

CAUSAL INFERENCES AND WORLD KNOWLEDGE

Leo Noordman

Wietske Vonk

Reinier Cozijn

Stefan Frank

To appear in: E.J. O'Brien, A.E. Cook, & R.F. Lorch (Eds.). *Inferences during reading*.

© Cambridge University Press

## Theoretical background

Research on inferences in discourse processing has been characterized for a long time by the question of what kinds of inference are likely made on-line and what kinds of inference are not. Considering that in principle an unlimited number of inferences can be made while the human processing capacity is limited, the question arises how the inferences are controlled. In the Seventies and Eighties, one tried to answer this question by classifying inferences, assuming that the different classes vary in the likelihood of being made. The classifications were in general based on textual characteristics, such as whether the inference is a forward or a backward inference in the text; whether it is necessary for coherence or just elaborative, whether it necessarily follows from the text or only possibly. Inferences were also classified with respect to their content, for instance, as inferences about instruments, causes, consequences, goals, time, or place. There was some consensus that inferences that are necessary for coherence, including backward bridging inferences, are made on-line and other kinds of inference are not (Singer, 1994, 2007; Van den Broek, 1994). But it was by then already clear that that conclusion fell short (Noordman, Vonk, & Kempff, 1992). Whether an inference is made does not only depend on its function in the text, but also on the availability of the required information and the reader's knowledge. This has also been acknowledged in the two main views that emerged in the Nineties: the minimalist view and the constructionist view. According to the minimalist view (McKoon & Ratcliff, 1992), only two classes of inferences are made during reading, those that are based on easily available information and those that are required for local coherence. Crucial here is how to establish what 'easily available information' is. The constructionist view (Graesser, Singer, & Trabasso, 1994) is characterized by a search (or effort) after meaning. Readers have a particular goal in reading: construct a coherent meaning representation –a mental model– and explain what is being read. In constructing such a mental model, world knowledge is activated. The main issue in both the minimalist view and the constructionist view was the question of which inferences are made on-line. Inferences were still considered as a rather isolated phenomenon rather than as a component process in discourse comprehension.

In the last decade, research on inferences got embedded in a broader cognitive framework. It seemed to be promising to conceive inferences as a component process in discourse comprehension and to relate inferences to the readers' world knowledge (Vonk & Noordman, 2001). This has been proposed in the memory-based approach to text processing (Gerrig & McKoon, 1998). In this approach, an inference is a process that takes place in active memory and makes contact with information in long term memory, which includes information from the previous discourse representation and world knowledge. Given that the controlling factor for inferences is 'easily available information', some rapprochement between the minimalist and constructionist views on

inferences emerges. Three observations may illustrate this rapprochement. First, since easily available information includes parts of the previous discourse (including knowledge about the goal of the protagonist) as well as world knowledge about the topic, it does not follow from 'easily available information' that only a very restricted class of inferences is made (Noordman & Vonk, in press). Second, the experiments by O'Brien, Myers, and colleagues demonstrated that the claim that readers attend to only local coherence was too strong because global inferences about the protagonist were made even though the text was locally coherent (Albrecht & O'Brien, 1993). O'Brien and colleagues argued that even though global information concerning a protagonist did not remain active' in memory (Myers, O'Brien, Albrecht, & Mason, 1994), such information was 'readily available' through a fast-acting resonance process (see O'Brien, 1995, for a review). In fact, a characteristic of the protagonist can remain "available" even if it is mentioned quite a distance ago. This is because information in a discourse is not only linearly but also hierarchically structured (van Dijk & Kintsch, 1983). Third, even the role of goal-directed behavior, which is a characteristic of constructionism, is not altogether incompatible with minimalism and can be accommodated within memory-based processing: The reader's goal dictates what information the reader is after, which, in turn, determines what information is currently in focus and available.

The position we advocate is also nicely articulated by Gerrig and O'Brien (2005) in their discussion of memory-based processing. They argue that whether an inference is made or not does not depend on the *kind* of inference, but on the activation of information in memory: "Inferences are encoded to the extent that information in active memory makes contact with relevant or necessary information from inactive portions of the discourse model and general world knowledge" (Gerrig & O'Brien, 2005, p. 236). They also acknowledge the role of the reader's goals in their approach. The crucial mechanism in their memory-based processing is 'resonance' (Ratcliff, 1978; Myers & O'Brien, 1998): Inactive information of the previous discourse representation and world knowledge becomes activated in proportion to its match with the incoming text input.

We take memory-based processing as our framework. Both textual information and the world knowledge of the reader can be easily available information. According to memory-based processing, concepts from earlier in the discourse representation and world knowledge resonate with the incoming text. But the notion of memory-based processing allows for other sources of activation than just the incoming text. There are factors related to the reader that can increase the accessibility of information. For instance, the reader's goal can influence what information is in focus. Concepts in current focus get more attention which increases their activation and accessibility. Furthermore, some information is intrinsically more available than other information. Specifically, human beings have the propensity to deal with causal relations (Noordman, 2005): Consecutive sentences are preferably

interpreted in a causal way, supporting Levinson's (1983) principle of informativeness. According to Kant, causality is one of the conditions that we impose on our experience: "The empirical knowledge of phenomena is only possible by the fact that we subject the succession of phenomena ... to the law of causality." (Kant, 1787, B234). Causality is a preferred mode of interpretation and in that sense causal information is easily available.

In this paper we review our research that demonstrates how causal inferences depend on world knowledge. The first part addresses experimental studies on knowledge and inferences. The second part addresses the modeling of knowledge-based inferences. In the first part, world knowledge is manipulated in two different ways. In the first section of that part, we vary the familiarity of the causal relations expressed in the texts while keeping the population of readers constant. Results show that inferences of unfamiliar causal relations, even if they are marked by the conjunction *because*, are not made on-line. It will turn out that inferences of unfamiliar causal relations are made only if they are made more available under the influence of the reader's goal. Inferences that concern familiar causal relations, on the other hand, *are* made on-line. But a prerequisite is that the relations are marked by a conjunction, in our case by the conjunction *because*. In the second section of the experimental part, we vary the expertise of the readers in a particular domain, while keeping the texts constant. Experts and non-experts in economics were compared while reading the same texts in the domain of economics. Results of these expert-vs.-non-expert studies support the conclusion from the experiments in which the familiarity of the causal relations is manipulated. Non-experts in economics do not make the inferences about the causal relations on-line, although they are familiar with the concepts in the relations. Experts, on the other hand, do make the inferences. Interestingly, the experts made the inferences even in the absence of the conjunction *because*. Apparently, experts read texts that concern their expertise in a more careful way.

The second part of the paper deals with modeling knowledge-based inferences. We first review the role of world knowledge in models of discourse processing and inferencing. We then discuss a formal model that we developed earlier in which inferences are considered as world knowledge applications (Frank, Koppen, Noordman, & Vonk, 2003). The model simulates the process of causal inferences and explains a number of findings from the reading experiments, but it does not yet account for all the empirical data. We discuss how further extensions of the model might explain the findings from our reading experiments.

## Experimental studies on knowledge and inferences

### *Manipulating knowledge: familiar and unfamiliar information*

In the experiments presented in this section, we investigated the role of world knowledge by varying the familiarity of the information to be inferred. The inferences concerned causal relations that were unfamiliar to readers (Noordman, Vonk, & Kempff, 1992) or familiar to readers (Vonk & Noordman, 1990; Cozijn, Commandeur, Vonk, & Noordman, 2011; Cozijn, Noordman, & Vonk, 2011).

### *Unfamiliar relations are not inferred*

Noordman et al. (1992) investigated whether backward causal inferences are made on-line during the processing of causal relations in text. Expository texts --with an average length of 7.6 sentences -- contained a causal relation that was signaled by the causal connective *because*. The causal relations were unfamiliar to the readers, as was checked in a questionnaire. For example, in a text on spray cans, the causal relation read "Chlorine compounds make good propellants, because they react with almost no other substance." Based on the connective *because*, this sentence can be analyzed in terms of syllogistic reasoning. The sentence expresses the conclusion of the syllogism -- chlorine compounds make good propellants- and the minor premise -- they (i.e., chlorine compounds) react with no other substances. What is missing from the syllogism is the major premise: Propellants must not react with the product in the spray can. The connective *because* signals that the information of the major premise has to be inferred in order to justify the causal relation. This inference is backward and contributes to the coherence of the discourse.

To test whether the inference was made during reading, the causal relation sentence was either preceded or not preceded by a sentence explicitly stating the major premise. This explicitly mentioned premise was supposed to facilitate the inference. If the explicit information is not given, it should take more time to make the inference. This should lead to a longer reading time of the *because* sentence. After reading the text, participants had to verify a statement that contained the explicit information. Because the inference was expected to be made on-line in both the explicit and the implicit conditions, no difference was predicted for the verification times.

The results did not provide evidence for inferential processing during reading: Reading times did not differ between the explicit and the implicit conditions (see Table 1, line A1). Apparently, the explicit information did not activate the inferential relation between the concepts.

The explicit information, however, did help readers to verify the inference sentence after reading the text. Verification times were shorter in the explicit than in the implicit condition (see Table 1, line A2), suggesting that the inferences were made during verification.

Table 1. Mean reading and verification times (ms) in the causal relation experiments (the codes in the first column are for textual reference).

	Task	Connective	Measure	Information		
				Explicit	Implicit	
A1	Comprehension	<i>because</i>	Reading time	3807	3921	
A2			Verification time	2657	2955	*
B1	Topic related question	<i>because</i>	Reading time	4700	5411	*
B2	Non-topic related question		Reading time	4243	4014	
C1	Inconsistency judgement	<i>because</i>	Reading time	5478	6093	*
C2			Verification time	3594	3867	
D1	Comprehension	<i>but</i>	Reading time	1473	1536	*
D2			Verification time	2169	2198	

\* Significant difference

The combined set of results indicated that the causal inference was not made during reading, but, instead, after reading the text when the task required the readers to do so.

To obtain confirming evidence for this result, a follow-up experiment was conducted with similar materials. But in this experiment, the explicit sentence with the major premise was changed into its opposite. For example, for the causal sentence “Chlorine compounds make good propellants, because they react with almost no other substance”, the explicit sentence read “Propellants must combine with the product in the spray can”, which contradicts the inference. Recall that the relations that had to be inferred were unfamiliar to the readers. In one condition, the participants were asked to read and check the texts for improvements. Seventy-six percent of the participants did not detect the inconsistency. In the other condition, participants were told that there might be an inconsistency in the text and, if so, that they had to correct it. Forty-five percent of the participants did not detect the inconsistency. This smaller percentage can be attributed to the effect of the instructed reading goal. But even in the latter case, a considerable percentage did not detect the inconsistency between the explicit information and the inference to be made, which is rather amazing. However, the fact that a substantial part of the participants did not make the inference on-line is consistent with the reading time results from the earlier experiment.

Conclusion: The results of the experiments with unfamiliar relations indicated that, unexpectedly, the inferences were not made during reading, even though the inferences would contribute to local coherence. These results suggest that the reader’s knowledge controls inferences.

*Unfamiliar relations are not inferred unless motivated by the reader’s goal*

Although readers did not make the causal inference *during* reading, even when the conjunction *because* would allow them to do so, they did make the inference after reading when they were instructed to verify the information. Furthermore, when readers were instructed to improve the texts, the inferences

were made, as was evidenced by the inconsistency detection. These two task-dependent effects suggest that on-line reading behavior may be affected by the reader's goal. The reader's goal can be defined in terms of the task demands, but also in terms of a specific topic. In agreement with memory-based processing, the reader's topic-goal may play a role in focusing the reader's attention to specific concepts and make these concepts more accessible. Both the goal in terms of task and in terms of topic were investigated in Noordman et al. (1992).

One experiment investigated whether the reader's topic-goal can motivate the reader to make inferences even if they are not familiar. As in the earlier experiments, participants read expository texts that dealt with unfamiliar causal relations. Participants had to give an oral answer to an explicit question about the text. The question was presented before reading the text. The texts had two types of question, but each participant was presented with only one question per text. The first type of question was about the causal relation. In case of the spray cans, the question was "How do spray cans work?" This question was intended to put the causal relation in current focus and thereby make it more accessible. Presumably, the causal relation is then more thoroughly processed and the inference made. The second type of question was about a different part of the text that occurred after the causal relation sentence. This question was not expected to elicit the causal inference during reading. The results supported the predictions: If the question was about the causal relation, the reading times on the causal relation sentences were shorter in the explicit condition than in the implicit condition (see Table 1, line B1). If the question was about a different part of the text, however, no difference between the conditions in the reading times on the causal relations sentences was observed (see Table 1, line B2). The interaction of text condition and type of question was significant. Apparently, having to answer a question about the causal relation urged the readers to make the inference on-line.

In another experiment, the reader's goal was defined by the task: Participants had to judge for each sentence whether it was consistent or inconsistent with its preceding context. To that end, one or more inconsistencies were inserted into the texts, but never into the causal relation sentence. The texts contained an explicit and implicit condition, as in the earlier experiments. Reading times of the *because* sentences were significantly shorter in the explicit condition than in the implicit condition and the verification times revealed no differences (see Table 1, lines C1 and C2), indicating that the task of checking sentences for inconsistencies elicited the inference to be made on-line.

Conclusion: These experiments showed that the goal of the reader is an important factor in determining whether inferences are made on-line. If the information to be inferred is unknown to the reader, the reader's goal can ensure that the inferences are made during reading. This has been

demonstrated for the goal defined in terms of topic and task. This inference behavior is tantamount to learning: Readers infer knowledge that they did not have before.

#### *Familiar relations are inferred*

The assumption that inferences are made if they are familiar to the reader has been addressed in Vonk and Noordman (1990). In this study, narrative texts on familiar topics were presented that contained a causal-contrastive relation sentence marked with the connective *but*. An example is the sentence “The room was large, but one was not allowed to make music in the room” in a text about a girl who is looking for a room to rent. As with the *because*-sentences above, the readers could make an inference based on the conventional implication of the connective, in this case the contrastive *but*. In the context of having to evaluate the acceptability of the room, the readers may infer that a large room is attractive for the girl, which is rather obvious, and that not being able to make music is unattractive for the girl, which implies that the girl plays an instrument. This inferential information was either given in the prior context or not. In the explicit condition, the prior text contained the sentence “The girl wanted to make music in the room.” In the implicit condition, this sentence was absent. After the text, a verification statement was presented that contained the information to be inferred from the text. If the inference was made on-line, the reading time on the *but* clause in the target sentence should be shorter in the explicit condition than in the implicit condition and there should be no difference in the verification times. The results were in line with the predictions (see Table 1, lines D1 and D2), supporting the conclusion that the inference was made on-line. The availability of prior knowledge is an important factor in determining whether inferences are made on-line.

The role of available knowledge in the processing of causal relations was also investigated in a listening study (Cozijn, Commandeur, Vonk, & Noordman, 2011). In a visual world paradigm experiment, participants listened to sentences that contained an implicit causality verb, like *to feel sorry for*: “The camel felt sorry for the octopus after the exam, because he could not get a pass mark for the work.” At the same time, their eye movements were monitored while they looked at a screen on which drawings of the two protagonists (camel and octopus) and a distractor item (a piece of paper with a pencil representing the exam) were presented (see Figure 1).



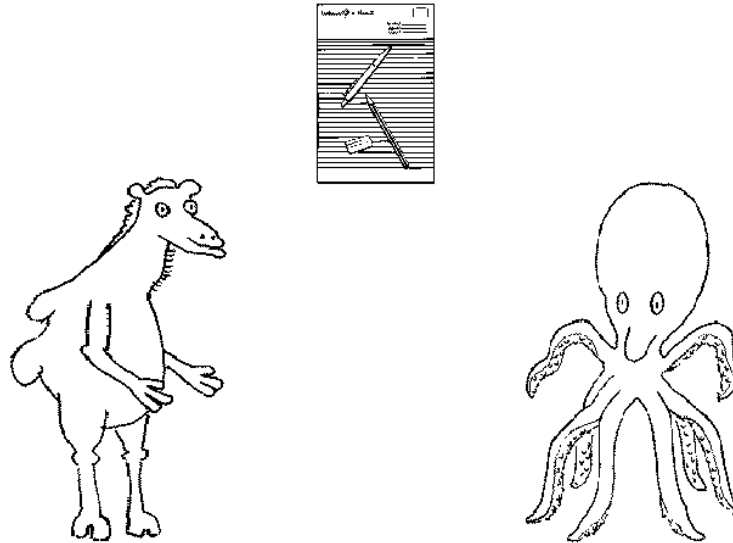


Figure 1. Example of a visual stimulus in the visual world paradigm experiment (vector traces of the original bitmap images). The three figures represent a camel, an octopus, and a piece of paper with pencils and an eraser representing the distractor ‘the exam’ (Cozijn et al., 2011, p. 385).<sup>1</sup>

The aim of the study was to find out how quickly comprehenders assign a referent to the ambiguous pronoun *he* in the *because* clause and, more specifically, whether they use the information in the verb to make a preliminary interpretation as soon as (or even before) they hear the pronoun. The preliminary interpretation of *he* in the present example is that *he* refers to the octopus. For present purposes, however, we are interested in the definitive interpretation of the pronoun that can only be made at the end of the sentence. The participants had to point with their index finger to the referent of the pronoun on the screen. The interpretation of the pronoun requires an inference. As explained above, understanding the connective *because* involves a backward causal inference; in this case, “A person who cannot get a pass mark is felt sorry for.” This inference is in line with our knowledge about exams (not getting a pass mark is a suitable reason for being felt sorry for by someone). In another condition, the pronoun had to be resolved by linking it to the other protagonist, the camel. For instance, “The camel felt sorry for the octopus after the exam, because he could not give a pass mark for the work.” This pronoun resolution goes against the preliminary interpretation. The eye-movement data showed that the inference was made at the end of the sentence: When the second clause was “because he could not get a pass mark for the work”, participants started looking at the correct referent at an earlier moment than when the second clause was “because he could not give a

---

<sup>1</sup> Reprinted from Reinier Cozijn, Edwin Commandeur, Wietske Vonk, and Leo Noordman. The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language*, 2011, 64 (4), 381-403 with permission from Elsevier.

pass mark for the work”, in which case world knowledge had to redress the preliminary pronoun assignment. With respect to the main question of the experiment, a preliminary interpretation of the pronoun was made very early, long before the disambiguating information was heard.

Conclusion: The results of the studies discussed in this section show that the availability of knowledge controls whether inferences are made during reading or listening: Inferences are made on-line only if their content is part of the reader’s knowledge.

The causal inference process can be described in the framework of memory-based processing. Words in the text activate concepts in the reader’s long term memory. If the relations between the concepts are familiar to the reader (i.e., if they are stored in long term memory), these relations are activated; they are checked against the reader’s world knowledge and added to the text representation. The explicit information facilitates the inference by activating the causal relation. If the relation between the concepts is not familiar but is signalled in the text (for example, by a causal conjunction), the relation can in principle be derived and added as new information to the text representation. The latter process is learning, since new information is then acquired from the text. Our results indicate that learning is not the normal mode in which readers process a text. Under normal reading circumstances, backward causal inferences, even if they are signalled by *because*, are not made during reading if they concern unfamiliar information. This result contrasts with our everyday belief that learning from text is self-evident and occurs all the time. Apparently, what is learned from text in normal reading depends on the reader’s knowledge: Inferences about unfamiliar relations are not made. New information is learned only if readers are motivated to really understand the text.

#### *Familiar relations are inferred only if supported by a conjunction*

In the introduction we claimed that humans have the propensity to interpret consecutive sentences in a causal way. The experiments showed that, for familiar relations, causal inferences were made on-line. How strong is this tendency to make causal inferences? Do the inferences depend on the presence of a conjunction such as *because*? This was investigated in Cozijn, Noordman, and Vonk (2011). Readers were presented with short narrative texts (seven sentences) on familiar topics that contained a familiar causal relation either signaled by *because* or not signaled. They processed causal relation sentences such as, “Mr. Smith was delayed because there was a large traffic jam on the highway”, while their eye movements were monitored. As explained above, the inference to be made when reading this sentence is the major premise of a syllogistic chain of reasoning. In this case, the inference is: “A traffic jam causes delay”. This inference is part of the reader’s knowledge. On the basis of the previous experiments, this inference should be made in presence of the conjunction, leading to an

increase in reading time. This will be manifest at the end of the sentence when the reader has the information to make the inference. If the inference depends on the presence of the conjunction *because* (Singer & O’Connell, 2003), the reading time at the end of the sentence should be shorter when the conjunction was absent than when it was present. To test this prediction, half of the texts did not contain the causal connective. The means of the regression path reading times are shown in Table 2, line A.

Table 2. Mean reading and verification times (ms) in the *because* experiments (the codes in the first column are for textual reference).

	Paradigm	Measure	Connective		
			Present	Absent	
A	Eye tracking	Reading time	512	448	*
B1	Self paced	Reading time	581	547	*
B2		Verification time	2317	2406	*

\* Significant difference

The results showed that if the connective *because* was present, reading times on the final words of the sentence were significantly longer than if the connective was absent. This indicates that the inference was made on-line only when the connective was present in the sentence. To ascertain whether the reading time difference could indeed be attributed to inferential processing, a second experiment with the same materials was conducted with a verification task that addressed the inferential information. Reading was self-paced. The results of the reading times analysis were in line with those of the first experiment: The reading times at the end of the sentence were longer if the connective was present than if it was absent (see Table 2, line B1). Importantly, the verification times were shorter if the connective was present in the text than if it was absent (see Table 2, line B2). This is in line with the reading time results, because if the inference was made during reading, verifying a statement about that information is easier than when the inference was not made.

Conclusion: The results of these experiments suggest that the presence of the connective *because* is necessary for readers to make the causal inference on-line.

#### *Manipulating knowledge: expert and non-expert readers*

In the preceding section, the effect of the reader’s knowledge on inferences was investigated by varying the familiarity of the inferences. In the current section, the effect of the reader’s knowledge on inferences is investigated by varying the knowledge of the readers. The aim is to show that inferences related to the reader’s knowledge, and only those, are made during reading. This is achieved by having two groups of subjects, experts and non-experts in a particular domain, read the same texts in

that domain. Both the experts and the non-experts should be familiar with the concepts that are used in the texts, but only the experts should be familiar with the causal relations between the concepts, since these relations constitute the inferences. We selected the field of economics as the domain of expertise. In two experiments, the knowledge of experts and non-experts was elicited and represented in knowledge networks (Noordman, Vonk, & Simons, 2000; Simons, 1993).

*Knowledge representation of experts and non-experts.*

Participants were either experts or non-experts in economics. The experts were advanced doctoral students in economics. The non-experts were advanced doctoral students in disciplines not related to economics, such as psychology and linguistics. In the first experiment, ninety economic concepts were selected from the most frequent concepts that appeared during one month and a half in a Dutch quality newspaper. The experts and non-experts were asked to freely associate to the economic concepts and to label relations between the concepts. The participants were requested to write down in one minute as many associations to the economic concepts as stimulus words as they could. In addition, they had to describe the relation between the stimulus word and the responses. Free associations were collected because the associations reflect the participant's knowledge representation (Graesser & Clark, 1985). The average number of word associations for experts (5.94) was significantly greater than for non-experts (4.01). In describing the relations between the stimulus words and the responses, experts described significantly more associations (4.48) than non-experts (2.65), and their associations were much more frequently described in terms of causal relations than those of the non-experts.

In the second experiment, 80 different concepts from the previous experiment were put into 100 different pairs. Experts and non-experts judged the familiarity of the 80 concepts on a seven-point scale. Subsequently, their knowledge about the relations between the 100 concept pairs was investigated: They were requested to judge the familiarity of the relations on a seven-point scale, to describe the nature of the relation, and, in particular to indicate whether the relation between the concepts was direct or indirect, whether a bidirectional or unidirectional influence existed between the two economic entities, and whether or not the relation was restricted to specific conditions. The familiarity of the concepts was significantly higher for experts (5.99) than for non-experts (5.39). Similarly, the relations between the concepts were more familiar to the experts (6.16) than to the non-experts (4.18). Seventy of the 100 relations were described as causal by the experts and only 39 by the non-experts.

On the basis of the relations between the concepts, semantic networks were constructed for the experts and the non-experts. Parts of the networks are presented in Figures 2 and 3. The boxes

(nodes) in the networks are the economic concepts. In the corners of the network, the concepts *consumer*, *producer*, *bank*, and *government* are added that give some ordering in the networks. The lines represent relations between the concepts. A relation between two concepts was assumed to exist only if that relation was formulated by at least 75% of the participants. The figures show whether the relations were causal or definitional, whether they were unidirectional or bidirectional, and whether there was a positive influence or a negative influence between the two concepts.

Experts were more familiar with the relations than non-experts. The number of causal relations was much larger for the experts than for the non-experts.

Parts of these networks were isolated as “knowledge schemes” to be used in the reading experiments. A knowledge scheme consisted of three concepts (A, B, and C) and two unidirectional relations between them (A-B; B-C). They can be expressed in one complex sentence, for example, “The increase in *inflation* (A) deteriorates the *competitive position* (B) and this leads to a decrease in the *exports* (C)”. Based on such a knowledge scheme, the following sentence can be formulated: “The exports are decreasing because the inflation weakens the competitive position”. On the basis of the elicited data, 30 knowledge schemes were identified that satisfied the following two criteria: The concepts in the scheme were familiar to both the experts and the non-experts and the relations in that scheme were familiar only to the experts. Forty-one concepts were involved in these knowledge schemes. For the 30 knowledge schemes, the familiarity of the concepts was the same for experts (6.05) as for non-experts (5.95), while the relations were significantly more familiar to the experts (6.22) than to the non-experts (3.91).

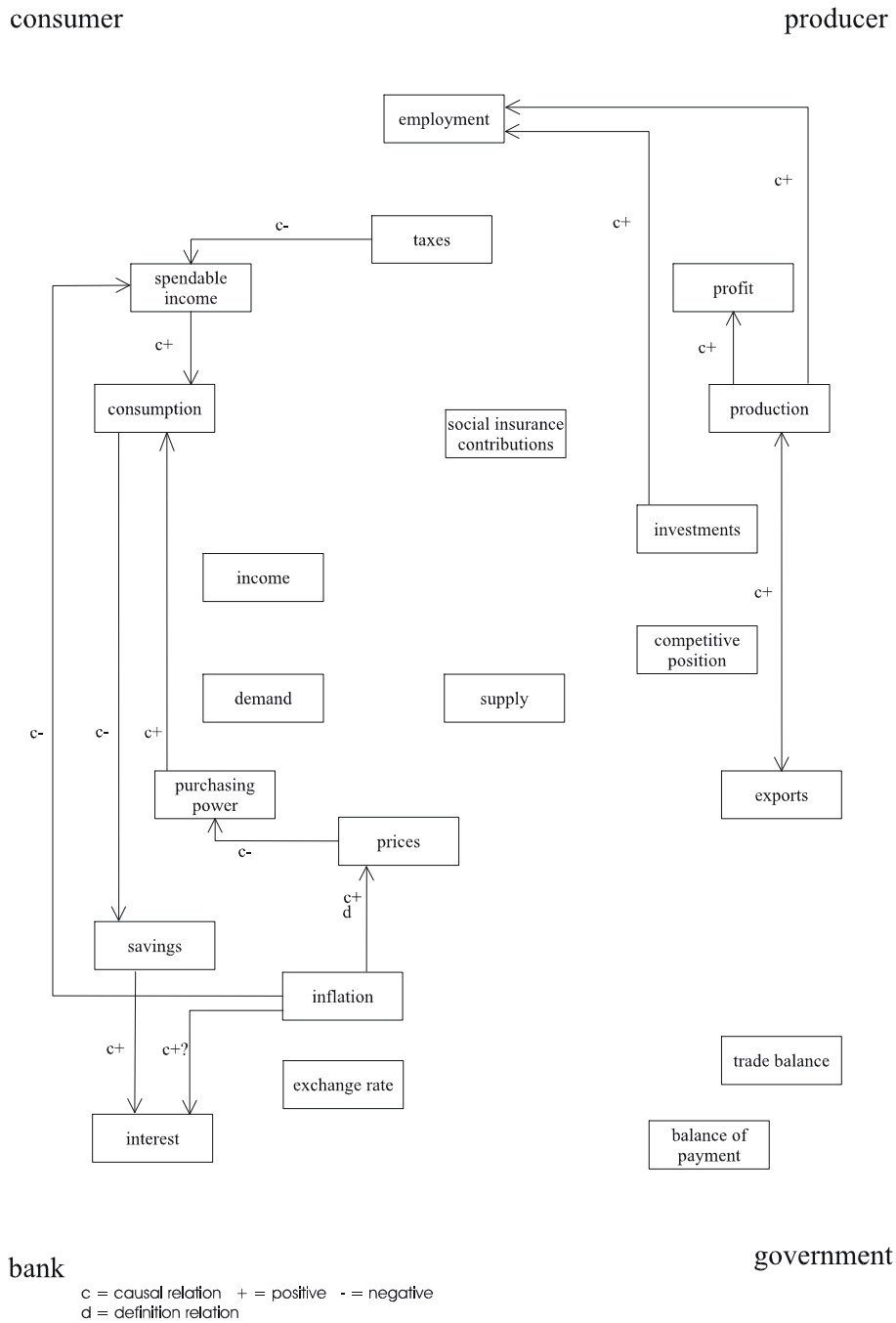


Figure 2. Knowledge representation of non-experts (Noordman et al., 2000, p. 252).<sup>2</sup>

<sup>2</sup> Reprinted from Reinier Cozijn, Edwin Commandeur, Wietske Vonk, and Leo Noordman. The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language*, 2011, 64 (4), 381-403 with permission from Elsevier.

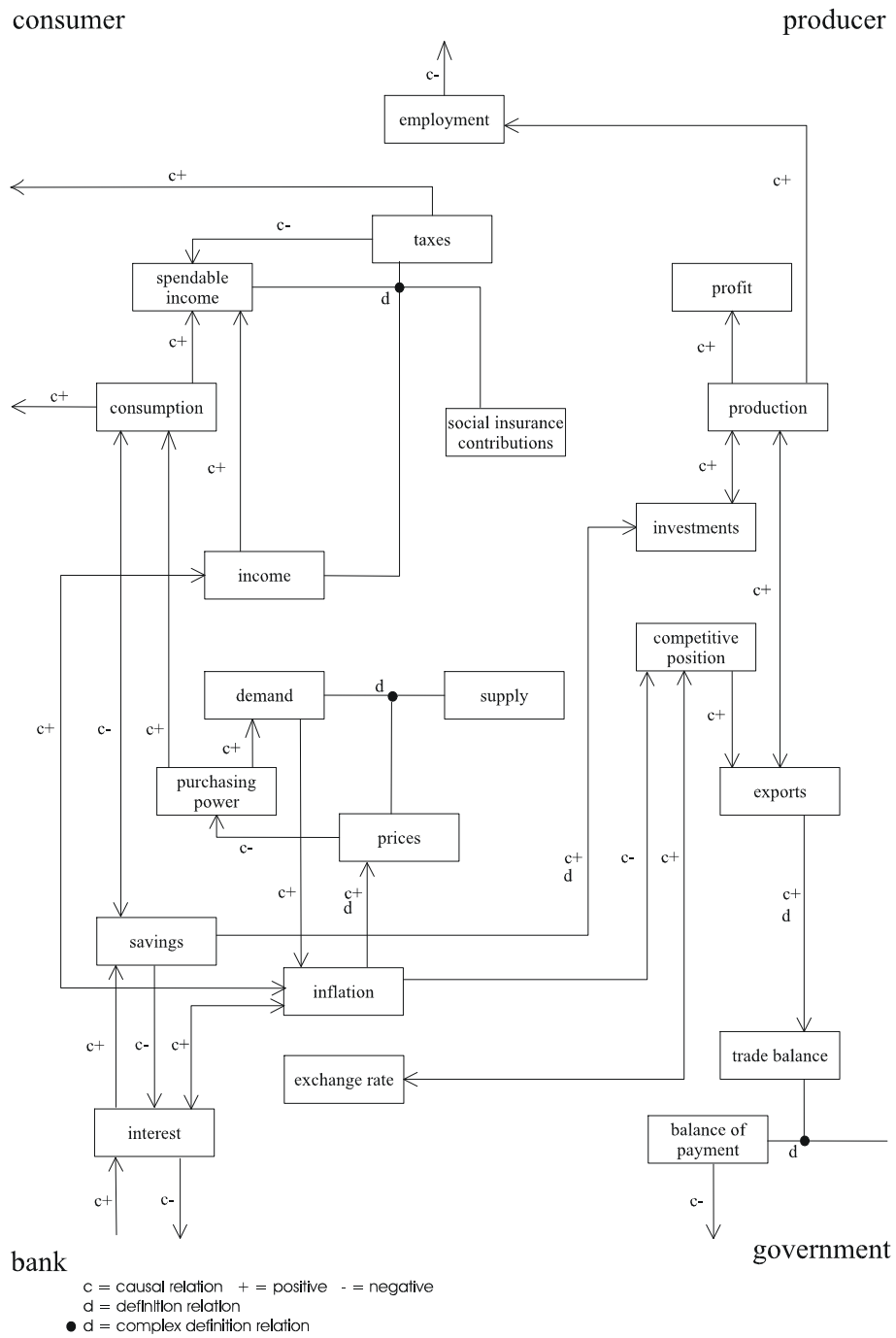


Figure 3. Knowledge representation of experts (Noordman et al., 2000, p. 253).<sup>3</sup>

<sup>3</sup> Reprinted from Noordman, L. G. M., Vonk, W., & Simons, W. H. G. (2000). Knowledge representation in the domain of economics. In L. Lundquist & R. J. Jarvella (Eds.), *Language, text, and knowledge: Mental models of expert communication* (pp. 235-260). Berlin/New York: Mouton de Gruyter with permission from Mouton de Gruyter.

### *Reading by experts and non-experts*

The knowledge schemes of the experts and non-experts formed the basis for the reading materials. The length of the texts was on average nine sentences. An example of a causal sentence based on the earlier mentioned scheme is the sentence: "The American exports have been suffering a decline in the last few months, because rising inflation impairs the competitive position of the U.S.A" in a text on the American economy. This sentence expresses the relation between three economic concepts as it applies to the specific situation described by the text, schematically: "C, because A leads to B". The conjunction *because* triggers the major premise inference that a weakened competitive position decreases the exports. This inference concerns the direction of the causal relation between the concepts (B leads to C) as well as the sign of the relation (there is a positive correlation between B and C). It is assumed that the relation between "competitive position" and "exports" is activated only for the experts, because only for the experts is the relation familiar. This general relation is then applied to the American economic situation described by the text, and becomes part of the text representation of the experts. This constitutes the inference.

The inferences were measured by presenting the *because* sentence in two conditions, similarly to earlier experiments (Simons, 1993). In the explicit condition, the *because* sentence was preceded by a sentence that expresses the causal relation implied by the target sentence (B leads to C). An example is: "Generally speaking, the competitive position of a country has a strong influence on the volume of its exports". In the implicit condition, the explicit information was not given. Experts are expected to make the inference and, therefore, the reading time for the target *because* sentence should be shorter in the explicit condition than in the implicit condition. This difference is predicted to occur during the reading of the *because* clause. It is unlikely that the inference is made during the reading of the main clause. For the non-experts, the relation between the two concepts is not available and, therefore, it is expected that they do not make the inference, neither in the implicit nor in the explicit condition. Consequently, there should be no reading time difference between the two conditions.

In addition to the reading task, a verification task was used. One statement referred to the inferred relation between competitive position and exports. This statement used the concepts of the target sentence. For the example above, this statement was "The impaired competitive position caused a decline of the exports". Because experts are expected to make the inference during reading in the implicit as well as in the explicit condition, there should be no difference in verification times between the explicit and implicit conditions. Non-experts are expected not to make the inference during reading, but at the moment that they have to verify the relation between the two concepts. As in the earlier experiments, the effect of the explicit information for the non-experts should be that the verification time is shorter in the explicit condition than in the implicit condition. For the experts, but



not for the non-experts, the reading times of the target subordinate clause were significantly longer in the implicit condition than in the explicit condition (see Table 3, lines A1 and A2).

Table 3. Mean reading and verification times (ms) in the expert-non-expert experiments (the codes in the first column are for textual reference).

Delay	Reader	Measure	Information			
			Explicit	Implicit		
A1	Expert	Reading time	3027	3449	*	
A2	Non-expert	Reading time	3734	3845		
A3	Expert	Verification time	4438	4901	*	
A4	Non-expert	Verification time	3948	4431	*	
B1	Expert	Reading time	3942	4222	*	
B2	Immediate	Expert	Verification time	4569	4919	*
B3	Delayed	Expert	Verification time	4656	5950	*

\* Significant difference

The difference for the experts was significantly greater than for the non-experts. The verification latencies were shorter in the explicit condition than in the implicit condition for the non-experts, but also for the experts (lines A3 and A4). The reading times support the conclusion that the experts made the inference during reading and the non-experts did not. The verification latencies for the non-experts confirmed this interpretation: They made the inferences only if they were explicitly requested to do so. The verification latencies for the experts showed a non-predicted difference between the explicit and implicit conditions. These data suggest that experts make the inference both during reading and at verification. How can this result be explained? A likely interpretation is that the verification task sets the goal for the experts. They take the verification task as a challenge to exploit their expertise in evaluating statements in their expertise domain. As we have seen earlier, the goal can lead to more thorough processing and to inferencing. This is true for an instructed goal as well as for a self-imposed goal. According to this interpretation, the inference is activated during reading, but is not sufficiently encoded for the verification task. The experts compute the inference again at verification and this is due to their motivation in evaluating a statement in their expertise domain. This motivation-explanation is strengthened by the fact that the verification times for the experts are longer than for the non-experts: The experts become more critical in the verification task. In agreement with this interpretation is the fact that no verification difference was observed in the earlier experiments in which a reading time differences indicated an inference; in those experiments expertise did not play a role.

This interpretation presupposes that inferred information is not robustly encoded and decays over time more so than explicit information. The interpretation reminds us of results found for

predictive inferences: Keefe and Mc Daniel (1993) found decreasing activation for predictive inferences over time; Cook, Limber, and O'Brien (2001) observed that predictive inferences are not robustly encoded so that they have to be computed again. The coding of the inference after reading is no longer sufficiently activated for performing the verification task. This leads to the prediction that the verification time difference between the explicit and implicit condition is greater in delayed verification than in immediate verification.

### *Decay of inferences*

In an experiment similar to the previous one, a verification task was administered immediately after the target sentence was read or after the whole text (Simons, 1993). If the activation of the inference decreases over time, one expects that the difference in verification times between the explicit and implicit conditions increases over time and is greater after reading the text than after reading the target sentence. The reading times for the target subordinate clause in the implicit condition were longer than in the explicit condition (see Table 3, line B1). The verification latencies were significantly longer in the implicit condition than in the explicit condition (see Table 3, lines B2 and B3). Crucially, this difference was greater in the delayed condition than in the immediate condition.

The finding of shorter reading times in the explicit condition replicates the earlier results and confirms the theory that inferences that correspond to available knowledge are made during reading. The significant interaction between implicit versus explicit condition and the moment of verification confirms that the inference is made during reading, that the inferred information is available immediately after reading the target sentence, but that the information after a delay is less available than information that was expressed in the text. Indeed, in the explicit condition there was no difference between the immediate and the delayed verification.

Conclusion: The results of the experiments in this paper are firm evidence for the conclusion that the knowledge of the reader controls the on-line inferences. This conclusion is based on experiments in which the availability of knowledge about the topic has been established in an empirical way. Difference in knowledge was operationalised by manipulating the familiarity of the text materials and by manipulating the expertise of the readers. Next, we turn to the issue of how such effects might be incorporated into a computational cognitive model.

## **Modeling knowledge-based inference**

### *Knowledge and inferences in computational models*

Computational simulations of inference for text comprehension can be highly valuable for uncovering the cognitive processes and representations that underlie the relation between knowledge, reading

goal, and inferences. Because the generation of inferences depends to a large extent on the reader's (or listener's) world knowledge, models of this process should incorporate a knowledge base and be able to apply the subset of knowledge required for understanding the particular text under consideration. The problem here is that people have access to a vast amount of knowledge, any part of which may turn out to be relevant. The human discourse-comprehension system rarely seems to experience any problem finding exactly the required knowledge. In computational modeling, however, this is still an unsolved problem: How can a very large knowledge base be implemented and efficiently accessed in a cognitively plausible manner?

#### *Early approaches*

Several early models of discourse comprehension (e.g., Golden & Rumelhart, 1993; Kintsch, 1988; Van den Broek, Risden, Fletcher, & Thurlow, 1996) dealt with (or, rather, avoided) this problem by including the required piece(s) of knowledge and (if anything) not much more. As argued by Frank, Koppen, Noordman, & Vonk (2008), such an ad hoc approach results in models whose predictive value cannot be properly evaluated. This is because the behavior of such a model is usually demonstrated by applying it to a single text from some experimental study. As a consequence, the outcome is merely anecdotal: There is no set of results whose statistical reliability can be assessed. However, even if the model is tested on a range of texts (from a single experimental study), the inference work for each text is essentially done by the modeler rather than the model. This is because each text requires a different subset of total world knowledge, which needs to be hand-selected before the model can run. In such cases, it remains unclear how the model would fare if it would have to find the knowledge that is relevant for a particular text itself.

An alternative approach arose from early work in Artificial Intelligence: "Frames" and "scripts" (Charniak, 1977; Schank & Abelson, 1977) were conceived to organize human knowledge into discrete packages, each dealing with a particular type of event or situation (e.g., there could be different scripts for "going to a restaurant", "going to the cinema", etc.). The system would then be able to process all stories that only depend on knowledge within the implemented frames and scripts. Although such knowledge structures became much less popular for AI systems in later years, they have been applied more recently as well (Mueller, 2004, 2007). As Kintsch (1988) rightly argues, however, the problem with frames and scripts is that they cannot model both the creativity and precision of discourse comprehension: They either employ too much knowledge and reasoning power (making them too inflexible) or not enough (making them too unconstraining).

### *The “microworld” approach*

The “microworld” method of knowledge implementation (Frank et al., 2008) is similar to the frames-and-scripts approach in that the knowledge base is reduced to a small, well defined domain, and all knowledge of this domain is incorporated in the model. The difference is that frames and scripts select a realistic part of human knowledge and (erroneously) treat it as if it were isolated from all the rest. In contrast, the microworld approach is to select a smaller world rather than a small part of the real world. This yields a small, but unrealistic, knowledge base that can be fully included in a model. That is, rather than assuming that the unimplemented knowledge is irrelevant (as in frames-and-scripts approach), the assumption is that there *is no* unimplemented knowledge.

In the microworld approach, there is no attempt to model a realistic language user and to simulate the processing of realistic texts. This may appear like a serious drawback, but it does away with the problem of frames and scripts because texts that require unimplemented knowledge can never occur in the microworld approach. Consequently, the microworld approach allows one to run many simulations without ad hoc adaptation of the implemented knowledge and, therefore, to statistically evaluate a model’s general predictions. A quantitative comparison between numerical measures of model and human behavior is not possible because the two knowledge bases are incomparable (in size, but most likely also in content). However, as Frank et al. (2008) explained, the same problem plagues models that incorporate only text-specific knowledge. Moreover, *qualitative* comparisons between microworld-model predictions and experimental findings are both possible and valuable. In such a qualitative comparison, a parameter of the model (or of its input) is manipulated in a way that corresponds to the experimental manipulation. To the extent that the effect of the manipulation on the model mirrors that of a human experiment, the model simulation was successful. This is demonstrated, for example, by the Distributed Situation Space model (Frank, Koppen, Noordman, & Vonk, 2003), which will be discussed below.

### *Implementing realistic world knowledge*

Needless to say, the microworld approach to knowledge implementation is not optimal. Ideally, *all* of a reader’s knowledge would be implemented, so that model predictions can directly (and quantitatively) be evaluated against human experimental data. Currently, the only feasible method for implementing realistically-sized knowledge bases is to extract them automatically from large text corpora. Techniques for extracting word meaning from corpora have been around for quite some time; the best known being Latent Semantic Analysis (Landauer & Dumais, 1997) and Hyperspace Analogue to Language (Lund & Burgess, 1996). However, access to lexical semantics clearly does not suffice for discourse comprehension because a *propositional* semantics is needed. There have been

several suggestions for combining word-meaning representations to yield the meaning of phrases, sentence, or propositions (e.g., Grefenstette & Sadrzadeh, 2011; Kintsch, 2001; Mitchell & Lapata, 2010) but it remains to be seen whether such an approach can ever yield knowledge about the meaning of propositions. After all, a proposition's meaning (and in particular its relation to other propositions) cannot usually be derived from the meanings of its predicate and arguments. For example, the fact that traffic jams can cause being late for work is a piece of world knowledge that is not inherent to the meaning of the words *traffic jam*, *work*, and *late*. Hence, we are skeptical about claims that combining word-meaning representations can give rise to the type of world knowledge required for text comprehension (see Frank et al., 2008, for a further discussion of this issue). This is not to say that automatic extraction of propositional knowledge from texts is impossible. In fact, recent work on this problem has yielded promising first results (e.g., Kintsch & Mangalath, 2011; Padó, Crocker, & Keller, 2009), although it did not yet lead to models that simulate knowledge-based causal inferences.

The question remains whether it is even theoretically possible to extract a significant part of a person's world knowledge from corpora. According to Glenberg & Robertson (2000), language comprehension often relies on our embodied experiences with the real world – knowledge that cannot be extracted from mere texts. Hence, knowledge extraction from text corpora may need to be integrated with extraction from data sets in other modalities (Andrews, Frank, & Vigliocco, in press). As such work is still in its infancy, the best available current option seems to be the microworld approach.

#### *The Distributed Situation Space model*

The Distributed Situation Space (DSS) model (Frank et al., 2003) simulates the generation of inferences not as logical reasoning but as the result of pattern matching to knowledge of the microworld. This knowledge is divided into knowledge of possible states of the world and knowledge of causal/temporal relation between states. Because the model makes no distinction between temporal and causal relations, we will simply refer to this part of knowledge as causal. Hence, story situations are states in the microworld and causal knowledge relates the states at consecutive points in story time. The model does not make use of surface form or textbase representations: Everything takes place at the level of Kintsch & Dijk's (1978) situation model. Recent evidence indeed suggests that readers represent causal relations at the situational level only (Mulder & Sanders, 2012).

Formally, the state-of-affairs at one moment in a story is represented as a vector in high-dimensional *situation space*. This space is organized such that it encodes the probabilistic relations between all possible microworld situations. For example, the vector representation of 'Bob plays soccer' also represents the fact that (with very high probability) Bob is outside, because (in this

microworld) soccer is only played outside. Likewise, the same vector represents that it is very unlikely that Bob is inside. Note that these relations indeed occur at the level of the situation model and not in the textbase: They are relations within the microworld, not within a text.

Knowledge about the causal relations between time-consecutive microworld situations is encoded in a matrix that relates all situation-space dimensions to each other. For example, this matrix contains the knowledge that if Bob plays soccer he is likely to be tired soon but unlikely to have been tired just before. Hence, the temporal sequence of story events 'Bob plays soccer (and then) Bob is tired.' is more coherent than 'Bob is tired (and then) Bob plays soccer.'

The DSS model simulates the process of causal inference as follows: First, the situations described by a story text (so far) are encoded as situation-space vectors. Next, these vectors adapt to the causal knowledge matrix, such that they come to fit this knowledge more closely without contradicting the story statements. As a result, the represented sequence of microworld situations gradually changes, increasing the probability of particular states-of-affairs and decreasing the probability of others. This comes down to inferring (to a certain extent) what is (and what is not) the case at each point in story time. This process continues until the vectors' total rate of change drops below a threshold value, set by the *depth-of-processing* parameter. The larger this parameter's value, the lower the threshold and hence the longer the inference process takes to halt. The required amount of time before the threshold is reached serves as the model's prediction of the time needed to comprehend (i.e., read) the current story statement.

As explained above, the microworld approach makes it impossible to directly compare human behavioral data on experimental texts to model predictions on the same texts. Nevertheless, the DSS model has been able to explain a number of important findings from the discourse-comprehension literature. For example, a story statement is read faster if it is more causally coherent with the previous statement (Myers, Shinjo, & Duffy, 1987) and the same relation holds in the model (Frank et al., 2003). Also, an extension to the DSS model for resolving referential ambiguity (Frank, Koppen, Noordman, & Vonk, 2007) correctly predicted, among others, how implicit causality affects reading times and pronoun-resolution error rates, and how this relation depends on processing depth.

In short, the DSS model simulates knowledge-based inferences in discourse comprehension and explains a number of findings from reading experiments. This raises the question of whether the model can also account for some of the results presented earlier in this chapter. As will become clear, the DSS model "as is" does not explain these results. However, it seems likely that its general framework, with relatively minor changes, does allow for simulating the reading experiments and that the results will be consistent with the experimental findings.

### *Simulating the effect of reading goals*

To some extent, the effect of reading goal is captured by the model's depth-of-processing parameter. Increasing its value increases both modeled processing time and amount of inference (Frank et al., 2003). Noordman, Vonk, & Kempff (1992) compared reading for comprehension to reading for consistency checking (which, supposedly, requires deeper processing) and indeed found that the latter condition resulted in slower reading and more inferences (see Table 1, lines A1 and C1).

When participants were asked to read a text with the goal to answer a particular question, this facilitated the drawing of inferences only for facts that were relevant to the question at hand. This question-dependent effect of inference cannot be modeled by varying the one-dimensional depth-of-processing parameter, as it will affect the overall level of inference. What seems to be required is a mechanism for dynamically adapting processing depth to the current content of the text; something which lies far beyond the model's scope.

### *Simulating the effect of the presence of a conjunction*

We have seen that the presence of a conjunction slows down reading at the end of the sentence and reduces off-line verification time (see Table 2), indicating that it facilitates the drawing of an inference. In a sense, it is as if the conjunction signals to the reader that comprehension effort must be temporarily increased. As such, the presence of a conjunction can be simulated in the DSS model in the same manner as the occurrence of deep processing due to a specific reading goal: By increasing the value of the depth-of-processing parameter. The only difference would be that the increase in processing depth would be short-lived, only applying to the statements linked by a conjunction.

### *Simulating the effect of a reader's knowledge*

The differences between non-experts' and experts' causal knowledge can be modeled in a relatively straightforward manner by varying the content of the causal knowledge matrix. Although this has not yet been attempted, the outcome is easy to predict. Recall that the model processes a statement more quickly if it is more coherent with the previous statement. If the text states that "Q because P" and there is indeed a positive causal relation from P to Q encoded in the matrix (i.e., it models an expert), simulated reading time will be shorter than if the causal relation is not present in the matrix (i.e., it models a non-expert). This indeed corresponds to the experimental results (see Table 3, lines A1 and A2).

How about the effect of the presence of an explicit statement about the causal relation between P and Q? For experts, this statement is said to activate the causal knowledge. For non-experts, the explicit statement may lead to new world knowledge. The DSS model was never intended to simulate

the update of world knowledge due to text statements, and “knowledge activation” has no meaning within the model. Nevertheless, we can speculate about how these operations can be captured within the DSS framework.

First, note that in the experiment described under “reading by experts and non-experts”, the explicit statement mentioned that a causal relation from P to Q exists without including the direction of the effect. That is, the statement was ambiguous between “if P then Q” and “if P then not Q”. In either case, the statement would lead to the selection of the situation space vectors representing world states P and Q. The causal knowledge matrix can then be updated such that the relation from P to Q is strengthened.<sup>3</sup> However, the two states are not yet linked in the knowledge matrix of a non-expert and since the direction of the effect is absent from the explicit statement it is unknown whether the connections from P to Q should receive positive or negative values. Hence, the matrix cannot be updated. This problem does not occur in the expert knowledge matrix because P and Q are already linked, be it positively or negatively. The result is that only when simulating an expert does the explicit statement increase the coherence of the text “Q because P”. Indeed, this is what the experiment showed: The explicit condition resulted in faster reading by experts, but not by non-experts (see Table 3, lines A1 and A2).

### Conclusion

The experiments presented in this chapter indicate that inferences about causal relations are made during reading only if the relations are part of the reader’s knowledge: Inferences depend on the availability of knowledge. How can we account for this role of world knowledge? The memory-based approach to discourse processing seems to offer a plausible framework, but a problem for this approach is to specify what available knowledge is. Available knowledge should be specified independently of the reading results, otherwise one makes a circular argument: if certain inferences are made during reading, for example concerning a visit to a restaurant, the explanation is that the restaurant script is available, but if no instrumental inferences are made for stirring coffee, the explanation is that the instrument is not easily available. In the research we discussed in this paper, world knowledge was identified in an empirical way and independently of results of a reading task. World knowledge was manipulated in two ways: by varying the texts and by varying the reader’s knowledge. In the first experiments, the familiarity of the inferences in the texts was manipulated. In the subsequent experiments, the knowledge of the readers was manipulated. The results of both series

---

<sup>3</sup> In much the same way, the matrix’ values are originally learned from many examples of consecutive microworld situations.



of experiments converged: Inference behavior depends on the availability of world knowledge. This can nicely be accommodated in the memory-based processing framework.

Inferences about *unfamiliar* causal relations are not made during reading. This is the case even if the presence of a conjunction signals the inference to be made. However, these inferences are made if required by the reader's goal. Two types of goal have been investigated: goals defined in terms of a specific topic that the reader should pay attention to and goals defined in terms of the demands of the task, for example, a question-answering task or a verification task. The explanation of the effect of topic-goal is not that the goal activated the causal relations that have to be inferred, because these relations were not familiar. The explanation of the effect of both kinds of goal is that they urge the reader to process the information in the texts more deeply. Inferences about *familiar* causal relations are made during reading. But making these inferences requires the presence of a causal conjunction.

The inference behavior of experts and non-experts confirms these results. We first investigated the knowledge representation of experts and non-experts in economics and identified relations between economic concepts that were part of the knowledge representation of experts but not of non-experts. Experts made the inferences about these causal relations during reading, while non-experts did not. Again, the non-experts did make the inference when they were requested to do so in a verification task. Interestingly, the experts made the inference both during reading and during the verification task after reading. This means that experts, when asked to verify a statement in their expertise domain, are extra motivated to thoroughly process the information and to make the inference. The effect of the explicit information, diagnostic for an inference, was stronger in delayed verification than in immediate verification. This suggests that inferred information decays over time. Can we account for these results in a computational model? The answer is no, because we don't have an implementation of the economic knowledge of experts and non-experts. However, our Distributed Situation Space model can simulate knowledge-based inferences and explain a number of phenomena that we observed in the experiments. The effects of the reading goal, defined in terms of instructed task, can be explained in terms of the depth of processing parameter. Increasing its value leads to slower reading and more inferencing. The same is true for the effect of the conjunction, although there is no mechanism yet for restricting the effect to the clauses connected by the conjunction. The fact that the explicit information speeds up the reading time for the target sentence for experts and not for non-experts can be accommodated in the DSS framework by the fact that the relation between two situations in the model is strengthened for experts and not for the non-experts, since the two states are not connected for the non-experts. This same difference in knowledge representation in the model can explain the shorter reading time for a *because* sentence by the experts than by the non-experts.

The present chapter highlights the importance of the reader's knowledge for inference processes and illustrates how current computational models fall short in this respect. Although the explanation of the experimental findings described here requires a further elaboration of the DSS model so as to include a very large knowledge base that can be accessed in a cognitive plausible way, the model can in principle simulate causal inference processes and explain a number of findings of discourse processing studies.

## References

- Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1061-1070.
- Andrews, M., Frank, S. L., & Vigliocco, G. (in press). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science*.
- Charniak, E. (1977). A framed PAINTING: The representation of a common sense knowledge fragment. *Cognitive Science*, *1*, 355-394.
- Cook, A. E., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language*, *44*, 220-234.
- Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. M. (2011). The time course of the use of implicit causality information in the processing of pronouns: A Visual World Paradigm study. *Journal of Memory and Language*, *64*(4), 381-403.
- Cozijn, R., Noordman, L. G. M., & Vonk, W. (2011). Propositional integration and world-knowledge inference: Processes in understanding *because* sentences. *Discourse Processes*, *48* (7), 474-500.
- Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2003). Modeling knowledge-based inferences in story comprehension. *Cognitive Science*, *27*, 875-910.
- Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2007). Coherence-driven resolution of referential ambiguity: A computational model. *Memory & Cognition*, *35*, 1307-1322.
- Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2008). World knowledge in computational models of discourse comprehension. *Discourse Processes*, *45*, 429-463.
- Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes*, *26*, 67-86.
- Gerrig, R. J., & O'Brien, E. J. (2005). The scope of memory-based processing. *Discourse Processes*, *39*, 225-242.
- Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43*, 379-401.
- Golden, R. M., & Rumelhart, D. E. (1993). A parallel distributed processing model of story comprehension and recall. *Discourse Processes*, *16*, 203-237.
- Graesser, A. C., & Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood: Ablex.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395.

- Grefenstette, E., & Sadrzadeh, M. (2011). Experimental support for a categorical compositional distributional model of meaning. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 1394–1404). Edinburgh, UK: Association for Computational Linguistics.
- Kant, E. (2004). *Kritiek van de zuivere rede* (J. Veenbaas & W. Visser, Trans.). (Original work, *Kritik der reinen Vernunft* [Critique of pure reason], published in 1781). Amsterdam: Boom.
- Keefe, D. E., & McDaniel, M. A. (1993). The time course and durability of predictive inferences. *Journal of Memory and Language*, 32, 446–463.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163–182.
- Kintsch, W. (2001). Predication. *Cognitive Science*, 25, 173–202.
- Kintsch, W., & Mangalath, P. (2011). The construction of meaning. *Topics in Cognitive Science*, 3, 346–370.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394.
- 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, 211–240.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Lund, K. & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, 28, 203-208.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440–466.
- Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science*, 34, 1388–1429.
- Mueller, E. T. (2004). Understanding script-based stories using common sense reasoning. *Cognitive Systems Research*, 5, 307–340.
- Mueller, E. T. (2007). Modelling space and time in narratives about restaurants. *Literary and Linguistic Computing*, 22, 67–84.
- Mulder, G., & Sanders, T. J. M. (2012). Causal coherence relations and levels of discourse representation. *Discourse Processes*, 49, 501–522.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, 26, 131–157.
- Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 876-886.
- Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, 26, 453–465.
- Noordman, L.G.M. (2005). *Taal en Kennis*. [Language and Cognition]. Valedictory lecture. Tilburg University.
- Noordman, L. G. M., & Vonk, W. (1992). Readers' knowledge and the control of inferences in reading. *Language and Cognitive Processes*, 7, 373–391.
- Noordman, L. G. M., & Vonk, W. (in press). Inferences in discourse, psychology of. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd edition). Amsterdam: Elsevier.

- Noordman, L. G. M., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language*, 31, 573–590.
- Noordman, L. G. M., Vonk, W., & Simons, W. H. G. (2000). Knowledge representation in the domain of economics. In L. Lundquist & R. J. Jarvella (Eds.), *Language, text, and knowledge: Mental models of expert communication* (pp. 235-260). Berlin/New York: Mouton de Gruyter.
- O'Brien, E. J. (1995). Automatic components of discourse comprehension. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of Coherence in Reading*, (pp. 159-176). Hillsdale, N. J.: Erlbaum.
- Padó, U., Crocker, M. W., & Keller, F. (2009). A probabilistic model of semantic plausibility in sentence processing. *Cognitive Science*, 33, 794–838.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.
- Simons, W. H. G. (1993). *De regulering van inferenties door de kennis van de lezer* [The control of inferences by the knowledge of the reader]. Doctoral dissertation, University of Nijmegen.
- Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479-516). San Diego, CA: Academic Press.
- Singer, M. (2007). Inference processes in discourse comprehension. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp.343-359). Oxford: Oxford University Press.
- Singer, M., & O'Connell, G. (2003). Robust inference processes in expository text comprehension. *European Journal of Cognitive Psychology*, 15, 607-631.
- Van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539-588). San Diego, CA: Academic Press.
- Van den Broek, P., Risdén, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Hillsdale, NJ: Erlbaum.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vonk, W., & Noordman, L. G. M. (1990). On the control of inferences in text understanding. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 447-464). Hillsdale, NJ: Erlbaum.
- Vonk, W., & Noordman, L. G. M. (2001). Inferences in discourse, psychology of. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 7427- 7432). Amsterdam: Elsevier.