon and its use in study of perceptual, cognitive, and response processes. *Memory & Cognition, 1*(2), 106-120.

Ehri, L. C. (1976). Do words really interfere in naming pictures. *Child Development, 47*(2), 502-505.

Fagot, C., & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1058-1079.

Falkovych, K., Cena, F., & Nack, F. (2006). Combining coherence and adaptation in discourse-oriented hypermedia generation, *Fourth International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH2006)* (pp. 274-278). Dublin.

Farahat, A., Pirolli, P., & Markova, P. (2004). *Incremental methods for computing word pair similarity* (technical report No. TR-04-6). Palo Alto, CA: Palo Alto Research Center.

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1187-1199.

Findlay, J. M., & Gilchrist, I. D. (2005). Eye guidance and visual search. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (2nd ed., pp. 259-282). New York: Oxford UP.

Fink, J., & Kobsa, A. (2002). User modeling for personalized city tours. *Artificial Intelligence Review, 18*(1), 33-74.

Foltz, P. W., & Dumais, S. T. (1992). Personalized information delivery - an analysis of information filtering methods. *Communications of the ACM, 35*(12), 51-60.

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., & Saenz, S. P. (2007). Repetition priming endurance in picture naming and translation: Contributions of component processes. *Memory & Cognition, 35*(3), 481-493.

Fu, W. T., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the world wide web. *Human-Computer Interaction, 22*(4), 355-412.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650-669.

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In G. Gigerenzer, P. M. Todd & the ABC Group (Eds.), *Simple heuristics that make us smart*. New York: Oxford UP.

Gigerenzer, G., Todd, P. M., & the ABC Group. (1999). *Simple heuristics that make us smart*. New York: Oxford UP.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 18*(4), 381-398.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 875-894.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: 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.

Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning. *Machine Learning, 3*(2), 95-99.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518-565.

Gray, W. D. (2007a). Composition and control of integrated cognitive systems. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 3-12). New York: Oxford UP.

Gray, W. D. (2007b). *Integrated models of cognitive systems*. New York: Oxford UP.

Greenwald, A. G., & Shulman, H. G. (1973). On doing two things at once II: Elimination of psychological refractory period. *Journal of Experimental Psychology, 101*(1), 70-76.

Griffiths, T. L., Steyvers, M., & Firl, A. (2007). Google and the mind: Predicting fluency with PageRank. *Psychological Science, 18*(12), 1069-1076.

Grootjen, M., Neerincx, M. A., & Veltman, J. A. (2006). Cognitive task load in a naval ship control centre: From identification to prediction. *Ergonomics, 49*(12-13), 1238-1264.

Hansen, J. P., Hansen, D. W., & Johansen, A. S. (2001). Bringing gaze-based interaction back to basics, *Proceedings of Universal Access in Human-Computer Interaction (UAHCI 2001)*. New Orleans, LA.

Hatala, M., & Wakkary, R. (2005). Ontology-based user modeling in an augmented audio reality system for museums. *User Modeling and User-Adapted Interaction, 15*(3-4), 339-380.

Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance, 28*(3), 527-545.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response-time distributions: An example using the Stroop task. *Psychological Bulletin, 109*(2), 340-347.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498-504.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243-271.

Hendriks, P., Van Rijn, H., & Valkenier, B. (2007). Learning to reason about speaker's alternatives in sentence comprehension: A computational account. *Lingua, 117*(11), 1879-1896.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*(1), 11-26.

Hornof, A. J., & Cavender, A. (2005). Eyedraw: Enabling children with severe motor impairments to draw with their eyes, ACM *CHI: Conference on Human Factors in Computer Systems*: ACM Press.

Hsi, S., & Fait, H. (2005). Rfid enhances visitors' museum experience at the exploratorium. *Communications of the ACM, 48*(9), 60-65.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), 23-32.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194-203.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(11), 1254-1259.

Jacob, R. J. K. (1991). The use of eye-movements in human-computer interaction techniques - what you look at is what you get. ACM *Transactions on Information Systems, 9*(2), 152-169.

Janssen, C. P., Van Rijn, H., Van Liempd, G., & Van der Pompe, G. (2007). User modeling for training recommendation in a depression prevention game. In T. Goosen & J. Janssen (Eds.), *1st NSVKI Student Conference* (pp. 29-35). Nijmegen: Utrecht University Press.

Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, (1), 249-256.

Jolicoeur, P., & Dell' Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*(2), 138-202.

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.

Juvina, I., & Taatgen, N. A. (2009). Adding distractors improves performance by boosting top-down control. In N. A. Taatgen & H. Van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society*. Amsterdam, NED.

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. C. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science, 282*(5386), 108-111.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology, 77*(4), 576-588.

Koch, C., & Ullman, S. (1985). Shifts in selective visual-attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219-227.

Kohonen, T. (2001). *Self-organizing maps*. Berlin: Springer.

Konstan, J., Riedl, J., Borchers, A., & Herlocker, J. L. (1998). Recommender systems: A GroupLens perspective. In *Recommender systems: Papers from the 1998 workshop (AAAI technical report ws-98-08)* (pp. 60-64). Menlo Park, CA: AAAI Press.

Kootstra, G., Nederveen, A., & De Boer, B. (2008). Paying attention to symmetry, *Proceedings of the British Machine Vision Conference (BMVC2008)* (pp. 1115-1125). Leeds, UK.

Kowler, E., & Blaser, E. (1995). The accuracy and precision of saccades to small and large targets. *Vision Research, 35*(12), 1741-1754.

Kuipers, J. R., La Heij, W., & Costa, A. (2006). A further look at semantic context effects in language production: The role of response congruency. *Language and Cognitive Processes, 21*(7-8), 892-919.

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., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33*(1), 1-64.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211-240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284.

Lebiere, C., & West, R. L. (1999). A dynamic ACT-R model of simple games, *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 296-301): Erlbaum.

Lee, M. D., Loughlin, N., & Lundberg, I. B. (2002). Applying one reason decision-making: The prioritisation of literature searches. *Australian Journal of Psychology, 54*(3), 137-143.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*: 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.

Levy, J., & Pashler, H. (2001). Is dual-task slowing instruction dependent? *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 862-869.

Li, D., Babcock, J. S., & Parkhurst, D. J. (2006). openEyes: A low-cost head-mounted eye-tracking solution, *Proceedings of the 2006 Symposium on Eye tracking research and applications* (pp. 95-100). San Diego, CA: ACM.

Lindsey, R., Veksler, V. D., Grintsvayg, A., & Gray, W. D. (2007). Be wary of what your computer reads: The effects of corpus selection on measuring semantic relatedness, *8th International Conference on Cognitive Modelling*. Ann Arber, MI.

Livesay, K., & Burgess, C. (1998). Mediated priming in high-dimensional semantic space: No effect of direct semantic relationships or co-occurrence. *Brain and Cognition, 37*(1), 102-105.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 565-572.

Lovett, M. C. (2002). Modeling selective attention: Not just another model of Stroop (NJAMOS). *Cognitive Systems Research, 3*(1), 67-76.

Lovett, M. C. (2005). A strategy-based interpretation of Stroop. *Cognitive Science, 29*(3), 493-524.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103-189). New York: Wiley.

Luce, R. D. (1986). *Response times*. New York: Oxford UP.

Lupker, S. J. (1979). Semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485-495.

Mackworth, N. H., & Morandi, a. J. (1967). Gaze selects informative details within pictures. *Perception & Psychophysics, 2*(11), 547-547.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203.

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.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4*(10), 383-391.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*(2), 197-237.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision, 52*(1), 29-71.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of strategies. *Psychology and Aging, 22*, 796-810.

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.

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence.

McClelland, J. L. (1979). On the time relations of mental processes: Examination of systems of processes in cascade. *Psychological Review, 86*(4), 287-330.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition, 79*(1-2), 115-134.

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.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Lawrence Erlbaum.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283-308). New York: Academic Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard UP.

Niessen, M. E., Van Maanen, L., & Andringa, T. C. (2008). Disambiguating sounds through context. *International Journal of Semantic Computing, 2*(3), 327-341.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General, 137*(3), 434-455.

Oaksford, M., & Chater, N. (1998). *Rational models of cognition*. Oxford: Oxford UP.

Oyekoya, O. W., & Stentiford, F. (2007). Perceptual image retrieval using eye movements. *International Journal of Computer Mathematics, 84*(9), 1379-1391.

Pachur, T., Bröder, A., & Marewski, J. N. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making, 21*(2), 183-210.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). *The PageRank citation ranking: Bringing order to the web*. Stanford, CA: Stanford University.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*(2), 220-244.

Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science, 29*(4), 559-586.

Petrov, A. A. (2006). Computationally efficient approximation of the base-level learning equation in ACT-R. In D. Fum, F. Del Missier & A. Stocco (Eds.), *Proceedings of the Seventh International Conference on Cognitive Modeling* (pp. 391-392). Trieste, ITA.

Pirolli, P. (2005). Rational analyses of information foraging on the web. *Cognitive Science, 29*(3), 343-373.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643-675.

Pirolli, P., & Fu, W. T. (2003). SNIF-ACT: A model of information foraging on the world wide web, *Proceedings of the Ninth International Conference on User Modeling*.

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

Puolamäki, K., Salojärvi, J., Savia, E., Simola, J., & Kaski, S. (2005). Combining eye movements and collaborative filtering for proactive information retrieval, *28th ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 146-153). Salvador, Brazil: ACM.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 216-270). Cambridge, MA: MIT Press.

Qvarfordt, P., & Zhai, S. (2005). Conversing with the user based on eye-gaze patterns, *CHI 2005*. Portland, OR.

Raptis, D., Tselios, N., & Avouris, N. (2005). Context-based design of mobile applications for museums: A survey of existing practices, *MobileHCI'05* (pp. 153-160). Salzburg, Austria: ACM.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111*(1), 159-182.

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.

Rayner, K., & Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture word interference task. *British Journal of Psychology, 77*, 207-222.

Ries, J. E., Su, K., Peterson, G., Sievert, M. C., Patrick, T. B., Moxley, D. E., et al. (2001). Comparing frequency of content-bearing words in abstracts and texts in articles from four medical journals: An exploratory study. In V. L. Patel, R. Rogers & R. Haux (Eds.), *MEDINFO 2001: proceedings of the 10th World Congress on Medical Informatics* (pp. 381-384).

Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review, 13*(2), 310-315.

Ritter, F. E., & Young, R. M. (2001). Embodied models as simulated users: Introduction to this special issue on using cognitive models to improve interface design. *International Journal of Human-Computer Studies, 55*(1), 1-14.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1-3), 107-142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*(3), 249-284.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modelling attentional control in the Stroop task. *Psychological Review, 110*(1), 88-125.

Rosenbloom, P. S., Laird, J. E., & Newell, A. (1993). The Soar papers: Readings on integrated intelligence: MIT Press, Cambridge, MA.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217-236.

Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin & Review, 8*(1), 73-80.

Salton, G., & McGill, M. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.

Salton, G., Wong, A., & Yang, C. S. (1975). Vector-space model for automatic indexing. *Communications of the ACM, 18*(11), 613-620.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review, 115*(1), 101-130.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3*(1), 1-17.

Schenk, D. (1997). *Data smog: Surviving the information glut*. New York: Harper and Collins.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 59*(2), 132-138.

Schumacher, E. H., Lauber, E. J., Glass, J. M., Zurbriggen, E. I., Gmeindl, L., Kieras, D. E., et al. (1999). Concurrent response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 791-814.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., et al. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*(2), 101-108.

Shah, K., Rajyaguru, S., St. Amant, R., & Ritter, F. E. (2003). Connecting a cognitive model to dynamic gaming environments: Architectural and image processing issues. In F. Detje, D. Dörner & H. Schaub (Eds.), *Fifth International Conference on Cognitive Modeling* (pp. 189-194). Bamberg: Universitäts-Verlag Bamberg.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM - retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*(2), 145-166.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to study of information processing in judgment. *Organizational Behavior and Human Performance, 6*(6), 649-744.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General, 109*(4), 373-392.

Sparacino, F. (2002). *The museum wearable: Real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences.* Paper presented at the proceedings of Museum and the Web 2002, Boston, MA.

Starker, I., & Bolt, R. A. (1990). A gaze-responsive self-disclosing display. *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, 3-10.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-315.

Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Kruger, A., Kruppa, M., et al. (2007). Adaptive, intelligent presentation of information for the museum visitor in PEACH. *User Modeling and User-Adapted Interaction, 17*(3), 257-304.

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

Sun, R. (2006). The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In R. Sun (Ed.), *Cognition and multi-agent interaction*. New York, NY: Cambridge UP.

Sun, R. (2007). The motivational and metacognitive control in CLARION. In W. D. Gray (Ed.), *Integrated models of cognitive systems*. New York, NY: Oxford UP.

Taatgen, N. A. (2005). Modeling parallelization and flexibility improvements in skill acquisition: From dual tasks to complex dynamic skills. *Cognitive Science, 29*(3), 421-455.

Taatgen, N. A. (in press). Consciousness in the ACT-R architecture. In *Oxford companion to consciousness*: Oxford UP.

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., & Anderson, J. R. (2008). Constraints in cognitive architectures. In R. Sun (Ed.), *Handbook of computational psychology* (pp. 170-185): Cambridge UP.

Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (in press). Too much control can hurt: A threaded cognition model of the attentional blink. *Cognitive Psychology*.

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.

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14*, 1-36.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599-606.

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

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

Turney, P. D. (2001). *Mining the web for synonyms: PMI-IR versus LSA on TOEFL.* Paper presented at the the Twelfth European Conference on Machine Learning, Freiburg, Germany.

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. (2007). Mediating expert knowledge and visitor interest in art work recommendation, *Proceedings of LWA 2007* (pp. 367-372). Halle (Saale), Germany.

Van Maanen, L., Janssen, C. P., & Van Rijn, H. (2006). Personalization of a virtual museum guide using eye-gaze, *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (pp. 2620). Vancouver, BC.

Van Maanen, L., & Van Rijn, H. (2006). An accumulator model account of semantic interference in memory retrieval. In D. Fum, F. Del Missier & A. Stocco (Eds.), *Proceedings of the 7th International Conference on Cognitive Modeling* (pp. 322-327). Trieste, ITA.

Van Maanen, L., & Van Rijn, H. (2007a). Accounting for subliminal priming in ACT-R. In R. L. Lewis, T. A. Polk & J. E. Laird (Eds.), *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 1-6). Ann Arbor, MI.

Van Maanen, L., & Van Rijn, H. (2007b). An accumulator model of semantic interference. *Cognitive Systems Research, 8*(3), 174-181.

Van Maanen, L., & Van Rijn, H. (2008). The picture-word interference effect is a Stroop effect after all. In V. Sloutsky, B. Love & K. McRae (Eds.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 645-650). Washington DC.

Van Maanen, L., Van Rijn, H., & Borst, J. P. (submitted). Stroop and picture-word interference are two sides of the same coin.

Van Maanen, L., Van Rijn, H., & Taatgen, N. A. (submitted). Accumulators in context: An integrated theory of context effects on memory retrieval.

Van Maanen, L., Van Rijn, H., Van Grootel, M., Kemna, S., Klomp, M., & Scholtens, E. (in press). Personal Publication Assistant: Abstract recommendation by a cognitive model. *Cognitive Systems Research*.

Van Ossenbruggen, J., Amin, A., Hardman, L., Hildebrand, M., Van Assem, M., Omelayenko, B., et al. (2007, march 2007). *Searching and annotating virtual heritage collections with semantic-web techniques.* Paper presented at the Museums and the Web, San Francisco, CA.

Van Rij, J., Van Rijn, H., & Hendriks, P. (submitted). Cognitive architectures and language acquisition: A case study in pronoun comprehension.

Van Rijn, H., & Anderson, J. R. (2003). Modeling lexical decision as ordinary retrieval. In F. Detje, D. Dörner & H. Schaub (Eds.), *Proceedings of the Fifth International Conference on Cognitive Modeling* (5 ed., pp. 207-212). Bamberg. GER: Universitätsverlag Bamberg.

Van Rijn, H., & Anderson, J. R. (2004). *The pseudohomophone baseword frequency effect explained.* Paper presented at the Architectures and Mechanisms for Language Processing conference (AMLaP) Aix-en-Provence.

Van Rijn, H., & Taatgen, N. A. (2008). Timing of multiple overlapping intervals: How many clocks do we have? *Acta Psychologica, 129*(3), 365-375.

Van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *Neuroimage, 27*(3), 497-504.

VanSelst, M., & Jolicoeur, P. (1997). Decision and response in dual-task interference. *Cognitive Psychology, 33*(3), 266-307.

Velichkovsky, B., Pomplun, M., & Rieser, J. (1996). Attention and communication: Eye-movement-based research paradigms. In W. H. Zangemeister, H. S. Stiehl & C. Freksa (Eds.), *Visual attention and cognition* (pp. 125-154). Amsterdam: Elsevier Science.

Vukovic, R. K., Wilson, A. M., & Nash, K. K. (2004). Naming speed deficits in adults with reading disabilities: A test of the double-deficit hypothesis. *Journal of Learning Disabilities, 37*(5), 440-450.

Wagenmakers, E. J. (in press). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*.

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., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., Van Rijn, H., & Zeelenberg, R. (2004a). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology, 48*(3), 332-367.

Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. P. 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. G. W. (2004b). Nonword repetition in lexical decision: Support for two opposing processes. *Quarterly Journal of Experimental Psychology Section a: Human Experimental Psychology, 57*(7), 1191-1210.

Welford, A. T. (1967). Single-channel operation in brain. *Acta Psychologica, 27*, 5-&.

Welford, A. T. (1980). The single-channel hypothesis. In A. T. Welford (Ed.), *Reaction time* (pp. 215-252). San Diego, CA: Academic Press.

West, R. L., Lebiere, C., & Bothell, D. (2006). Cognitive architectures, game playing, and human evolution. In R. Sun (Ed.), *Cognition and multi-agent interaction: From cognitive modeling to social simulation* (pp. 103-123). New York, NY: Cambridge UP.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual-search. *Psychonomic Bulletin & Review, 1*(2), 202-238.

Wooding, D. S. (2002). Fixation maps: Quantifying eye-movement traces, *symposium on Eye tracking research & applications*. New Orleans, LA: ACM Press.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

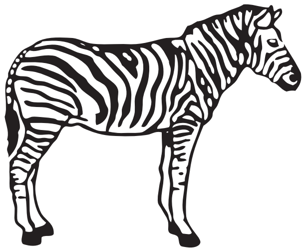
Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning Memory and Cognition, 32*(1), 1-14.

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.

Samenvatting

Het geheugen is één van de belangrijkste eigenschappen van de mens. Het geheugen speelt niet alleen een belangrijke rol bij het onthouden van feiten, of bij het leren van nieuwe vaardigheden, maar ook bij heel veel dagelijkse bezigheden. Voor bijvoorbeeld het lezen van dit proefschrift, of voor het voeren van een gesprek is het belangrijk om de betekenis van woorden te herkennen. In feite is dat een vorm van herinneren. Andere voorbeelden van zaken waarbij geheugen een rol speelt zijn het herkennen van objecten – waarbij je je de naam of de functie van het object moet herinneren (Figuur 1) – of het benoemen van eigenschappen van objecten, bijvoorbeeld de kleur, waarbij je je de naam van de kleur van het object moet herinneren.

Figuur 1. Een voorbeeld waarbij je geheugen nodig hebt: "Wat voor dier is dit?"



Binnen de cognitieve psychologie wordt onderzoek gedaan naar de werking van het geheugen. De cognitieve psychologie onderzoekt hoe cognitie ("het denken") je gedrag bepaalt. Geheugen is een belangrijk aspect van cognitie. Om te onderzoeken welke aspecten van cognitie op welke manier je gedrag beïnvloeden, worden meestal experimenten gedaan. Hiermee wordt onderzocht waarom de deelnemers aan die experimenten zich gedragen zoals ze zich gedragen. Deelnemers wordt dan gevraagd om zo snel maar ook zo correct mogelijk te reageren op een stimulus, bijvoorbeeld een afbeelding. Analyses van de snelheid van reageren en het foutenpatroon kunnen ons veel leren over de onderliggende mechanismes die het gedrag van mensen bepalen.

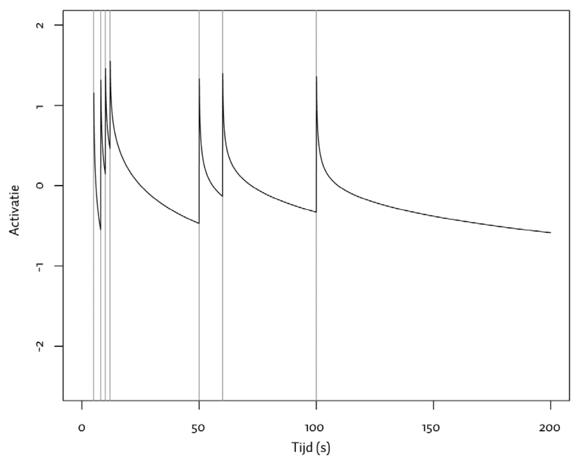
Dergelijke experimenten leiden in veel gevallen tot nieuwe theorieën. Omdat de theorieën in de cognitieve psychologie steeds complexer worden, worden tegenwoordig vaak cognitieve modellen ontwikkeld. Een cognitief model is een computersimulatie van die aspecten van menselijke intelligentie die een rol spelen bij het uitvoeren van het experiment. Als het model klopt, kan het hetzelfde gedrag produceren als de menselijke deelnemers aan het experiment. Het voordeel van zo'n computerprogramma is dat je er precieze voorspellingen voor gedrag mee kunt doen. Aan de ene kant dwingt dat je om je theorie heel precies te formuleren (want anders krijg je incorrecte voorspellingen), aan de andere kant kun je die voorspellingen ook weer toetsen met experimentele studies. Daarnaast is een voordeel van cognitieve modellen dat ze ook gebruikt kunnen worden in toepassingen waarbij het van belang is dat de computer de mens assisteert. Een cognitief model van een gebruiker van een computerprogramma kan precieze voorspellingen doen voor het gedrag van die gebruiker. Als het gedrag van een gebruiker van een computerprogramma bekend is, dan kan dat programma beter inspelen op de wensen van de gebruiker, en dus beter assisteren.

In deel 1 van dit proefschrift heb ik vooral het eerste aspect van cognitieve modellen benut. In dit deel beschrijf ik een aantal studies waarmee ik heb onderzocht hoe het geheugen werkt door cognitieve modellen te ontwikkelen die geheugenprocessen beschrijven. In deel 2 van het proefschrift komt het tweede aspect aan bod. In dit deel beschrijf ik een aantal toepassingen waarin computermodellen van geheugen een belangrijke rol spelen.

DEEL 1: HOE WERKT HET GEHEUGEN OP DE MILLISECONDE?

In deel 1 van dit proefschrift heb ik onderzocht hoe het *proces* van herinneren precies in zijn werk gaat. Eerder onderzoek beschrijft het wel of niet herinneren van een feit als het gevolg van hoe vaak en wanneer je dat feit eerder gezien hebt. Volgens deze “rationele theorieën” is het doel van geheugen om dat feit te vinden wat de hoogste kans heeft relevant te zijn in de huidige context. Dus als ik je vraag om zo snel mogelijk een dier te noemen, zullen de meeste mensen gewoonlijk eerder “Paard” dan “Zebra” zeggen. Dit komt omdat je normaal gesproken vaker paarden hebt gezien en dan zebra’s. Welke dierennaam het eerste bij je opkomt is echter contextafhankelijk. Omdat je op de vorige pagina een plaatje van een zebra hebt gezien en dit een hele recente ervaring is, is de kans dat je nu “Zebra” zegt ook groter. Eén implementatie van dit idee is geïntegreerd in de cognitieve architectuur ACT-R (Adaptive Control of Thought – Rational). ACT-R probeert verschillende aspecten van cognitie met elkaar in verband te brengen. Dat wil zeggen dat ACT-R-verklaringen van gedrag de nadruk leggen op de interactie van verschillende kernprincipes van cognitie, zoals langetermijn geheugen, perceptie, planning en timing. De mogelijkheden tot interactie van deze principes bepalen in grote mate hoe mensen zich gedragen. In ACT-R worden feiten – dingen die uit je langetermijn geheugen gehaald kunnen worden – weergegeven door middel van een activatiewaarde, die aangeeft hoe relevant dat feit op dit moment is (Figuur 2). Een recent feit is over het algemeen relevanter dan een feit dat je lang niet nodig hebt gehad, en een feit dat je veel nodig hebt is ook relevanter. Rationele theorieën (en ook ACT-R) geven een goede verklaring waarom je je bepaalde zaken makkelijker kunt herinneren dan andere. Maar deze theorieën hebben geen verklaring voor *hoe* je je iets herinnert.

Naast de rationele theorieën waarin het herinneren wordt beschouwd als een manier om de potentieel meest relevante feiten te vinden, bestaan er ook theorieën waarin het herinneren wordt beschreven als een proces. Deze theorieën worden wel “sequential sampling” modellen genoemd. Sequential sampling stelt dat er tijdens het ophalen van een bepaald feit een competitie is tussen verschillende feiten, die allemaal in meer of mindere mate waarschijnlijk zijn. Voor elke mogelijkheid wordt continu bewijs verzameld dat dat de meest relevante herinnering is, totdat voor één van de mogelijkheden veel meer bewijs is verzameld dan voor de andere. Het herinneren van de naam van het dier in Figuur 1 zou je kunnen zien als een competitie tussen “Zebra” en “Paard” (en misschien nog wel meer mogelijkheden). Er is

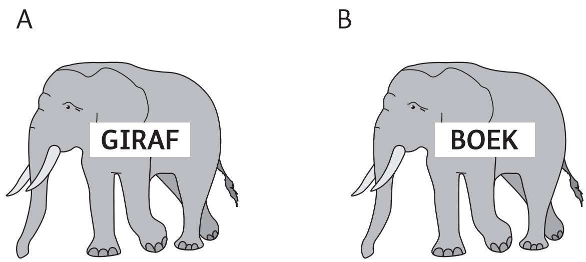


Figuur 2. De ACT-R activatie van een feit in langetermijn geheugen. De verticale lijnen geven de momenten aan waarop dit feit gebruikt wordt.

bewijs voor de “Paard”-mogelijkheid, omdat het afgebeelde dier vier benen heeft, en hoeven, en manen. De “Zebra”-mogelijkheid is echter waarschijnlijker, omdat het dier ook nog strepen heeft. De meeste mensen zullen dus “Zebra” zeggen in antwoord op de vraag “Wat is dit voor dier?”, omdat er meer bewijs is voor “Zebra” dan voor “Paard”. De sequential sampling theorie neemt aan dat de bewijslast sequentieel (dat wil zeggen stapje voor stapje) toeneemt, wat het proces van het ophalen van een feit uit je geheugen beschrijft.

De belangrijkste bijdrage van dit proefschrift is een specificatie van het langetermijn geheugen model in ACT-R. Door dat uit te breiden met een sequential sampling model kunnen we meer gedrag verklaren binnen één theorie (ACT-R) dan vroeger. Wij laten bijvoorbeeld zien dat het zogenaamde Stroop-effect (genoemd naar een experiment dat de Amerikaan John Ridley Stroop in 1935 heeft uitgevoerd) met deze theorie verklaard kan worden. Het Stroop-effect is de bevinding dat het heel moeilijk is om de kleur van een woord te noemen, als dat woord een kleur beschrijft (bijvoorbeeld, het woord “groen” in rode letters). Om het juiste antwoord (“rood”) te kunnen geven, moet je de naam van de kleur van het woord uit je geheugen ophalen. Voor zowel rood als groen wordt echter bewijs verzameld, omdat het allebei kleuren zijn. Daarom ben je dus langzamer en maak je meer fouten wanneer je de kleur van kleurwoorden wilt zeggen, dan wanneer je de kleur van andere woorden wilt zeggen.

Het Stroop-effect is een voorbeeld van een experiment waarbij de competitie tussen verschillende feiten in het geheugen (in dit geval kleuren) zorgt voor meer fouten en langzamer antwoorden. Hetzelfde principe geldt ook als je gevraagd wordt een afbeelding te benoemen waar een woord doorheen geschreven staat (een zogenaamde PWI-taak, voor Plaatje-Woord Interferentie). Een afbeelding van een olifant is bijvoorbeeld moeilijker te herkennen wanneer het woord “giraf” erdoorheen geschreven staat, dan wanneer een woord dat niets met olifanten te maken heeft erdoorheen geschreven staat (bijvoorbeeld “boek”, Figuur 3).



Figuur 3. Een voorbeeld van een stimulus waarbij er competitie is tussen verschillende eigenschappen van de stimulus. A. Wel competitie; B. Geen of minder competitie.

In deze taak blijkt dat de reactietijden afhangen van het interval tussen het aanbieden van het woord en het aanbieden van de afbeelding. De reactietijden zijn het grootst als een gerelateerd woord (“giraf”) door een afbeelding geschreven wordt, 100ms nádat de afbeelding getoond wordt. Deze moeilijk te verklaren observatie past binnen onze theorie, als we er vanuit gaan dat woorden sneller gelezen worden dan afbeeldingen benoemt worden. Vanwege dit verschil in verwerkingssnelheid kan het woordverwerkingsproces nog tijd goedmaken, waardoor het grootste effect optreedt als het woord na de afbeelding wordt aangeboden.

Een andere belangrijke observatie is dat als mensen *tegelijkertijd* met een Stroop of PWI-taak een tweede taak moeten uitvoeren, een ander effect in de PWI-taak optreedt dan in de Stroop-taak. Hieruit werd in de literatuur geconcludeerd dat PWI en Stroop effecten veroorzaakt werden door verschillende mechanismen. Onze theorie laat zien dat er geen fundamenteel verschil is tussen de PWI-taak en de Stroop-taak. Wij verklaren de verschillen

A B C

Figuur 4. A: De Nachtwacht; B: De anatomische les; C: Het korporaalschap van kapitein Albert Bas en luitenant Lucas Conijn.

door te kijken naar de verschillen tussen kleuren en afbeeldingen. Als we aannemen dat kleuren sneller worden herkend dan afbeeldingen, kunnen we de verschillen tussen de Stroop-taak en de PWI-taak verklaren zonder dat we uit hoeven te gaan van verschillende mechanismen.

Naast de hier besproken voorbeelden, hebben we onze theorie ook nog toegepast op lexicale decisie experimenten en subliminale perceptie om de validiteit te verstevigen. Tijdens lexicale decisie experimenten wordt proefpersonen gevraagd om van een letterstring (BALC of BALK) te zeggen of het een correct Nederlands woord vormt of niet. Je kunt het gedrag in deze taak beschrijven door aan te nemen dat hiervoor geheugen nodig is. Een cognitief model dat wij voor deze taak ontwikkeld hebben voorspelt correcte reactietijden voor een aantal fenomenen die je bij lexicale decisie waar kunt nemen. In een subliminale perceptie experiment wordt de stimulus zo kort aangeboden dat de proefpersoon zich niet bewust is van de waarneming. Toch kan de proefpersoon dan nog boven kansniveau zeggen wat hij of zij heeft waargenomen. Ook dit kan gemodelleerd worden met onze theorie. Gezamenlijk dragen de cognitieve modellen die worden besproken in deel 1 van dit proefschrift bij aan de betrouwbaarheid van de geheugentheorie die in dit proefschrift ontwikkeld is.

DEEL 2: GEHEUGENMODELLEN IN DE PRAKTIJK

Naast meer begrip over de werking van het geheugen levert een cognitief model van geheugen nog iets anders op: Je kunt het model gebruiken om te voorspellen welke begrippen mensen in een bepaalde situatie waarschijnlijk zullen gebruiken. Een computerprogramma dat dat weet, kan daar op in spelen. Met behulp van cognitieve modellen zou je bijvoorbeeld kunnen voorspellen welke informatie mensen op internet willen opzoeken. Een zoekmachine zoals Google zou dan al suggesties voor nieuwe zoektermen kunnen doen, op basis van het huidige zoekgedrag. In deel 2 van dit proefschrift wordt dit idee geïllustreerd aan de hand van twee voorbeelden, de Publicatie Assistent en de Virtuele Museumgids.

Uit eerder onderzoek is gebleken dat het gebruik van woorden behoorlijk goed voorspeld kan worden door te kijken wanneer en hoe vaak een bepaald woord al eerder gebruikt is. Hoe vaker je een woord gebruikt hebt, hoe groter de kans op hergebruik, maar hoe langer geleden je het woord gebruikt hebt, hoe kleiner de kans op hergebruik. Daarnaast speelt de context waarin je je bevindt en waarin je die woorden eerder gebruikt hebt ook een rol.

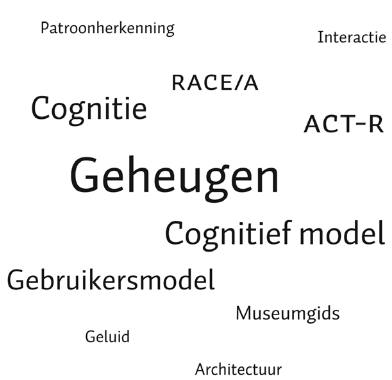
Op basis van deze drie eigenschappen hebben wij een computerprogramma ontwikkeld dat gebruikersprofielen voor museumbezoekers opstelt. De Virtuele Museumgids bouwt een profiel van een museumbezoeker aan de hand van beschrijvingen (woorden) van kunstwerken die de bezoeker gezien heeft en de interesse die de bezoeker heeft getoond voor die kunstwerken.

De interesse van de bezoeker hebben we onder andere gemeten door naar de oogbewegingen van de bezoeker te kijken. De aanname was dat museumbezoekers vooral kijken naar de dingen die ze interessant vinden. Gegeven dat deze aanname klopt, kunnen we dan het gebruikersprofiel verbeteren door de gemeten oogbewegingen op te nemen in het model? Ondanks dat we wel aanwijzingen hebben gevonden dat museumbezoekers voornamelijk kijken naar delen van schilderijen waarin ze geïnteresseerd zijn, en zich niet laten "sturen" door het verhaal dat de museumgids houdt bij een schilderij, bleek het nog niet eenvoudig om dit principe ook echt te gebruiken in een Virtuele Museumgids applicatie.

Naast het gebruikersprofiel van de bezoeker, heeft de Virtuele Museumgids ook een geheugen voor alle kunstwerken die er in het museum zijn, met de samenhang daartussen. De Virtuele Museumgids combineert zijn kennis van de museumcollectie met het gebruikersprofiel van de bezoeker, om een zinvolle suggestie te doen welk kunstwerk bezichtigd moet worden. Het volgende voorbeeld illustreert dit principe. Stel de museumbezoeker geeft aan dat hij of zij "De Nachtwacht" van Rembrandt erg mooi vindt. De Virtuele Museumgids berekent vervolgens welk kunstwerk daar het meest mee te maken heeft. Dit kan bijvoorbeeld "De anatomische les" van Rembrandt zijn. Als de bezoeker nu aangeeft dat hij of zij dit niet zo mooi vindt, dan probeert de Virtuele Museumgids een kunstwerk te vinden dat veel samenhang vertoont met "De Nachtwacht", maar weinig met "De anatomische les". Een voorbeeld is "Het korporaalschap van kapitein Albert Bas en luitenant Lucas Conijn", geschilderd door Govert Flink. Dit werk beeld ook een schutterscompagnie af, net als De Nachtwacht, maar is niet door Rembrandt geschilderd (Figuur 4).

Een belangrijke eigenschap van kunst die helaas niet goed door het cognitieve model kon worden meegenomen bij het inschatten van interesses, is dat kunst vaak mooi of interessant wordt gevonden vanwege een bepaald kleurgebruik, of een bepaalde stijl of techniek. Doordat deze aspecten niet makkelijk te kwantificeren zijn, hebben we een tweede toepassing, de Publicatie Assistent, ontwikkeld waarbij dergelijke affectieve aspecten een kleinere of geen rol spelen.

Voor de Publicatie Assistent hebben we gebruikt gemaakt van een vergelijkbaar idee als voor de gebruikersprofielen in de Virtuele Museumgids. Deze museumgids-profielen voorspellen welke woorden een museumbezoeker interessant vindt. De profielen in de Publicatie Assistent voorspellen welke woorden een wetenschapper waarschijnlijk belangrijk acht of interessant vindt. Het belang van elk woord voor een bepaalde wetenschapper wordt ingeschat op basis van hoe vaak hij of zij dat woord gebruikt heeft, wanneer, en in welke context. De aanname achter de Publicatie Assistent is dat deze drie eigenschappen voorspellen of iemand een bepaald woord belangrijk vindt (Figuur 5).



Figuur 5. Een "tag-cloud" waarbij de grootte van de begrippen aangeeft wat de wetenschappelijke interesse van de auteur hierin is. De grootte is bepaald op basis van het gebruik van deze begrippen in wetenschappelijke publicaties van de auteur.

De profielen kunnen gebruikt worden om relevante wetenschappelijke literatuur te vinden. Deze kunnen gebruikt worden om te kijken of de belangrijke woorden ook voorkomen in nieuwe publicaties van andere wetenschappers. Als de voor een bepaalde wetenschapper belangrijke woorden in een publicatie voorkomen, dan is dat waarschijnlijk een interessant artikel om te lezen. We hebben de Publicatie Assistent vergeleken met andere systemen die ook literatuursuggesties kunnen doen. Daaruit bleek dat onze geheugengebaseerde methode goed de keuzes van individuele onderzoekers voorspelt. Daarnaast laat onze studie ook zien dat er veel variatie zit in de selectiecriteria die verschillende mensen gebruiken om hun keuzes op te baseren. Een nieuwe uitdaging is om deze individuele verschillen ook in het cognitieve model op te kunnen nemen.

CONCLUSIES

De nieuwe cognitieve modellen die besproken worden in dit proefschrift schetsen een beeld van cognitie waarin geheugen een cruciale rol speelt. Zonder herinnering van eerder opgedane feitenkennis kunnen mensen niet functioneren. De theorie in dit proefschrift beschrijft het proces van ophalen van feiten als een sequential sampling proces waarbij continu de waarschijnlijkheid wordt geschat dat een bepaald feit relevant is. Daarnaast zorgt het mechanisme achter herinneren er echter ook voor dat in speciale gevallen het proces juist moeizamer gaat. Zoals bij het Stroop-effect, waarbij de betekenis van de letters van een woord (“rood”) de kleur (groen) juist moeilijker te herinneren maken. Met name wordt in dit proefschrift ook de interactie tussen geheugenprocessen en andere aspecten van cognitie benadrukt. De verschillende experimenten en cognitieve modellen laten zien dat deze interactie cruciaal is voor om een compleet beeld te krijgen van het gedrag van mensen.

Begrip van de werking van het geheugen betekent niet alleen meer begrip van de werking van cognitie. Het betekent ook dat we betere computerprogramma’s kunnen ontwikkelen die gebruik maken van de kennis die we hebben opgedaan over de mechanismes achter het ophalen van feitenkennis uit je geheugen. Een eerste aanzet daartoe wordt gegeven in dit proefschrift. De toepassingen die ontwikkeld zijn, laten zien dat het gebruik van cognitieve modellen in computerprogramma’s waarbij een persoonlijke selectie uit informatie gemaakt moet worden zinvol is, en een belangrijke toevoeging kan zijn aan reeds bestaande technologie.

Dankwoord

Iedereen zegt altijd dat AI@Rug zo'n leuke club is, maar dat is ook echt waar! Ik wil graag iedereen bedanken die de afgelopen vier jaar tot een hele fijne tijd hebben gemaakt, zowel in de Bernoulliborg als daarbuiten:

Als eerste wil ik graag Hedderik bedanken. Niet alleen was je mijn dagelijkse begeleider en nu copromotor, maar daarnaast hebben we ook veel samengewerkt, samengereist en samengepraat. Je hebt me op ontelbare manieren geholpen om wetenschapper te worden.

Daarnaast wil ik ook graag mijn promotor Lambert en copromotor Niels bedanken, omdat ze beiden – gedurende het hele promotietraject, maar vooral ook in de laatste fase – zinvolle en belangrijke feedback op het proefschrift en de promotie hebben gegeven.

Gerard, Jeroen, and Wayne, thank you for taking the time to thoroughly read my thesis and assess its value.

Julian, thanks for a great collaboration, which lead to Chapter 8 of this thesis. Maarten, Stephanie, Martin en Erwin, bedankt voor jullie bijdrage aan het paper wat tot Hoofdstuk 7 van dit proefschrift leidde.

Jelmer en Gert, super dat jullie mijn paranimfen willen zijn!

Bovenstaande mensen zijn in meer of mindere mate verantwoordelijk geweest voor het welslagen van mijn proefschrift en de verdediging daarvan. Daarnaast hebben velen gezamenlijk een omgeving gecreëerd waarin ik mij heel erg thuis voelde (en voel). Allereerst de Cognitive Modeling groep, Hedderik, Niels, Fokie, Jelmer, Jacolien, Chris en Ben, bedankt voor alle feedback op mijn soms maar half affe presentaties, manuscripten, en experimenten. Alle KI-Aios (en een aantal anderen, die ik toch maar onder deze groep schaar), bedankt voor alle uren tijdens en na het werk waarin we soms over KI en cognitie spraken, maar meestal ook niet: Anja, Axel, Bea, Dirkjan, Elske, Gert, Hedde, Ilse, Jacolien, Jelmer, Jolie, Karin, Maria, Marius, Renante, Sjoerd, Tijn en Wouter. Speciaal uit deze groep moeten diegenen genoemd worden die omdat ze mijn kamergenoot waren ongevraagd van hun werk werden gehouden (alhoewel het soms ook andersom was, Chris): Anja, Chris, Jacolien, Jelmer, Karin en Wouter, tezamen met Anton, Ben en Lennart. Overige KI'ers ook bedankt: Arnold, Bart, Bart (met name dank voor de samenwerking aan de newsletter), Esther, Jennifer, Marco, Rineke, Ronald, Ronald, Sietse, Sonja, Sujata en Tjeerd.

Maarten, dank voor het vele werk dat je in de lay-out van het proefschrift hebt gestoken. *Bijna* alle pagina's zien er fantastisch uit! Rudi en Marjan, bedankt voor het vertrouwen en enthousiasme dat jullie hadden toen ik aankondigde mijn vaste aanstelling in te ruilen voor promotieonderzoek.

Speciale dank ook voor de vrienden die altijd geïnteresseerd bleven in wat ik deed en doe, ondanks dat het soms moeilijk uit te leggen is (maar niet altijd!): Carl, Ellen, Eva, Freek, Irene, Laurens, Rob, Sandra, Stefan en Wieke.

En natuurlijk mijn lieve Anneke, jij eigenlijk bedankt voor Alles!

Notes