THE ACCESS TO CAUSAL RELATIONS IN SEMANTIC MEMORY

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultäten

der Georg-August-Universität zu Göttingen

vorgelegt von

Daniela Bettina Sellner

aus Griesbach

Göttingen 2002

D 7
Referent: Prof. Dr. M. Waldmann
Korreferent: Prof. Dr. G. Lüer

Tag der mündlichen Prüfung: 29. Oktober 2002

Acknowledgements

1.   Abstract

In my dissertation I investigated the dissociation between the **causal model view** and the **associationistic view** of the representation of causal relations in semantic memory. The **associationistic view** postulates that causal relations as well as other relations are represented as association in the human mind. The **causal model view** on the other hand claims that causation and other relations are separately represented including all their specific features, especially causal directionality.

I empirically tested these two views in six experiments where the retrieval process from semantic memory was manipulated. In the first three experiments I investigated the *retrieval asymmetry* for causal relations. I presented participants with pairs of words one after another either describing events that referred to a cause (e.g., spark) or an effect (e.g., fire). I manipulated the temporal order in which the words were presented, and the question participants had to respond. Moreover, the causal word pairs had a symmetric strength of association for the predictive (cause-effect) and for the diagnostic (effect-cause) causal direction to avoid a retrieval asymmetry based on associative asymmetry. The results show that questions referring to the existence of a causal relation are answered faster if the first word referred to a cause and the second word to its effect than vice versa. Then I extended the investigation of the causal retrieval asymmetry by testing a *task specific access* to causal relations. That is, the same causal relations had to be accessed via causation or via association. No asymmetry was observed if an association had to be determined between the causal word pair, but the afore mentioned asymmetry pattern was found for the causal access, indicating that participants were capable of accessing different types of relational knowledge in a task specific fashion.

The distinction between the **causal model** and the **associationistic view** was further investigated in an exploratory *function magnetic resonance imaging study* (fMRI), where first evidence could be found that the task specific access to causal relations not only yields a difference in the behavioral data, but also exhibits different patterns of cerebral blood flow.

The results from all experiments indicate that causation is stored as a separate relation in semantic memory including its directionality and it can be accessed as such. This view is postulated by the **causal model theory**.

2.   Introduction

In daily life we are confronted with a vast number of facts, events, and things. We know how things function, we act and interact with them, and we verbally relate to them. Information about facts and events, their relationships to one another, their order of appearance, and their potential explanation and even prediction, is helpful for our survival. We know that lightning can harm us and therefore seek shelter, when we see a storm approaching. When we hear the thunder rolling we can conclude that lightning has struck somewhere, because lightning is accompanied by the sound of thunder. Given this ability it is plausible that these facts, events, and things including all their features, functions, consequences, and possible interconnections must somehow be represented and stored in our long-term memory. The memory system regarded responsible for storage of knowledge of that kind is referred to as *semantic memory*. Accordingly, semantic memory is the database from which we can retrieve the knowledge concerning the specific relationship between lightning and thunder. When I hear thunder I can activate the relation between thunder and lightning in my semantic memory and act appropriately (i.e., seek shelter).

The relation between lightning and thunder has an associative characteristic: Lightning is accompanied by thunder. The associative relation is, however, not the only relation connecting these two events, as the relation also has a causal characteristic: Lightning is the cause of thunder. Consequently, different kinds of relations need to be represented in long-term memory.

In my dissertation I will investigate:

1. How the specific relationship of *causality* is represented in the semantic memory system and

2. how different relationships (associative vs. causal) that connect the same events are retrieved from semantic memory.

3.   Theoretical background of causality and semantic memory

"Any particular event that we might wish to explain stands at the end of a long and complicated causal history" (David Lewis, 1993, p.182).

There is a longstanding debate in philosophy whether the world can be explained via causality or if non-causal explanations are possible (Hempel, 1965). David Lewis (1993) claims that the explanation of an event is accomplished by gaining information about its causal history. According to his proposal, every fact, event, and thing, existing in the world possesses its own causal history and can therefore be explained via that history. Thus, causation is not only capable of explaining the world it is indeed the very fabric that holds the world together.

Given these claims the representation of the concept of causality in the human mind becomes inevitable. However, before I can start investigating the representation of causality in long-term memory, I need to describe the concept of causality as it is discussed in philosophy and psychology. This is important, because the specific features of causality demand a corresponding type of storage and retrieval.

This chapter is divided in three sections: *Causality*, *Semantic memory* and *Causal semantic memory*. The first section provides an introduction to the concept of causality, especially causal asymmetry, as seen from philosophical and psychological perspectives. Additionally the acquisition of causal knowledge within an associative and a causal relational framework is discussed. In the second section, the integration of causal relation and causal asymmetry in semantic memory models is investigated. And finally in the third section, the retrieval processes of causal relations from semantic memory are differentiated for the **associative** and the **causal** relational **view**.

3.1   Causality

Causality and causation has been of great interest in philosophy and psychology. The manner in which it is viewed, however, differs greatly between the two. Philosophy tries to answer the question what causality really is, whereas psychology deals with how humans perceive causation and utilize it in making judgments, predictions, attributions, etc. Nevertheless, these two domains cannot be regarded separately. Causality has unique features, which the two different types of inquiry have to address. One feature, which is important for my investigation, is causal asymmetry. It is discussed in section 3.1.2 for both domains, philosophy and psychology.

3.1.1   Ideas about Causation

Causation in Philosophy

The concept of causation in philosophy goes back as far as Aristotle and continued via Ockham and Hume to modern Philosophers like Mackie and Suppes, to name a few.

The concepts of cause and of causal relation have been regarded under different perspectives over time. Causal relations have been viewed as a connections between two events possessing certain properties: They underlie regularities and necessities (i.e., temporal, spatial contingencies, constant conjunction), they can be deterministic (A always causes B) or probabilistic (the appearance of A increases the likelihood with which B will occur), and they can be viewed as productive relations (A has the power to cause B). Causation has also been regarded as a process that changes the state of something (Salmon, 1984).

Regularities and necessities

Aristotle (1929) considered a cause to be something particular, such as an oak tree. He differentiated four different types of causes that are responsible for the existence of any given entity: 1) Material cause: the material constitution of the entity (wood), 2) formal cause: the contribution to the being of a thing of its form or shape (shape of an oak tree), 3) efficient cause: effects of antecedent events (lightning causes fire to the tree), and 4) final cause: the end, the reason, the purpose of that being (providing shade). Over time, however, the concept of causation became increasingly narrow and currently only the efficient cause

is utilized for scientific explanation. Aristotle also proposed that causes could be definitely known and that causal relations were relations of necessity.

The concept of relation of necessity was rejected by William of Ockham (1980). He is considered one of the representatives of the medieval, theological view of causation, which regards God as the cause to all effects. This statement implies that there cannot be a necessary connection between contingent events, because God always has the power to intervene. According to Ockham, one cannot distinguish between causes coming from God's interference or other physical causes. Thus, causal relations cannot be known with certainty.

Later, David Hume (1739) defined causation as a construct of the mind. A causal relation in itself cannot be perceived. The regularity between two events, their spatial and temporal contiguity, is perceivable, but knowledge about causality is superimposed by the mind. Hume further argued that the construction of causality is possible when the regularity between two events has been experienced more than once. Therefore, we would be unable to identify the relation between two events as causal if exposed to them only once. People have the idea of causation as some power connecting the events. But this idea is formed by the mind from the repeated occurrence of the conjunction of the two objects.

Other philosophers have viewed causal relations from a different point of view. Mackie (1965, 1974) regarded a cause as a condition in which an effect occurs. In Mackie's theory a cause is an **I**nsufficient but **N**ecessary part of a scenario that is **U**nnecessary but **S**ufficient for the effect to occur (INUS). Take a fever caused by a bacterial infection for example. Bacteria belong to an implicit scenario that may include a wound and unhygienic environment. The scenario is not a necessary cause of fever, because fevers can be caused by other things, too. The scenario could be sufficient to cause fever. Bacteria per se are, however, not sufficient to cause fever; other conditions such as a weak immune system are required. Nevertheless, bacterial infection is a necessary part of the scenario. Accordingly, the bacterial infection is an INUS condition for the fever to occur. Mackie (1965, 1974) placed his concept of causation within a causal field, in which the number of possible causes is limited. Within a causal field the effect sometimes occurs and sometimes does not. To analyze a causal relation one has to compare the differences between the times in which the effect occurred and at times it did not within the causal field.

Probabilistic relationship

Only a small number of causal relationships are deterministic relationships (i.e., every time the cause is present the effect can be observed). Accordingly philosophical theories also view causation in probabilistic terms (i.e., not necessarily every time when the cause is present the effect occurs). Thus the relationship between cause and effect is characterized using tools of probabilistic theory. The idea is to develop a theory of causation that does not presuppose physical determinism. The central idea of a probabilistic approach to causality is that causes raise the probability of their effects, everything else being equal. An event C is the cause of event E if and only if $P(E|C) > P(E|\neg C)$. The probability for an effect to occur is higher if the cause is present than if it is absent. This definition of causality is, however, plagued by the difficulty of differentiating between "genuine" and "spurious" correlations. For example, if C and E are both caused by a third event F, then it possible that $P(E|C) > P(E|\neg C)$ even though C does not cause E. Therefore it is necessary to distinguish between "genuine" and "spurious" causes. For example, both the probability of lung cancer and the probability of a person having yellow-stained fingers is higher among the group of smokers than in the general population. Because of the correlation between yellow-stained fingers and cancer, yellow-stained fingers could be taken as a "spurious" cause for lung cancer, even though the factor smoking is the "genuine" cause. In the example, smoking is the third event (F) that factors out the probability relationship between lung cancer and yellow-stained fingers (Hitchcock, 1997). Reichenbach (1956) addressed this type of probabilistic relationships by introducing the term "screening off". If $P(E|C \wedge F) = P(E|F)$, then F is regarded as screening off C from E. According to Reichenbach (1956) a "no screening off"-assumption needs to be added to the probability raising condition: Factor C occurring at time t, is a cause of the later factor E if and only if:

$P(E|C) > P(E|\neg C)$ and there is no factor F, occurring earlier than or simultaneously with C that screens off C from E.

Causation as action and manipulation

Causal relationships are considered to be exploitable for the purpose of manipulation and control. For example, if C is genuinely a cause of E, and I am able to manipulate C, I should via C be able to manipulate or change E (von Wright, 1971). According to Menzies and Price (1993, p. 187) "An event A is a cause of a distinct event B just in case bringing about the occurrence of A would be an effective means by which a free agent could bring

about the occurrence of B". These so called agency theories face their limits explaining causation under circumstances where human manipulation is not possible. For these cases they argue that humans rely on analogies between situations where a human manipulation is possible and situations where it is not, to extend their causal judgment.

Causal processes

Salmon (1984) treats causality primarily as a characteristic of continuous processes rather than as a relation between events. Processes have greater temporal duration and spatial extension than events. "A baseball colliding with a window could count as an event; the baseball traveling from the bat to the window, would constitute a process" (p. 139). Causal processes are processes that are able to transmit information, which he calls a mark. He differentiates between two concepts of causation: causal production and causal propagation.

An example of a causal production would be the following: If a light beam consists of white light (or a suitably restricted set of frequencies), we can put a filter in front of the light source, separating out only the red frequencies. The light beam after it passes through the filter will bear the "mark" of having done so: it will now be red in color. Contrast this with the case of the light spot traveling along the wall: if we put a red filter at one point in the process, the spot will turn red for just that moment and then carry on as if nothing had happened. Interfering with the traveling process will leave no "mark".

Causal propagation is given as a conditional definition in terms of the interaction between two causal processes. For example, the baseball that hits the window is considered to be a causal process and the persistence of the window is considered to be a causal process. Thus, the collision of the ball and the window is an interaction of the two causal processes.

To summarize, causation has been regarded under different aspects in philosophy. It has been characterized with different processes and properties such as regularities, the manipulative influence of causes on their effects, and probabilistic relations.

Causation in Psychology

On the other hand, psychologists are interested in the way people understand and perceive causation, how they make causal inferences, judgments, and attributions, and how we acquire and store causal knowledge.

There are different approaches to investigating human's ability to derive causal judgments and learn new causal relations. In this section I will focus on *associative*, *contingency*, and *causal power models*. These models focus primarily on learning situations and causal judgment, not the perception of causation itself. In psychology, perception of causal relations refers to the question how can causality be inferred for two connected events, if observed only once (Michotte, 1963, White, 1988). In causal perception single case attributions are of interest. In contrast, causal learning is accomplished by numerous observations and is therefore the relevant concept for my investigation. The following causal learning models make different assumptions about how causal knowledge is acquired and, more importantly, how it is used to derive causal judgments.

Without the acquisition of causal knowledge, causality could never be stored in semantic memory and subsequently could not be retrieved from it. Given these different theories of causal learning, it is not surprising that the assumptions regarding storage and retrieval of causal knowledge differ between the various models.

Causation in an associative model

In the associationistic view causation is inferred from cues to causality, such as the regular succession of the effect, the temporal contiguity of cause and effect, and the spatial contiguity of cause and effect. These characteristics have also been postulated in (animal) learning theories (Hull, 1943; Rescorla & Wagner, 1972). The association between two events is strengthened when they are contiguous and is weakened when an event occurs by itself. Thus, the learning of cause-effect relationships solely relies on regularities.

It is also noteworthy that the terminology used in these models differs from the terminology used by philosophers. The words "cause" and "effect" have disappeared and the words "cues" and "outcomes" are introduced instead. This change in language is due to the behaviorist background assumptions (Pavlov, 1927; Thorndike, 1911), which postulate that organisms respond to stimuli regardless of the type of event the stimuli represent.

One of the most prominent exemplars of associative theory is the *Rescorla-Wagner model*. The two researchers Rescorla and Wagner specified in 1972 the development of

strength of association between two events over time on the basis of contiguity and contingency. In the following example of the model (see Figure 1) there are two input nodes (A and B) that are explicit cues and one output node (outcome). The individual has to learn that the position of two switches has an influence on the light of a bulb. There are several trials in the experiment in which the individual receives information about the condition of the switches and the light bulb. The associative strength of switch A and switch B with the light bulb is calculated with the following equation:

$$\Delta V_i = \alpha_i \beta_{outcome}(\lambda - \Sigma V)$$

where ΔV_i is the change in associative strength of the cues (switch A and B), α_i is a learning rate parameter (how fast the knowledge is acquired) that is unique to each cue and roughly represents the salience of that cue, $\beta_{outcome}$ is the learning rate parameter for the outcome variables, λ is the maximum associative strength that the outcome will support (i.e., 0 if the light is out and 1 if the light is on), and ΣV is the sum of the associative strengths of all cues present in a given trial.
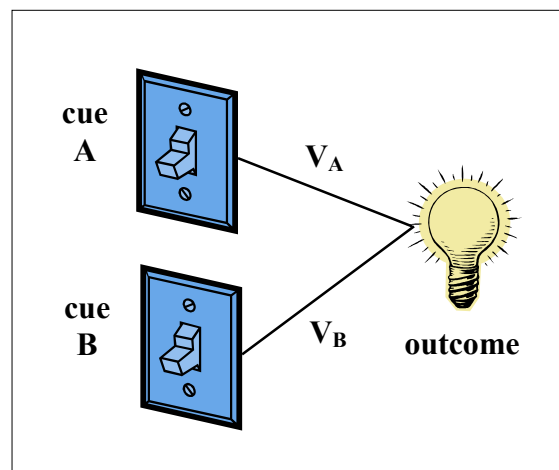
Figure 1: Associative learning model for the association of two switches and a light bulb. V_A refers to the associative strength between cue A and the outcome; V_B refers to the associative strength between cue B and the outcome.

If a cue is not present, its associative strength is not affected. Thus, if a switch is off in a particular trial it cannot gain associative strength during that trial. Accordingly, if one switch is always on and the other is always off, the latter one cannot develop any associative strength.

Given the associative character of causal knowledge acquisition in the Rescorla-Wagner model, information about causation would accordingly have to be represented in an associative fashion in the human mind.

Causation in a contingency model

The contingency model theory argues that humans, in order to infer causality, compute the covariation between cause and effect and use this information as part of a causal model (Cheng & Novick, 1992). This means that the absence of the cause is also taken into account. Table 1 depicts a contingency table for four events: absence and presence of the cause, and absence and presence of the effect. The cells represent the frequencies of the four possible conjunctions (e.g., A represents the number of times the cause and the effect is present.).

Table 1:

A 2x2 contingency matrix of a binary cause-effect condition.

	Effect	**No Effect**
Cause	A	B
No Cause	C	D

Given these frequencies, the covariation of cause and effect can be computed. The corresponding statistic is called Δp (Allen 1980). It is defined as the difference between the conditional probability of an effect given the presence of a cause and the conditional probability given the absence of the cause:

$$\Delta p = p(E|C) - p(E|\neg C) \qquad \text{(Definition of contingency)}^{[1]}$$

Therefore, an individual trying to determine whether one event is the cause of another event, would compare the corresponding conditional probabilities to derive the judgment. If the value of Δp is larger then 0, the cause is generative (e.g., too much alcohol can cause headache), if the value is below 0, then the cause is preventive (e.g., vaccination can prevent illness).

---

[1] Δp takes probabilistic causal relations into account, which were discussed by many Philosophers (Hitchcock, 1997; Reichenbach, 1956).

A problem of contingency models, like the one above, is that mere covariation does not have to imply causation. Events that covary are not necessarily connected via a causal relation. For example, though sunrise follows the crow of a rooster, the crowing does not cause the sun to rise. Thus covariation of two events might be a cue to causality, but there are additional factors that influence the judgment of a relation being causal. Cheng and Novick (1990) proposed a *probabilistic contrast model* as a generalized contingency model. This model was used to provide a descriptive account of the use of statistical regularity in natural causal induction. Additionally, the notion of a focal set was introduced. A focal set constitutes predisposed conditions under which a causal relation needs to be examined. In a focal set only events are included that might contribute to the explanation of a cause-effect relationship. Thus, a potential cause is evaluated by its contingency computed over that focal set. The size of the focal set depends on the reasoner's knowledge. For example, a focal set for the explanation of lung cancer could include smokers, and people who live in an industrial area. Cheng (1997) developed the probabilistic contrast model further into the *causal power model*.

Causation in a causal power model

The causal power model assumes that there are things that have the power to produce other things or to prevent other things from happening. They have the causal power to influence the occurrence of the effect. Therefore, causal power is a theoretical construct used to explain observed connections. This construct is derived from observation. According to Cheng (1997) and others (Buehner & Cheng, 1997; Wu & Cheng, 1999) causal power is defined as the probability of a cause to produce the effect if no other (alternative) cause is present. Unfortunately, in real life, alternative causes are the rule and not the exception. Therefore, the causal power of an event can only be estimated, but never fully determined. In order to estimate the conditional probability of a given (candidate) cause C, all alternative causes A have to be taken into account. If one assumes that all causes Cs and As produce the effect E, then the alternative causes can be viewed as having a generative character. If the candidate cause C and the alternative causes A prevent the effect, then they can be viewed as having an inhibitory character. Causal power also makes the additional assumption that the influence of Cs and As on the effect is independent from one another.

Thus, the probability of the effect E to occur given the presence of cause C is the sum of the causal power of the candidate cause and the causal power of the present alternative causes minus the intersection of the influence of both types of causes.

Table 2:

Causal power calculation for generative and preventive causes.

Causal Power Calculation[2]	
Generative cause	*Preventive cause*
$$P_c = \dfrac{\Delta p}{1 - P(E|\neg C)}$$	$$P_c = \dfrac{-\Delta p}{P(E|\neg C)}$$

As can be seen in Table 2, causal power is a measure of the effectiveness of a cause that relies on the contingency, but takes the influence of additional causes into account and corrects for them. The idea being that if an effect frequently occurs in association with alternative causes, the candidate cause can rarely show its influence. The influence of the candidate cause can only be observed if the alternative causes do not occur.

Causal power theory is one of the most influential approaches to explaining human causal judgments. It not only relies on observation, but also draws on people's prior knowledge about the causal texture of the world. Nevertheless, causal power theory is quite vague on the assumptions it makes about prior causal knowledge, especially on how causal direction might have an explicit influence on causal learning.

So far, the psychological approach to causality was discussed for causal learning and judgment. However, causation is an important relation, which allows us to manipulate things, to predict and explain effects, etc. and therefore representing it is useful and necessary. The introduced models make no assumptions about how causality is represented in long-term memory, whether by association or, more specifically, by covariation or contingency.

---

[2] An exact derivation of the formula can be found in Cheng (1997, p. 373 f)

3.1.2 Causal asymmetries

Causation has one important feature that distinguishes it from other types of relations, its asymmetry. Causality is direction dependent. This characteristic feature has given rise to philosophical explanation accounts and to psychological investigations, especially on the acquisition of causal knowledge.

Causal Asymmetries in Philosophy

According to Hume (1739) the perception of a connection between two events as a causal connection (relation) underlies three assumptions.

- *Contiguity*: Cause and effect have to be contiguous in space and time.
- *Priority of time*: The cause must be prior to the effect.
- *Necessary and Constant Conjunction*: There has to be a constant union between the cause and the effect.

Asymmetry is one of the prominent features of causation and it could be attributed to the following characteristics of causation:

Temporal priority can be regarded as an asymmetric feature: Causes precede their effects. A glance at causal chains (i.e., a chain of causally linked events) supports the view of temporal asymmetry in Hume's approach. However, there are cases where the cause remains present during occurrence of the effect or continues on even long after the effect has occurred. Therefore, the concept of temporal priority refers to the onset of the cause preceding the onset of the effect. What if cause and effect appear simultaneously? For example, two playing cards are leaned against each other so that they support each other: The right card holds the left card and vice versa. The position of each card causally depends on the position of the other. As can be seen, there are exceptions to this claim of temporal priority that make it difficult to explain asymmetry in causation on grounds of temporal asymmetries. Hence, there should be more to asymmetry of causation than only the temporal asymmetry. Otherwise causation could be reduced to temporal asymmetry.

Probabilistic independence states that while causes of a common effect are probabilistically independent of one another (e.g., a virus and/or a bacterial infection can

cause fever independently from each other), the effects of a common cause are probabilistically dependent on one another (e.g., a flu can cause cough and a stuffy nose).

For example, to determine the causal structure in a scenario including three events where no information is given about the temporal order of the occurrence of the events, and their connection, the correlation for each pair of events needs to be calculated. It turns out that in a common-cause structure with three events such as flu, cough, and stuffy nose, the events cough and stuffy nose are unconditionally dependent and conditionally independent of each other if the third event flu is held constant. Accordingly, if I have the flu and a cough, it is very likely that I will also have a stuffy nose (Figure 2). In contrast, in a common-effect structure the two causes are unconditionally independent and conditionally dependent, if I hold the third event, which is the effect, constant. For example, if I have the flu and I know that I was exposed to a virus, then it is rather unlikely that bacteria were also involved causing the outbreak of the flu (Figure 2).
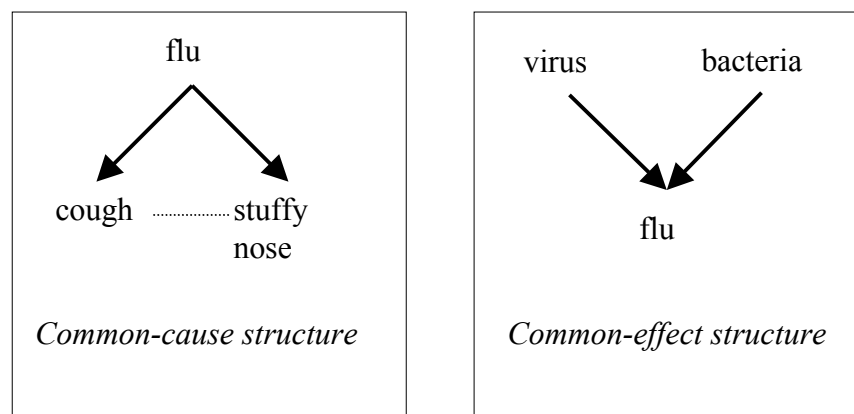
Figure 2: Causal asymmetry determined by the causal structure. Left: common cause; right: common effect.

Another way of dealing with this causal asymmetry in probabilistic relationships comes from Reichenbach (1956). He also postulated that this causal asymmetry is based on probabilistic relations and cannot be reduced to a temporal priority issue. Consider two factors E and F which are positively correlated:

$P(E \wedge F) > P(E) \cdot P(F)$

and a factor C with the following properties:

1. $P(E \wedge F|C) = P(E|C) \cdot P(F|C)$
2. $P(E \wedge F|\neg C) = P(E|\neg C) \cdot P(F|\neg C)$
3. $P(E|C) > P(E|\neg C)$
4. $P(F|C) > P(F|\neg C)$

These formulas describe a common-cause situation. Reichenbach (1956) now argues that if event C occurs prior to E and F and there is no event which satisfies the same conditions and occurs after E and F, then CEF form a *conjunctive fork* open to the future. If C occurs after E and F and satisfies the condition the fork is said to be open to the past (Figure 3). He claims that the direction from cause to effect is the direction in which open forks predominate and in our world forks are predominantly open to the future.
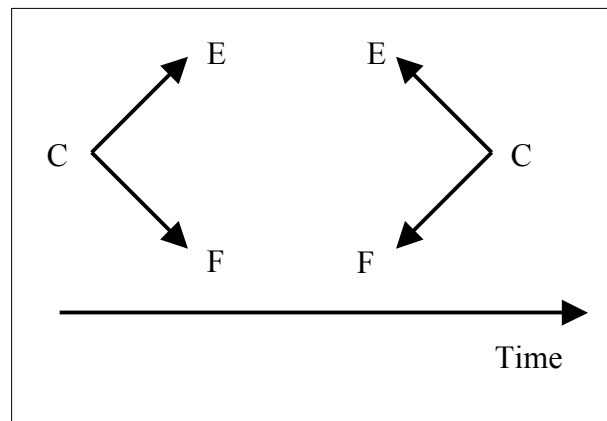
Figure 3: Left: fork open to the future; Right: fork open to the past.

A third asymmetry found in causation is the *asymmetry of action and manipulation*. Causes can be used to manipulate their effects, but effects cannot be used to manipulate their causes. Events C and E are causally related, if manipulation of C causes changes in E. In research, for example, we manipulate the independent variable in order to find different expressions of the dependent variable in due to that manipulation. A manipulation of the effect, however, has no influence whatsoever on the cause. Thus, a voluntary manipulation of the reaction time (e.g., by a bored subject) has no influence on the amount of hours this person has or has not slept before (hours of sleep being the independent variable).

Asymmetry in causal processes is discussed by Dowe (1992a). He argues that causal interactions satisfy conservation laws and causal processes involved in any interaction can, therefore, be classified into two groups: incoming and outgoing processes. One of the groups is regarded as being prior to a process of the other group. Causal interaction could be considered to be linked in a net without loops, therefore, direction can be applied to all causal relations within the net. The classification of causal processes as being causally prior to other could also be accomplished in the direction of entropy increase and time irreversibility. This kind of causal asymmetry is based on physics and the asymmetric feature of time.

Asymmetry is a characteristic property of causality, it can be found (1) in action and manipulation processes, in which causes manipulate their effects, (2) in probabilistic dependences, in which causes of a common effect are independent, but effects of a common cause are dependent, and (3) in temporal order, in which the cause precedes the effect. Given the above, asymmetry can be regarded as a crucial feature of causality. It is, therefore, plausible that asymmetry is represented along with causation in long-term memory. Furthermore, it can be assumed that asymmetry guides our way of processing causal knowledge.

Causal asymmetries in Psychology

Results from studies about the learning of causal relationships provide evidence that the asymmetries mentioned above also have an influence on how we acquire knowledge about causality.

Asymmetries in associative and contingency model theories
Some researchers (Shanks & Dickinson, 1987; Wasserman, 1990) supporting the associationistic view argue that learning of causal relationships is a special case of associative learning and is therefore symmetrical: Cues are associated with the outcome, independently of their causal status. The direction of the causal relation is not of importance, hence no difference in learning from cause to effect (*predictive direction*) and from effect to cause (*diagnostic direction*) should be obtained. Moreover, temporal order is supposed to have no influence on associative learning. Cues are events that occur temporally prior to the outcome and play the role of eliciting responses, outcomes are events to which responses refer. Due to this reduction of learning to the acquisition of associative strength between cues and outcome, predictive and diagnostic learning are regarded as being identical. In a predictive learning condition cues correspond to causes and outcomes correspond to effects, whereas in a diagnostic learning condition cues correspond to effects and outcomes correspond to causes. Within this framework there should be no difference between the learning and mental representation of the two conditions, provided that the cues and outcomes are identical. In the associative learning condition the associative weights represent the strength of covariation between the learning events (Chapman & Robbins, 1990; Cheng, 1997). In causal learning situations cues and outcomes are mapped to causes and effects, but otherwise there is nothing special about causal learning that would set it apart from other associative learning tasks (Figure 4).

In contingency models causation is reduced to covariation, but covariation is also symmetric per se. However, causation is asymmetric: Causes influence their effects and not vice versa. Therefore, contingency models cannot account for causal asymmetries, because statistical correlations are symmetrical: A cause raises the probability of the effect and the effect then typically also raises the probability of the cause. Even though contingency models cannot account explicitly for causal asymmetry, they might still be able to take it into account implicitly, via the direction in which the contingency is computed. Just by the fact that I determine an event as a cause and the following event as an effect provides basis for the asymmetry. The event, which I determine as a cause is the one prior to the one I determine as effect. Nevertheless, contingency models make no clear prediction about causal asymmetries.

Asymmetries in causal model theory

The assumption about causal learning in the *causal model theory* is that people acquire knowledge about new causal relationships via processes that are guided by already existing knowledge about properties of causality. Causes are prior to their effects, accordingly, knowledge acquisition about new causal relations underlies that directionality. Therefore, in a predictive learning condition the cues correspond to the causes and the outcomes correspond to the effects. In a diagnostic learning condition the order is reversed: The cues correspond to the effects and the outcomes correspond to the cause (Figure 4).
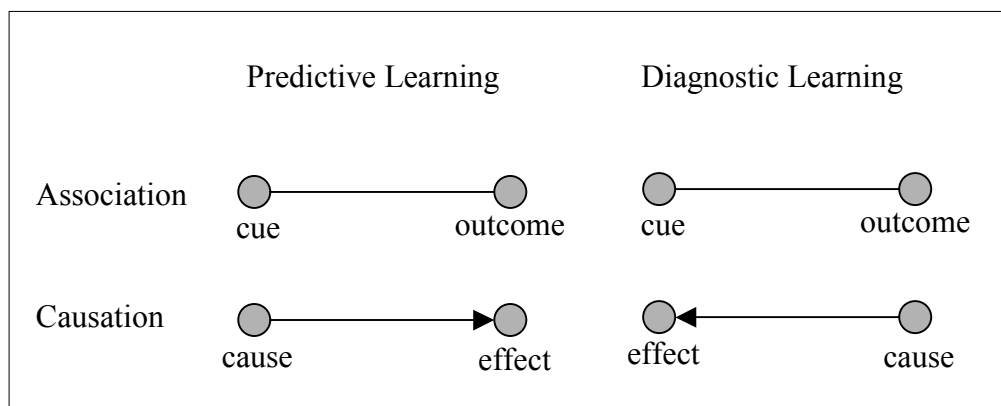
Figure 4: Differences in causal learning and associative learning conditions.

Waldmann and Holyoak (1992) could show that subjects were sensitive to the order of new causal relationships, which had to be learned. They used *a two-phase blocking paradigm*: In the first phase subjects learned that a predictive cue (e.g., a light button) is perfectly correlated with the outcome (e.g., the state of the alarm in a bank). In the second

phase another predictive cue (second light button) was always redundantly paired with the predictive cue from the first phase. Thus, whenever the two buttons were on, the alarm was on and whenever the two buttons were off, the alarm was off. Both switches had the same state (on or off) simultaneously, there was never one switch on and the other off. In the test phase subjects were asked to rate how predictive each switch was, individually, for the state of the light bulb. According to the Rescorla-Wagner theory (Rescorla & Wagner, 1972) the cue in the first phase is a perfect predictor of the outcome and thus possesses maximum associative strength. The additional cue in the second phase is always redundantly paired with the first and hence, according to their theory, has no chance of gaining associative strength, because the first cue is already perfectly associated with the outcome. Consequently, the first cue is *blocking* the second cue from acquiring associative strength.

The manipulation in Waldmann and Holyoak's Experiment 3 (1992) involved the following cover stories. In the predictive condition, the buttons were described as the potential causes of the alarm (common effect), and in the diagnostic condition the structure between the three events was reversed. The alarm was now described as a (common) cause of the two buttons. Only the cover stories varied in this design; the learning trials and the test questions were identical across both conditions. According to associative theories (Rescorla & Wagner, 1972) blocking would occur in both conditions. Because learning for the second cue would not occur.

However, the results of Waldmann and Holyoak (1992) yielded a different outcome. They found that the blocking effect showed an interaction with the causal status of the cues and the outcome: Blocking was only found in the predictive condition, where the two cues were the causes. In the diagnostic condition no blocking effect was observed. This finding was predicted by the causal model theory (Waldmann & Holyoak, 1992; Waldmann, 1996), which postulates that assumptions about abstract causal models interact with the processing of the learning input. In the predictive condition the cues are assigned the role of potential causes and the outcome the role of the common effect. Assessing causal strength within the common-effect model requires holding the potential influence of alternative causes constant. A typical feature of the blocking paradigm is that the redundant cue can never be observed in the absence of the predictive cue, making it impossible to assess the individual causal power of the redundant cue. Although the redundant cue can be observed in the presence of the alternative cue, this cue represents a deterministic cause so that the potential additional impact of the redundant cue cannot possibly display itself. Therefore, it is expected that participants would be uncertain about the potential causal power of the

redundant cue. By contrast, in the diagnostic condition, the cues are assigned the role of potential effects of a common cause. Assessing causal strength within a common-cause model does not require holding alternative effects constant. Hence, participants should have learned that the common cause has two deterministic effects. Since no alternative causes of these effects were mentioned, both effects should be rated as equally diagnostic for their common cause.

To summarize, causality exhibits the feature of asymmetry, that is, it underlies the directional constraint from cause to effect. The acquisition of causal knowledge is discussed within an **associationistic view** (Shanks & Dickinson, 1987; Wasserman, 1990) and **casual model theory** (Waldmann & Holyoak, 1992; Waldmann, 1996). From the **associationistic** perspective causation is reduced to association between cues and outcomes, whereas in the **causal model theory** the causal direction is taken into account.

These two different opinions might as well propose different aspects of the representation of causal relations in long-term memory: In the **associative** perspective causal relations might simply be represented as association between two concepts, whereas in **causal model theory** causation needs to be represented with its asymmetric features and directionality. Nevertheless, these models deal only with causal learning and not with representation of causal knowledge in semantic memory, therefore I now turn to different models of semantic memory and investigate the possibilities they offer from the representation of causation and causal asymmetry

3.2   Semantic Memory

The term semantic memory was first introduced in Quillian's doctorial dissertation in 1968. Later E. Tulving (1972) drew a distinction between episodic and semantic memory systems. "Episodic memory receives and stores information about temporally dated episodes or events, and temporal-spatial relations among these events" (Tulving, 1972, p. 385). Semantic memory on the other hand "[…] is the memory necessary for the use of language. It is a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms, for the manipulation of these symbols, concepts, and relations" (Tulving, 1972, p.386). Tulving (1972) characterizes episodic memory performance as "remembering" and semantic memory performance as "knowing". But not only words or concepts are stored in semantic memory. Also pictorial and spatial information about how things look or what function they have can be included in semantic memory.

Before I focus on the semantic memory system, its structure, and the representation of causal asymmetry therein, the term *relation* needs to be introduced briefly. A relation is regarded as a connection between concepts. A relation between two concepts in the real world is represented by a connection between two concepts stored in semantic memory. These connections are called *semantic relations*. Semantic relations form a subset of relations. The term "semantic" refers to the meaning of a relation, especially its meaning in language. *Causal relations*, in turn, form a subset of semantic relations. They describe and represent the causal connection between two concepts.

Since causal relations exhibit the crucial feature of asymmetry, their representation might reflect this (i.e., they may be stored in an asymmetric fashion). Causal relations are typically not part of assumptions explicitly formulated about the structure of memory. However, in order to comprehend causal semantic relations, they need to be integrated in our semantic knowledge. Furthermore, the feature of causal asymmetry must also be represented in semantic memory, otherwise we would be not able to account for causal directionality when learning new causal relationships.

3.2.1   Representation of causation in semantic memory models

Different structural assumptions about the architecture of semantic memory are discussed in the literature. The most prominent semantic memory models are *network models* and *feature comparison models*. While network models (localistic and distributed) were designed to represent the relations between concepts, feature comparison models deal more with the internal structure (i.e., the features) of a concept.

Table 3 provides an overview of the structure of network and feature comparison models and their representation of relations.

Table 3:

Overview of three semantic memory models.

Model	Contents/Structure	Representation of relations
Localistic network	• labeled nodes representing a concept. • labeled, (sometimes directional) links between concepts.	Links between nodes.
Distributed network	• concept nodes, relational nodes. • input layer, hidden layers, output layer. • unlabeled, meaningless links.	Association of activational pattern at input layer (concept and relational nodes) with activational pattern of output layer via the links to and from the hidden layers.
Feature comparison	• semantic features of concepts.	Feature overlap of different concepts.

Causation in localistic network models

Collins and Loftus (1975) formulated the *spreading activation theory*, which is a further development of the hierarchical semantic network model introduced by Quillian (1969), to model human semantic memory with a computer program. This theory belongs to the class of models that regard the structure of semantic memory as a big network. The basic structure of the model consists of labeled nodes (e.g., "bird", "robin", "feathers", "animal", etc.) that represent concepts in semantic memory and labeled, directed links between the concepts. The links between two concepts describe the features of their relationship (e.g., "has", "is", taxonomic order "is_a", etc.). The model has the following properties:

- If concepts are activated, the activation spreads along the links to the next concepts and it attenuates as it travels along the links.

- The longer a concept is processed the longer the activation is released at a certain rate.

- The more concepts simultaneously activated, the less each individual one will be activated.

- The activation of different nodes sum up. For example, the sentence "A goldfish is orange" will activate the node "goldfish" and the node "orange". If a subject now has to decide, if this statement is true (like in a sentence verification task see section 3.2.3) a certain threshold needs to be reached to trigger that decision. Once that threshold is reached, the path is evaluated by tracing the activation back to its sources ("goldfish" and "orange") and the sentence can be verified.

In this theory, semantic memory is organized along the lines of semantic relatedness. The more two concepts have in common, the closer their relationships will be. If two concepts are highly related or associated with each other, they will have a high strength of connection or association. This strength of association and the activation of the concept nodes are accentuated, whereas the importance of the labels is de-emphasized in comparison to Quillian's model (1969).

Semantic and causal relations in classical semantic networks

Causal relations could be implemented in the spreading activation theory (Collins & Loftus, 1975). For any two events that share a causal relation a link labeled "causes" would be established between them. Causal events that are highly typical would have stronger links than others: The concepts "thunder" and "lightning" are more strongly associated with one another than the concepts "fire" and "destruction" (see Figure 5).
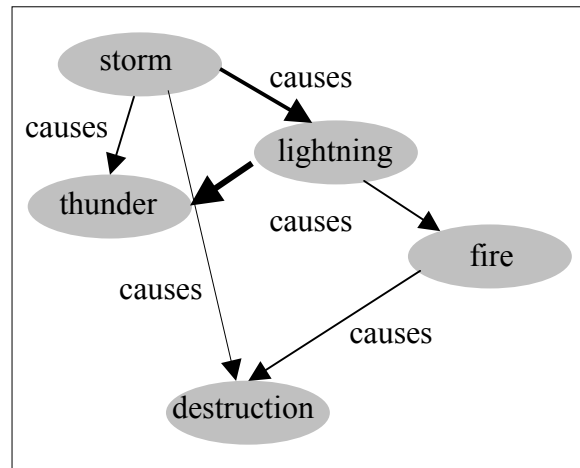
Figure 5: Imbedding of causal relations in spreading activation theory. Thicker arrows represent stronger connections than thinner arrows.

But, how are probabilistic causal relations implemented in the model? One answer could be that as soon as a causal relation is observed between two events, a connection is established between them, but the connection is not very strong. If more observations support the causal relation, the connection could grow in strength.

How could causal asymmetry be represented in localistic models?

Since causal asymmetry is one of the prominent features of causality, the localistic network models have to be examined under this aspect. The links within the localistic model are labeled with directional relations such as "has", "can", etc. Hence, causal asymmetry can be represented within this network models by labeling links as "causes". This label determines the direction from cause to effect and, therefore, implies causal asymmetry. However, the model makes no prediction about what would happen if a relation has to be accessed or retrieved against the direction of the link. Moreover, counter-directional access is not even postulated. Nevertheless, the most typical prediction would be that the evaluation of a relation against the link takes longer. But this explanation is quite weak, because the model does not provide any processural assumptions concerning directional difference. The statement that a diagnostic causal sentence is evaluated slower, because it is presented in the opposite direction is trivial. It would be just a re-description of the evaluation and not a consequence of the structure. In order to withstand this criticism the model would have to make clear statements about the process that makes the access of diagnostic relations more difficult without reducing the explanation to the arbitrarily assigned direction of the links.

Causation in distributed network models

Another group (McClelland & Rumelhart, 1986; Farah & McClelland, 1991) postulates that semantic relations are represented in a distributed semantic network. The main differences to the localistic semantic network models are as follows: (Figure 6):

- Information is presented in a distributed fashion, that is, the *pattern* of activation of different units represents a concept. The specific constellation of the activated units determines the features of the concept. This approach is quite flexible compared to the local representation of knowledge in hierarchical model and it regards the context under which knowledge was formed.

- The links between the units do not convey meanings (like "is"), they only help to spread the activation from unit to unit. Only the pattern of distributed activation represents information. The links just create an association between the units.

- These models need to be trained to simulate the acquisition of semantic knowledge (i.e., the model can learn new semantic relationships).

During training the untrained network gets input signals. The network output is then compared to the desired output. The deviation of the network output to the desired output is calculated and accordingly the weights of the input and output units are altered to minimize the discrepancy between the input and output pattern and to more closely approximate the desired output pattern. Thus, the network is given the input pattern, information about the desired output, and an algorithm to change the weights (between input and output nodes) accordingly. Depending on the predetermined weights between the units learning can be slow or fast.

Semantic and causal relations in connectionist semantic network models

As in the classical semantic network models causal events could also be embedded in a connectionistic structure. The main difference would lie in the structural assumptions of the connectionist model itself. The relational input layer would have to be equipped with additional nodes to represent "causes". The network could be trained with causal relations such as "fire causes heat". The input pattern about a causal relation between concepts such as "fire causes heat" and "fire causes smoke" would activate the cause-node ("fire") and the relational ("causes") node, the network then would be able to complete the output pattern by activating the effects-nodes "heat" and "smoke" (see Figure 6).
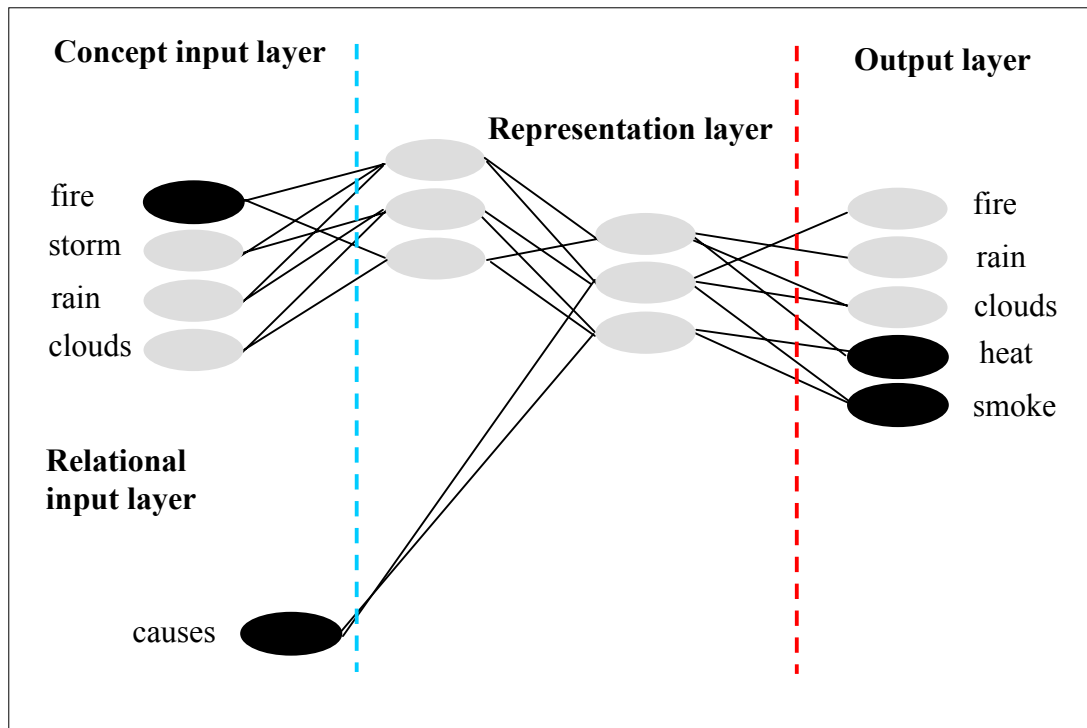
Figure 6: Imbedding of causal relations a connectionist network model of semantic memory. The network has learned that fire causes heat and smoke. The black nodes are activated during training.

How could causal asymmetry be represented in connectionist models?

One way of assessing if the models are capable of representing causal asymmetries is to feed the expected output pattern to the network and see if the model can create the correct (i.e., corresponding) input pattern. For example, the output pattern for the concept "smoke" could be fed to the network. Then the network might be able to trace back the corresponding input pattern to the output concept "fire". The correct input pattern would be an activation of the concept "fire" and the corresponding relation "causes". However, a network structured like the one in Figure 6 would fail for various reasons: a) the network has not previously learned anything about the concept of smoke, thus there are no connections for smoke in the input pattern, if smoke is not part of the network training. b) if "causes" is the only node that represents causality, the network cannot account for causal relations in the opposite direction, because it has not learned anything about diagnostic relationships; learning is processed from the input to the output layer and there is no connection from the output layer back to the input layer.

However, human beings have no problems whatsoever naming possible causes for a given effect. A good network model that claims to represent causality should be able to mirror this behavioral data. Therefore, in order to account for both causal directions an

additional relational node "is caused by", representing the diagnostic direction, would have to be added to the network (Figure 7) and the concepts of the output layer would have to be trained, that is, they would also have to be represented at the input layer.
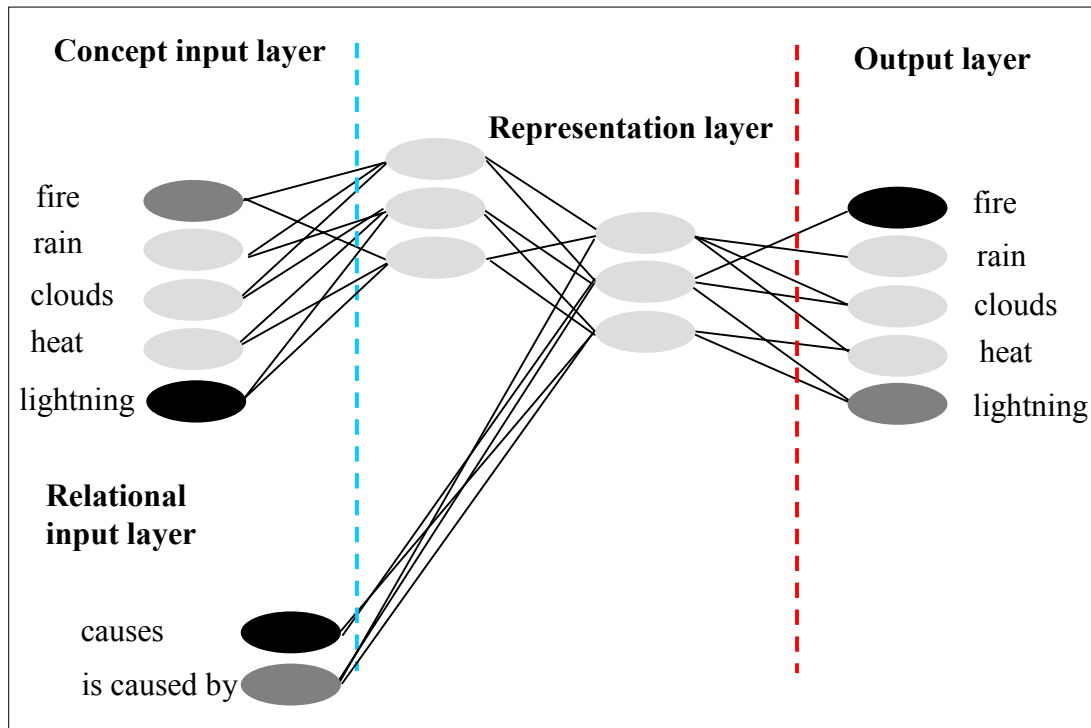
Figure 7: The representation of causal directions via two different relational input nodes. The network is trained for both directions: the black nodes indicate the activation for predictive training (lightning causes fire) the dark gray nodes indicate the activation for diagnostic training (fire is caused by lightning).

For example, the network is trained with the concept input pattern for lightning and the corresponding relational input "causes" to generate the corresponding output pattern "fire". In parallel, the concept of fire and the corresponding relational pattern "is caused by" is trained to generate the appropriate output pattern "lightning". Consequently the model would be able to represent both directions and thus be able to account for the empirical finding that human beings are able to retrieve effects for a given cause from memory and vice versa. However, this would also mean that causal asymmetry is not represented. The correct output pattern would be generated with the same speed for each direction and the output nodes will gain the same degree of activation.

In order to create asymmetry within that kind of model, a different assumption about the strength of connections between the units needs to be made. All stimuli activate their associated representations to some degree, depending on the strength of connections. Connection strength is a product of practice or learning. The acquisition of new causal

knowledge underlies prior stored abstract knowledge about causal directionality. Thus one possibility to account for causal asymmetry might be that causal relations are more frequently learned in the predictive direction. Within the context of the model this would mean that the connection weights from predictive training could be larger than from diagnostic training, because more predictive training would have been processed. Therefore, causal asymmetry would be represented in terms of different connection weights between the two directions. The different connection weights would then be responsible for the different degree of activation of the output pattern. That is, the output pattern representing the effect would receive a higher degree of activation than the output pattern representing the cause.

Nevertheless, this is just a vague idea, because it could also be the case that for some domains of knowledge the opposite is true: causal relations are more trained in the diagnostic direction. For example, a physician is highly trained to derive possible causes of a disease to treat it. A connectionistic model that would simulate this kind of expertise would have larger connection weights for the diagnostic direction.

Network models in general do not seem very well equipped to account for asymmetric semantic relations without serious modifications to their structure or providing post-hoc explanation. In the localistic network the links can be directional, but no further explanation is given about what would happen, if memory were retrieved against the link's direction. Furthermore, the activation is assumed to spread equally fast in each direction regardless of the link's labeling. Thus, links could be reduced to undirectional, associative connections between the concepts.

In distributed networks the input of the network is associated with the desired output. The network is trained with association of different nodes. If causal asymmetry is determined by different connection weights, the output nodes would get a weaker activation for the diagnostic direction. However, it is hard to explain how the degree of activation could be related to an underlying retrieval process. Another weak point of these models is that they do not have a memory component for the abstract feature of causal asymmetry. Causal asymmetry can only be incorporated via learning processes and not via stored (i.e., already gained) knowledge.

Causation in feature models

The feature comparison model developed by Smith, Shoben, and Rips (1974) has found widespread use (McCloskey & Glucksberg, 1979; McNamara & Sternberg, 1983) as an alternative to network models.

Feature comparison models are based on the following assumptions:

- The concept of a noun is divided in a subset of semantic features

- It is represented by a set of elements and not in a network structure.

- These features incorporate a relevant semantic dimension and a range of values.

- There is a distinction between two types of features: features that are *defining* and features that are *characteristic* for a concept (Smith et al. 1974). [3]

For example the concept "bird" is represented by an array of features such as: "biped", "beak", "sings", "flies", etc. (Figure 8).
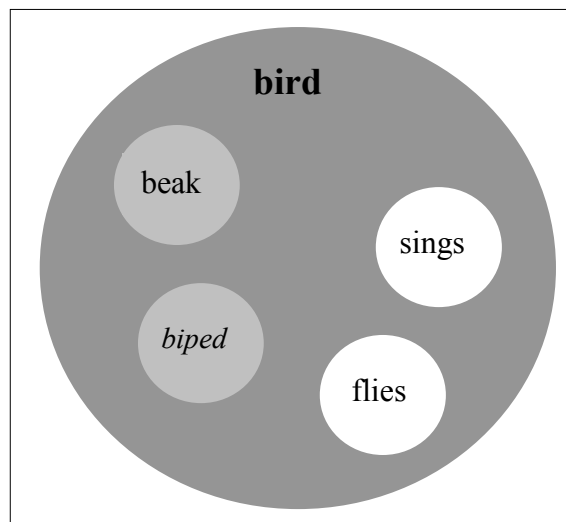
Figure 8: Representation of the concept "bird" including four of its features. Italic printed features are defining and regular printed features are characteristic features.

In feature comparisons models the relationship between two concepts can be inferred by the amount of features they have in common. The degree of overlap can be calculated by retrieving stored defining and characteristic features from semantic memory. If the overall featural overlap of two concepts is high, then they share a relationship such as belonging to the same category (e.g., a wolf and a dog have many features in common, thus

---

[3] Some authors chose another distinction: context dependent vs. context independent features Barsalou (1982); obligatory and facultatory (Klix, 1980a, Kluwe, Wolke & Bunge, 1982).

they are members of the canine category). Accordingly, the concepts are associated because they are highly similar.

Semantic and causal relations in feature models

Feature comparison models can represent different kinds of semantic relation, especially part-whole (a bird has a beak) and class inclusion (a car is a vehicle) relationships, but other types of semantic relations, such as ownership relations, drive these models to their limits. In ownership relations such as "people own cars" there is no overall similarity to be computed between the concept "people" and "car". People have different features than cars (e.g., head and legs vs. motor and wheels) and therefore a feature comparison model would yield a "no" as an answer to the verification of this sentence.

The same is true for causal relations: most causal relations consist of concepts that do not have a high featural overlap or other similarities. Consequently an overall similarity between the cause and the effect cannot be computed. Therefore, it seems unlikely that feature comparison model can represent causal relations.

How could causal asymmetry be represented in feature models?

Feature overlap models face problems when it comes to representing relations that do not require overlapping features of concepts. Causal relations belong to these kinds of relations. Given that these models cannot explain the representation of causal relations in general, there is only little doubt that they can also not explain causal asymmetries.

So far, the three different semantic memory models have difficulties in representing causal asymmetry explicitly instead of explaining it post-hoc.

3.2.2 Representation of causation in the relational element theory

A completely different perspective on the representation of different types of relation in semantic memory is given by the *relational element theory* (Hermann & Chaffin, 1986), in which a processing theory of semantic relations comprehension is proposed. The authors made various assumptions, based on linguistic and psychological research, about

the representation of relations in semantic memory, but not about the structure of semantic memory itself, which stands in contrast to the semantic memory models introduced before.

They make the following assumptions:

- Relations can be decomposed into simpler elements, derived from the meaning of the concepts (Figure 9).

- Relations can share one or more elements with other relations. These common elements can be achieved by a rating procedure of these relations.

- Relations have a hierarchical structure: They contain a basic element and qualifying elements that function in such a manner that the primary qualifying element modifies the basic element, the secondary qualifying element modifies the primary qualifying element and so on. Independent elements are not hierarchically related.

Language tasks require the processing of relation perception, which means the determination of the relation type between two concepts.
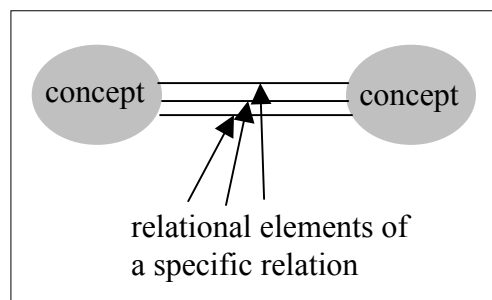
Figure 9: A relation between two concepts containing three relational elements indicated by the three thin lines.

To perceive a relation, specific relational elements have to be present or absent. A specific criterion for each type of relation needs to be established and then the relational elements of this relation are retrieved from long-term memory. After that each relational element has to be evaluated starting from the basic element to the qualifying element. If the process fails to find a relational element in a pair of concepts, the evaluation is terminated. The processing of independent elements can be accomplished without the order constraints. When there are no relational elements left to be evaluated, the decision can be made. The evaluation process is carried out depending on the strictness of the response requirements: if the requirements are very strict, then the pair has to possess each single relational element of the relation, if the requirements are less strict, only some of the relational elements are evaluated to provide a positive response.

Hermann and Chaffin (1986) validated their relational element theory on several studies investigating class inclusion, synonymity, and antonym verification. They found that, with an increasing number of elements to evaluate, the latencies of a "no" response increased (e.g., "slow-late" was rejected slower than "stubborn-angry" in the synonymity verification), the verification time decreased with an increasing number of elements (e.g., "insane-crazy" was verified faster than "abnormal-crazy" in the synonymity verification), and response latencies were longer for non-related word pairs that possessed characteristics similar to, but not the same as, the elements of the verification relation (e.g., "bird-moth" was rejected slower than "bird-soap" in a class inclusion verification).

Causal relations within the relational element theory

The approach Herrmann and Chaffin (1986) take has an interesting implication. Its foundation stems from network models, but it opens up a whole new dimension of relational processing. The links between the concepts are not just meaningless connections, they are concepts themselves, relational concepts. These relational concepts in turn can be decomposed into more elements to get the fine-grained structure of a relation. The relational elements can share similar properties with other relations. Thus, it is the overlap of relational not of conceptual features, that can be compared. A causal relation could be made of a spatial, temporal, and causal relational element. Temporal and causal relations share similar features such as directionality and order. Two events appear along a time line and the event A (the rooster crows) appears always before event B (the morning dawns). However, only the causal relational element contains the meaning of causation: event A (spark) not only appears before event B (fire), it is the reason why event B appears. Every time A appears, B will follow, thus A is necessary for B to appear; they form a constant conjunction (Figure 10). If subjects are instructed to look for a causal relation as a target relation, they have to consider the strict constraints of a causal relation, especially the causal element of that relation. Otherwise they would base their assumption on co-variance and temporal order and not on causation. Therefore, causality is the basic element and temporal order and contiguity would be qualifying elements as proposed by the relational element theory.
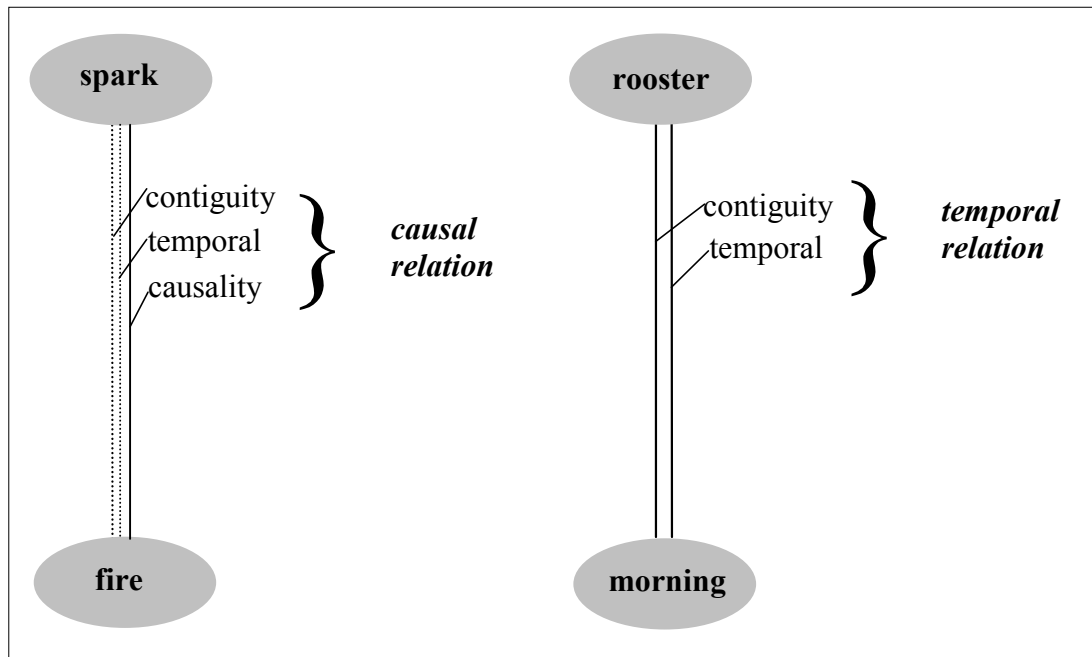
Figure 10: Example for the relational element theory: left: causal relation; right: temporal relation. The solid lines indicate basic relational elements and the dotted lines indicate qualifying relational elements.

As described above the retrieval of a causal relation from long-term memory could be accomplished by retrieving its relational elements. Each single element needs to be evaluated and then a decision can be reached. The evaluation of a causal relation would depend on the degree to which the elements possible of two events resemble the criterial elements such as temporal order, spatial and temporal contiguity and causality (i.e., necessary connection). If two events are connected by a relation that possesses these elements, they will be regarded as causal. For example if the two concepts "spark" and "fire", and a causal relation has to be evaluated, then the evaluation could start with retrieving the features of a causal relation from long-term memory (Figure 11). The following criteria would have to be met: there is a temporal and spatial contiguity in that spark and fire occur in a certain time and location frame; there is a temporal order in that the spark present itself prior to the fire; and finally, it is possible that a spark can cause a fire. The basic relational element of *causality* would be the most important one to determine a causal relation. Because once a causal relationship is established between two concepts, the only element that differentiates it from other relationships is causality.
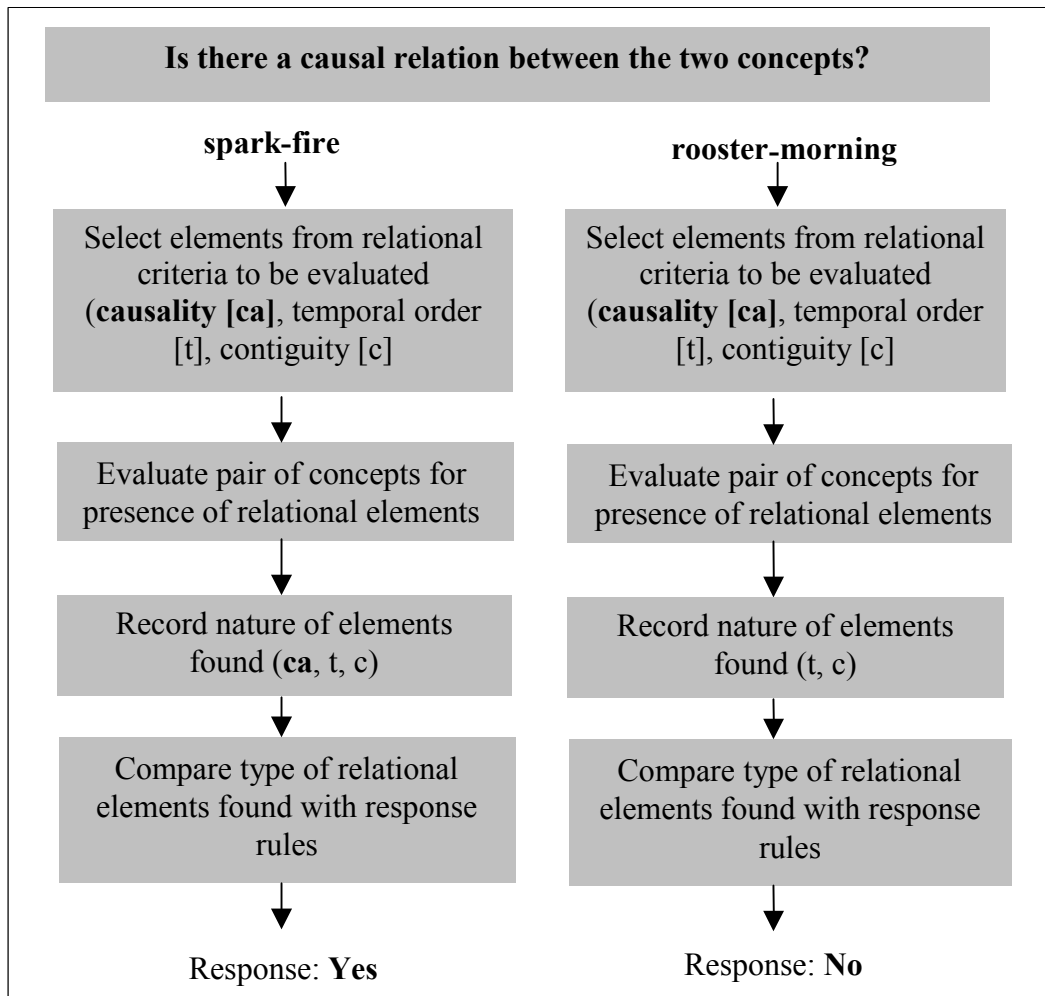
Figure 11: Retrieval process of a causal and a temporal relation in a causal relation task.

How could causal asymmetry be represented in the relational element theory?

The relational element theory has no problem representing causal relations. The representation of causal asymmetry, however, is more difficult within the theory. The causal relational elements represent the direction of the causal relation. An evaluation of a predictive causal relation can be processed by the model and would yield a positive outcome and hence a verification of the causal relation. In a diagnostic causal relation the direction is reversed. Consequently, the evaluation of causal relational elements would yield a negative outcome for the relational elements *temporal order* and *causality*. The relation would be rejected and the evaluation would elicit a "no" answer. Therefore the model would make different predictions for the outcome of the evaluation process of a predictive and a diagnostic evaluation. That is, the evaluation of a predictive causal relation yields a positive answer, whereas the evaluation of a diagnostic relation yields a negative answer. However, human beings are able to determine diagnostic causal relations between two

concepts as a causal relation, thus, relational element theory is not in accord with empirical data.

To summarize, the relational element theory can represent causal relations and causal asymmetries, but the predictions of the theory for diagnostic causal relations cannot account for empirical findings.

However, the relational element model could answer the question whether or not the relation between concepts in semantic memory might have different features and, moreover, if the relation can be accessed and evaluated according to its specific features. According to the theory the processing of a causal relation is task dependent.

3.2.3 Task specific access to causal relations

The models I discussed so far make various assumptions about the representation of causal relations and causal asymmetry in semantic memory. Now, I turn to the retrieval of causal relations represented in semantic memory. If they are stored then they also need to be retrieved.

The retrieval of semantic relations is typically tested with two different kinds of semantic memory tasks: *explicit* and *implicit*. The term *explicit* means that subjects are explicitly *probed* by a certain type of task to access semantic memory accordingly, whereas the term implicit means that the access to semantic memory is incidental and passive and not part of the task subjects have to process. Schacter and Buckner (1998) investigated the difference between explicit and implicit memory retrieval processes and found that for the implicit *priming* task the brain activation decreased in various brain areas depending on the task and modality, but for the explicit retrieval (e.g., recall of a studied word list) activation increased in prefrontal and medial temporal regions.

However, both tasks have in common that they test already stored knowledge. Subjects do not have to go through a learning phase as in other memory testing situations.

Explicit measurements of memory – Probing

Prominent examples for the explicit investigation are *categorization tasks* and *sentence verification tasks*. In a categorization task subjects have to decide, for example, whether or not the word "house" or a picture of a house is a member of the superordinate category "buildings". In this specific task the degree of relation between the instance and

the category is manipulated to assess the interference of the degree of relation on the categorization process. If the influence of typicality of an instance is of interest, then the typicality relation between instances and their category will be manipulated. Subjects will have to categorize typical and atypical instances (members) of a given category. Another manipulation concerns the size of the category: relatively small ("cars") vs. relatively big ("animals") categories.[4] The time subjects need to verify or falsify a category membership is used to infer the structure of semantic memory.

In a sentence verification task subjects have to judge whether statements such as "a car is a vehicle", "some animals are birds" or "a fish has gills" are true or false. Once again the response time is taken as the measure of structural access. The purpose of the sentence verification task is not only to gather information about category membership, but also about the integration of features such as "gills" and "fins" of the instance "fish".

Causal semantic relations and explicit measures

A sentence verification task could be suited to access information about causal relation in semantic memory. Subjects are asked to verify causal sentences ("A spark causes fire."). However, there are some caveats that need to be considered when applying an explicit semantic memory measurement to causal semantic relations.

The typicality of the cause-effect relationship could have an influence on the verification time. A high typicality of a cause-effect relationship could exist if it is deterministic (one event always causes the other) or if that relationship per se has a high occurrence frequency and therefore is prominent in semantic memory ("gravity causes things to fall"). The verification time for a typical cause-effect relationship might be shorter than the verification time for a very untypical cause-effect relationship ("A magnet causes deflection"). Typicality of a causal relation can be further differentiated: The cause can be typical for the effect or vice versa. If a cause ("bomb") is very typical for the effect ("explosion"), the verification process could be shorter, than the process for an untypical cause ("flour powder"). If an effect is very typical for a cause such as "Friction causes heat", the verification process could be fast, whereas for untypical effects such as "Friction causes charge" it could take longer.

Another factor that can influence reaction time in a causal sentence verification task, is the manipulation of the degree of non-causal relatedness between two events in a causally

---

[4] Refer to Chang (1989) for an exhaustive review of factors influencing the explicit access to semantic memory.

framed sentences. Events that are framed in a causal sentence might be *falsified* faster, if they have no relation whatsoever ("A door causes a pinball") and might be *falsified* more slowly, if they possess a temporal relationship ("Spring causes thaw"), which can be highly similar to a causal relationship.

Along the same lines, the context may have an influence on the verification time of causal relationships as well. If the relationships in the filler sentences are highly similar or dissimilar to the relationship in the target (causal) sentences, the verification time may increase or decrease, respectively. Thus, if the relationship in the causally framed filler sentences has only a temporal character ("Spring causes thaw") the verification time for the real causal sentences ("Heat causes thaw") and the falsification time for the filler sentences could increase. However, the verification time for correct causal sentences might be faster if the relational contrast between the target and the filler sentences is high, that is, if the filler sentences have no causal relation ("Door causes pinball").

Implicit memory measurements – Priming

The implicit approach to investigating semantic memory is typically applied in a priming paradigm. Priming is considered to be the facilitation or inhibition of the identification or classification of a word after the presentation of a related word (Neely, 1977, 1991). It was initially reported by Meyer and Schvaneveldt (1971) and ever since a huge discussion about the underlying mechanisms has taken place.

Implicit measures of semantic memory

A typical priming task is to present a word for a short period of time, followed by a related word, an unrelated word, a non word or a string of letters. The time period between the beginning of the presentation of the first stimulus (prime) and the beginning of the presentation of the second stimulus (target) is called *stimulus onset asynchrony* (SOA). Subjects can be asked, for example, to determine if the second word is a word or not (*lexical decision task*); or they may be requested to pronounce the second word (*naming task*). The reaction time is taken as a measure of the access speed to semantic memory. There are robust findings showing that subjects are faster in identifying a string as a word, if the preceding word was related in some way; for example, category membership (fish-trout) or part-whole relationships (fish-fins).

By applying a priming paradigm the focus is to find out, what kind of information is activated in semantic memory, if a prime word is presented. For example, does the word

"fish" prime the words "trout", "gills", "swim", and "water"? The priming activation can spread into different semantic structures, such as categories ("fish" could activate "trout", "cod", etc.), features ("fish" could activate "gills", "scales", etc.) or context situations ("fish" could activate "water", "sea", etc.).

Two different priming processes have been distinguished: associative and semantic priming (Moss, Ostrin, Tyler, & Marslen-Wilson, 1995). Associative priming is considered to be automatic and fast. It is found for pairs of words that have a certain frequency in following each other in spoken language (Plaut, 1995; McNamara, 1992). Words that co-occur in spoken language are assumed to establish a link between each other, and are therefore associated with one another. Hence words that are primed in an associative priming task are also most likely the same words that will be produced in a free association procedure (Nelson, McEvoy, & Schreiber, 1998). Consequently, associative priming can be attributed to strength of association between the two words (e.g. "butter"-"bread"), but not to sharing semantic relations. Associative priming occurs at a SOA of less than 500 ms. The reaction times (RT) it produces are also shorter than 500 ms.

Semantic priming on the other hand is regarded to occur for pairs of words that share some semantic features (e.g., "dolphin"-"shark"). And, most importantly, it is assumed that it does not rely on the spreading of activation, because the words do not necessarily exhibit a high strength of association between each other. Therefore it is the semantic relation between the prime (dolphin) and the target word (shark) that facilitates the processing of the target word. Semantic priming is considered to be a conscious process. The response times (RTs) are longer than 500 ms and longer SOAs (over 750 ms) are chosen to test semantic priming.

Priming is also commonly used as a tool to investigate the different memory systems in clinical studies. Patients who suffered from a variety of memory and speech impairments were tested in different priming experiments: associative priming and semantic priming. Studies showed priming in amnesic patients, who do not have any recollection of the explicit testing, but show clear priming effects (see Graf & Schacter, 1985; Tulving, Hayman & McDonald, 1991, Gabrieli, Fleischman, & Keane, 1995; Graf, Squire & Mandler, 1984). Tyler, Moss and Jennings (1995) also found semantic priming for concrete and abstract words in aphasics. It can be concluded that priming taps into a different memory system than other procedures such as cued recall. Moreover, it harbors a completely different method to access memory.

Analogical priming of semantic relations

As could be seen in priming studies, implicit access to semantic memory contents can be accomplished via various semantic relations, such as category coordinated, functional relationship, etc. (Moss, et al., 1995). A question, which priming tasks have not yet answered, is whether or not the same memory contents can be accessed via different semantic relations. For example, does the prime word "broom" facilitate the target word "sweep" via a script (cleaning situation) relation or via a functional (a broom is used to sweep) relation? The problem arising from this example is that we do not know what kind of relation exactly caused the priming effect. As long as participants do not know that they have to look for a specific semantic relation, the information about which type of semantic relation was responsible for priming cannot be gathered. Only if the relation remains the same over all trials (e.g., only functional word pairs are used), subjects could get an inkling about the type of relation and create an expectancy strategy to process the priming task. Spellman, Holyoak and Morrison (2001) found analogical priming via semantic relations. In their experiments they tested subjects in a relational priming paradigm. Participants were given a pair of words that possessed a specific semantic relation such as "lives in" for 400 ms. This word pair functioned as the prime (e.g., bear-cage). Then the first target word appeared on the computer screen. After 250 ms the second word was presented. The two target words had either the same (e.g., bird-nest) or a different relation (bird-egg) as the prime pair. Subjects had to name the second target word. Spellman et al. (2001) found priming only, when subjects were explicitly told to *note and use* the semantic relation of the prime word pair. If subjects were told to attend to the words only, no relational priming was found. It can, therefore, be concluded that in order to prime a semantic relation, subjects have to be explicitly instructed to take notice of the semantic relation.

Causal semantic relations and implicit measures

Priming tries to answer the question, what kind of semantic relations are activated when a word is read or heard out of context. As delineated above the differentiation between associative and semantic priming is of important. However, it is argued that associative relationships are only stronger semantic relationships where the associative response is elicited because the words possess a very strong semantic relationship (McRae & Boisvert, 1998). On the other hand Fischler (1977) argued that in many priming studies semantic priming could be attributed to an association between the prime and target and not to a semantic relationship per se. This debate still continues.

In the early models of semantic memory (Quillian, 1969) hierarchical categorical relationships were regarded as the basis of semantic memory organization. Consequently, many early priming studies investigated these category relations.

In a *lexical decision* priming study by Moss et. al. (1995) priming was investigated for three different types of semantic relations. The prime-target word pairs belonged either to the same semantic *category* (e.g., pig-horse), were part of a *script* (e.g., restaurant-wine), or shared an *instrumental* relation (e.g., broom-sweep). In addition, some of the semantic relations were also associatively related. They found significantly more priming for all three types of relations with additional association, if the words were presented auditorily with an interstimulus interval (ISI) of 200ms and 1000ms. The words were read to the subjects, whose task it was to decide whether or not the target word was a real word. The pairs containing an additional association produced more priming than the pairs without additional association. This effect was labeled *associative boost*. The data differed, however, for visual presentation of the prime-target words: the *category* word pairs primed only when they were additionally associated, the *instrumentally* related word pairs primed both with or without association, and the *script* relations showed no priming at all. For the visual presentation Moss et al. (1995) used a single word presentation, where the word remained present until a decision was made. In their study Moss et al. also replicated results from Shelton and Martin (1992), who found priming only for associated category coordinates. Thus, different types of semantic relations elicit different degrees of priming depending on the strength of their association, the specific priming task, and stimulus presentation.

In some neuropsychological studies it could be shown that different kind of semantic relations are impaired in aphasics and Alzheimer's patients (Chertkow, Bub & Seidenberg, 1989; Moss, Tyler & Jennings, 1997). Tyler and Moss (1997) also found that functional relationships (e.g., broom-sweep) between prime and target are relatively impervious to breakdown in semantic dementia and aphasia. Other studies investigated priming for category coordinates (Moss et. al., 1995), synonyms (Tyler, Moss, & Jennings, 1995), and perceptually related primes and targets (Schreuder, d'Arcais, & Glazenborg, 1984). They all found priming for these kinds of semantic relations.

The effect of priming on causal relations is not yet clear. To my knowledge there has been, to date, no study investigating causal priming. However, given the results of priming studies with other types of semantic relations, it seems plausible to assume the presence of priming effects for causal relations.

To summarize, explicit and implicit access processes to semantic memory differ and yield divergent empirical results. So far the access and retrieval of causal relations for both types of semantic memory measures have not been sufficiently investigated. A causal connection between two concepts could be represented based on the association between them without accounting for causal directionality. This would stand in accord with the associationistic perspective of causal learning (Shanks & Dickinson, 1987; Wasserman, 1990). But it is contrary to the perspective of the causal model theory (Waldmann & Holyoak, 1992), which postulates that causal relations in semantic memory would include information about causal direction. Therefore, a distinction needs to be drawn between these two perspectives. In an explicit semantic memory test, causal relations could be accessed via association and causation (including causal directionality), respectively and they might provide different results depending if the causal direction needs to be retrieved as well.

An implicit semantic memory task would test a different access mechanism of causal relations. Moreover, as can be seen in priming studies like Moss et. al.'s (1995), the degree of association between the words has an influence on the degree of priming, that is, priming is sensitive to different strengths of association. Therefore, the investigation of priming for causally related word pairs would also provide insight about potential associative connections between the causal prime and target words. Friedman (1990) argued that events are more strongly strong associated in temporal order. That would also mean that causally connected events are more strongly associated in the predictive direction than in the diagnostic direction. In a priming study this assumption could be tested. If indeed the predictive direction contains a higher degree of association, then the degree of priming would be higher for predictive causal prime-target pairs, whereas for diagnostic word pairs the degree of priming would be lower.

All these factors need to be considered when testing explicit and implicit retrieval processes of causal relations from semantic memory.

3.3   *Causal semantic memory*

After having discussed various models, ideas, and assumptions about semantic memory; I now turn to *causal semantic memory*. A potential theory about causal semantic memory should be equipped to account for findings of explicit and implicit semantic memory measures, structural assumptions, and empirical evidence from causal learning studies and they have to incorporate:

1. specific features of causal relations, especially, causal asymmetries

2. the storage of causal memory contents (i.e., concepts about causes and effects)

3. and the relevance of these features to access and retrieval processes from the database of semantic memory.

3.3.1 The retrieval of temporal relations from semantic memory

Causation has the feature of temporal order: causes precede their effects. In order to develop a model about the retrieval of causal relations from semantic memory, one needs to take a look at the retrieval process of temporal relations from semantic memory. Even though, as mentioned in section 3.1.2., causal asymmetries cannot be reduced to mere temporal order, empirical findings about the access to temporal relations could prove helpful.

Krüger and Van der Meer (1999) investigated the retrieval of *temporal relations* from semantic memory. They postulated that temporal relations of events are stored and retrieved from semantic memory in their natural order, which they called prospective direction (compare Friedman, 1990). They presented pairs of words, one word after the other, that possessed a temporal relation, that is, exhibited a certain regularity (e.g., burn-extinguish). Subjects were not instructed to check the words for the temporal relationship, but were to determine if there was a *meaningful* relationship between the words. Krüger and Van der Meer (1999) manipulated three factors in their study: the word pairs were either temporally related or had no relation at all, the SOAs between the words were either 200 ms or 1000 ms, and the order of the temporal word pairs was either prospective (burn-extinguish) or retrospective (burn-ignite). Their temporal word pairs were taken from a three word temporal chain (ignite-burn-extinguish). They collected the reaction times, errors, and they also measured the diameter of the cornea, because the dynamic of the pupils

is a sensitive indicator for the degree of cognitive demands (Klix, Van der Meer & Preus, 1984). The more demanding a cognitive process is, the higher the tendency of pupil dilation. They found that participants were significantly faster in determining the relationship when the temporal word pairs were presented in prospective order. This was found for both SOAs: the prospective reaction time was 858 ms and the retrospective reaction time was 912 ms at a SOA of 200 ms. At a SOA of 1000 ms the reaction times were 907 ms for the prospective and 952 ms for the retrospective direction. Pupil reaction interacted with the SOAs: There was a difference in pupil diameters depending on temporal direction in the short SOA, but no difference in the longer SOA. In the short SOA condition, the diameter was larger for the retrospective direction.

Krüger and Van der Meer (1999) explained their results with the fact that temporal relations are stored and accessed in semantic memory along the time arrow. Thus, access to temporally related events is accomplished via that time line. When the order of the presented events is reversed, the access to these relations is more difficult and cognitively demanding, since the reversed order needs to be integrated. Pupil dilation was overall larger in the short SOA condition, indicating a higher cognitive demand in that condition. However, no significant difference was found for the two temporal directions in the long SOA condition. Even though behavioral data show a significant reaction time difference, the pupil measure does not. Unfortunately, in no explanation was given for the missing support of the pupil measure.[5] Another problem with this study was that the stimulus material consisted of three word temporal chains, with the result that different word pairs were presented in the different temporal directions: prospective: burn-extinguish vs. retrospective: burn-ignite. Differing strengths of association between the word pairs were controlled for. Thus, the prospective and the retrospective word pairs were assumed have the same strength of association. Nevertheless, it would be more plausible to use the same item pair in each direction in order to be able to compare them directly.

And lastly, as provided in their example above, the word pairs might not only have had a temporal relation, but also contained a causal relation: ignition can cause burning and burning can cause extinguishing. Thus, their stimuli are based on *regularities* and *constant conjunctions*, which represent or are understood as causal relationships (Hume, 1739). Therefore, their results could also be attributed to the fact that their subjects inferred causation instead of a temporal relation to derive the meaningfulness of the word pairs.

---

[5] Their data were published as an abstract of a talk at the annual meeting of the Gesellschaft für Kognitionswissenschaften KogWiss99.

In an earlier experiment Van der Meer and Schmidt (1992) tested the inference mechanism for relations that had a final, causal, and temporal character. They define a final relation as a means or goal relation between two events, a causal relation as a cause-effect relationship, and a temporal relation as a relationship, in which two events follow each other on a time line. They also used filler word pairs that did not have any relationship at all. (Table 4).

Table 4:

Examples for the four different types of relations used in the experiment. The words are translated from German into English.

Type of relation	Examples
final	cooking-eating feeding-slaughtering
causal	die-burry speeding-accident
temporal	cashing-clearing laying brigs-plastering
meaningless	upholster-morse running-folding

Thus, one factor of manipulation involved the type of relation, the second factor was the direction of presentation: prospective or retrospective presentation, and lastly the SOA between the two words was manipulated: 250 ms vs. 1000 ms. Subjects, as in the experiment above, were instructed to decide whether a meaningful relationship existed between the two words.

Their results revealed that in the short SOA condition the prospective direction was faster for all three types of relations, whereas in the long SOA condition the difference between the prospective and the retrospective condition was found only for final relations. Moreover, at a SOA of 250 ms, the final items were compared to the temporal and causal items significantly faster in the predictive condition. Van der Meer and Schmidt (1992) argue that the results indicate that final relations are a part of stationary knowledge, which need not be inferred, whereas causal and temporal relations need to be inferred and, moreover, do not belong to the class of associative relations. Their retrieval is based on strategic processes. Thus, in the 1000 ms SOA condition the time interval between the two stimuli was enough to diminish the influence of the order, but in the short SOA condition,

the strategic processes could not be activated fast enough therefore, the temporal arrow, along which temporal and causal relations are coded in memory, became prominent.

A comparison between the two reported experiments shows that their results contradict each other: There was a difference for temporal relations in the long SOA condition in the later experiment (1999), but there was no difference in the earlier experiment (1992). In her article Van der Meer (1999) does not report her earlier results and hence, provides no explanation for the contradiction.

In a more recent study van der Meer, Beyer, Heinze and Badel (2002) investigated temporal order relations in language comprehension. They based their experiments on the assumption that semantics of time is encoded in grammar and the lexicon (Zwaan, 1996) and that syntax helps to construct a situation model (Kintsch, 1992). A situation model is regarded as a mental model of a situation, which could be described by a text information. The mental model helps to integrate the described information with the situational information represented in long-term memory. However, the temporal information is not always provided explicitly. For these cases it has been postulated that the default assumption in text comprehension would be that the temporal order of reported situations corresponds to the situations' chronological order. This assumption is called *iconicity* assumption (Fleischmann, 1990; Zwaan, 1996). Van der Meer et al. (2002) investigated the recognition of frequently occurring events in dependence of their temporal framing (Experiment 1). Subjects were given word pairs describing chronological events relating to a *script* (Shank & Abelson, 1977). These word pairs were modified with the words "before" and "after". The presentation was either in the correct chronological order (e.g., after bite off-chew) or in the reversed order (e.g., before digest-swallow). Subjects saw the first word and the modifier for 200 ms or 1000 ms before the second word was presented. They had to indicate by pressing a button, whether or not the word pair describes the event sequence corresponding to the temporal modifier. For example, the word pair "bite off-chew" would correspond correctly to the modifier "after", and the word pair "digest-swallow" corresponds correctly to the temporal modifier "before". Van der Meer et al. (2002) expected a facilitation of the response, if the events where in the chronological order, which would be supported by the *iconicity* assumption. That is, stimuli like "after bite off-chew" would be recognized faster than stimuli pairs like "before digest-swallow", because in the latter one the presented order is reversed to the chronological order of the real events. Their data showed that participants took significantly longer to respond, if the order was chronological reversed in both SOA conditions and that the response was also slower in the

short SOA condition. Van der Meer et. al. attributed this result to the establishment of a script-based situation model based on the temporal order of events in reality.

Causally related events could also be part of such a situation model, because they are based on temporal regularities, that is, order and sequence. The retrieval of causal situation might also depend on the *iconicity* assumption about the mental representation of a particular (causal) situation. I will discuss the direction of causal relations and sequences in the next section.

3.3.2 Assumptions and hypotheses about asymmetry in causal semantic memory

In *philosophy* causality is regarded as an important feature of the world. Moreover, it is considered to explain the world. According to Lewis (1993) every event in the world can be explained by its causal history. Most of these histories are long and complicated, but all are made up of causal chains with each link influencing its successor. None of it could stand alone or gain existence without any link before it.

These causal chains or histories the prominent feature that they are *asymmetric*. That is, they can only exert influence on each other in the forward direction, not in the backward direction. This idea was also discussed by Reichenbach (1956), when he introduced the notion of *conjunctive forks* open to the future, but not to the past (see section 3.1.2). In our world causation backward in time is hardly comprehensible like the irreversibility of time (but see Faye, 2001 for a different opinion). Thus, the physicality of our world could be regarded as being asymmetric and underlying causal and temporal directionality.

Given this important feature of the world, the question arises how it might be *represented* in the human mind. We are capable understand of understanding causal relations and use them for our means. We know about causal directionality and its irreversibility, but does this necessarily mean that the concept of causal direction and causal asymmetry is represented in our mind with this specific asymmetric feature?

Psychology investigates the question of causal representation and, among other things, human's ability to understand causation. There are different opinions, which postulate different representational assumptions of causal asymmetry in the domain of psychology. For the purpose of my studies I will differentiate between two: the **causal model view** with the postulate that causal directionality is represented in the human mind (i.e., human long-term memory), and the **associationistic view** that claims causality is

represented only as an association between two concepts. Since these two assumptions differ in their understanding of causal representation, it is quite likely that they will differ when it comes to providing predictions about the retrieval and judgment of causal relations.

Causal asymmetry-Causal model view

The **causal model view** (Waldmann & Holyoak, 1992) postulates that causal directionality is represented as an abstract feature of causality in human long-term memory. This representation of causal directionality between concepts is independent from the association between them. Evidence for this view could be found in studies about causal learning. Waldmann and Holyoak (1992) could show that participants were sensitive to the causal directionality during learning of new causal relations. The reasons why causal asymmetry would be represented in human memory becomes more plausible if we take a look at how we deal with causal directionality in common situations of our life: We are able to predict and anticipate events. This prediction is based on the temporal and especially causal order of events. For example, we can predict thunder from lightning or storm from dark clouds. Without the causal order it seems impossible to determine the causal status of events, thing, concepts etc. This would make the understanding and explaining of the world become very difficult. So, in our daily life predicting events seems to be highly trained, because new causal relations occur in the temporally forward (i.e., predictive) direction. Even animals seem to learn prediction. For example, a pigeon can learn that pressing a particular button "causes" the scientist to provide food. Therefore, the pigeon might be able to predict that every time when the scientist is present and it presses the button, it will be fed. Consequently, prediction helps to seek out positive events and to avoid harmful events.

Moreover we need to be able to predict the outcome of actions in order to adapt our behavior (e.g., eating a specific kind of mushroom would cause stomach ache) and to intervene if necessary. The possibility to intervene in a causal process is only plausible, if we know the direction of the causal process. Manipulating the effect instead of the cause is fruitless. For example, one cause of cavities is poor oral hygiene. A way of preventing cavities (i.e., influencing its occurrence) is improving oral hygiene (i.e., manipulating it). There is no way of manipulating cavities in order to alter oral hygiene back in time. Accordingly, there is only one direction of manipulation and consequently this direction also needs to be represented along with causal relations. Thus, the predictive direction (from

cause to effect) would be the only direction that makes sense to consider when manipulating causal relations.

Another evidence that causal relations are represented with their feature of directionality was introduced above on the section about the retrieval of temporal relations from long-term memory. As I mentioned before, it seems plausible that causal relations are also part of script-based mental representations of situations. Scripts reflect the chronological sequence of events. Some parts of the sequence are only temporally related, but other parts can also be causally related. For example, if I am in a restaurant, a possible sequence of events would be: (1) I order food, (2) the waiter brings it, (3) I take the first bite, (4) the food is too blend, (5) I put spices on it, (6) I take another bite, etc. The first three steps would refer to a chronological sequence of events without causal relations; the steps (4) and (5) would describe a chronological sequence of causally related events. These causal relations follow the chronological order, that is, the forward or predictive direction in time. According to the *iconicity* assumption (Zwaan 1996), causal sequences would also be represented in the situation model in their predominant direction, which is forward in time. This assumption is also supported by Freyds (1987) postulation that all mental representations of the external world emphasize the future time. The mental representation of causal events would therefore also emphasize the predictive direction.

Like the prediction of effects, the diagnosis of causes is based on causal directionality. Even though we might first gain information about the effect and then about the cause, the direction of the causal relation still remains predictive. For example, if a person gets the flu, it can be diagnosed that the cause of the flu was a viral infection. But the causal direction of these two events was still predictive: The viral infection came first causing the flu. Despite the fact that we know about diagnostic causal relations, it might be used rarely and as a consequence be less trained. Moreover, so far very little evidence has been found in animal research that animals can reason diagnostically. That is, if the scientist feeds the pigeon in the cage, it would probably not turn to the button and check if it is pressed.

Given these examples, it would seem plausible that causal directionality is represented in long-term memory, because in our daily life we utilize it. It seems that predictivity of causal relations might be more helpful and trained than the diagnostic direction (given the directionality of the physical world). This training advantage of the predictive direction and also the supposed represented causal directionality, lead to the

suggestion that the retrieval of predictive causal relations from long-term memory should be more efficient and faster.

Causal asymmetry - Associationistic view

However, even if causal directionality (i.e., asymmetry) is found in the physical world, it cannot necessarily be concluded that it is represented as causal asymmetry in the human mind. An alternative view, which can be traced back to Hume (1739) holds that causal relation can be reduced to association; this is view is called **associationistic view**. According to it causal asymmetry is represented by association any additional information about the directionality. In his book "About Time", Friedman (1990) argued that events, which are connected via the temporal arrow, have a stronger association in the forward (i.e., in time) direction. For example, if participants are asked to recall a sequence of events or talk about past events in their lives they typically order them along the timeline. These events are represented in memory along their timeline. In a developmental study Friedman (2002) also found that four and eight month-old children showed a significant preference for the forward presentation of a video in which water was poured into a glass.

Since in daily life events are usually presented in the direction forward in time, the associative connection between single events in this (forward) direction should be stronger than in the reverse direction. According to this view, the only information about a direction is provided by the associative strength and not by the temporal, causal, or any other kind of direction.

Applied to causal relations this would mean that predictive relations are represented as having a higher strength of association, because they are geared forward in time, whereas the association in the diagnostic direction for the same concepts is weaker. It is important to note that within this framework, causal asymmetry is reduced to an associative asymmetry: Causes are more strongly associated with their effects and not vice versa. Causal directionality in the physical world is reduced to a representation of an associative asymmetry in long-term memory.

According to the associative view, connections that have a higher strength of association will be retrieved faster from memory than connections, which have a lower strength of association. It seems plausible to infer that predictive causal relations will accordingly be retrieved faster than diagnostic causal relations.

To summarize both views: the **causal model** and the **associationistic view** predict that the access to a predictive causal relation should be processed faster than the access to a diagnostic causal relation. The **causal model view** bases this prediction on the representation of causal directionality in human long-term memory, whereas the **associationistic view** deducts this prediction based on an associative representation of causal relations in memory. A proof for that distinction between these to contradicting approaches is difficult to find on theoretical grounds. Therefore, this distinction between these two different views would need to be drawn empirically. In my experiments I will try to differentiate these two assumptions by testing and controlling for each separately. In the following I will develop my line of hypotheses and investigation.

Dissociation of the causal and associative views

There are different tasks that offer the possibility to differentiate between the **causal model** and the **associationistic view** of causal asymmetry.

Causal retrieval with symmetric association between concepts

According to the **associationistic view**, retrieval asymmetry for causally related concepts would arise because of an asymmetry in the associative strength between the predictive and diagnostic direction. However, it can be argued that causal relations exist, which are equally associated for both causal directions. Therefore, differences in the retrieval speed between the two causal directions for these equally associated causal relations could be attributed to asymmetric features inherent of causality.

In order to control for strength of association and to detect differences in the retrieval process for causally related concepts, concepts need to be found which are equally associated in both causal directions. That is, the associative connection needs to be matched for the predictive and diagnostic direction. The access to these causally related and directionally equally associated concepts could then be investigated. For example, word pairs that describe causal relationships could be presented in the predictive (e.g., spark-fire) and in the diagnostic direction (e.g., fire-spark). For each word pair it could be determined whether or not they describe a causal relationship regardless of the presented direction. For this experimental setting the causal and associative view would make different predictions:

Associative view: The retrieval speed depends on the strength of association between the concepts. If the strength of association is equal for both directions, no differences in the retrieval speed should be expected.

Causal view: The retrieval speed depends on the direction of the causal relation and not on the strength of association. Therefore, the retrieval speed will be different depending on the direction of the causal relationship.

In my investigation I support the **causal model view**: Causal directionality cannot be reduced to association, and therefore, the two influence factors can be differentiated:

1. **I postulate that the *causal access* to causal relations from semantic memory will exhibit a retrieval asymmetry between predictive and diagnostic causal relations, which will manifest itself in faster retrieval of predictive relations than diagnostic relations.**

This step focuses on the *causal access* to causally related and equally associated word pairs.

But the *associative access* to these concepts contributes to the distinction between causal and associative asymmetry as well. The causal put on the retrieval process of causal relations can now be disregarded, because the association is undirectional given the associatively matched word pairs. Moreover, an association merely determines whether some kind of connection exists between the concepts without taking other possible types of relations such as temporal and/or spatial relations into account. For example, the evaluation of a causal relation between two concepts has to regard causal directionality: spark → fire vs. fire → spark, whereas the evaluation of an associative connection between them can disregard causal directionality: spark ↔ fire. Therefore, the associative access to causally related and equally associated concepts would predict no difference between the two causal directions.

This associative access could be investigated with the presentation of causal word pairs in both directions, but now an *associative* relation would have to be determined. That means it would have to be determined, if the connection between two concepts describes an associative relationship. Both views would predict that there would be no differences in the retrieval speed between these two directions. However, each view would base its prediction on different assumptions:

Associative view: There is no retrieval difference between the two causal directions, because the word pairs are equally strongly associated in both directions.

Causal view: There is no retrieval difference between the two causal directions, because the retrieval process can be executed without taking causal directionality constraints specified by the task into account.

The argument that the access to causal relations depends on the task specification (i.e., associative access is different from causal access) seems plausible and is empirically testable, because by using different retrieval tasks the distinction between two retrieval assumptions could be investigated. In support of the **causal view** I propose my second hypothesis:

2. **The *associative access* to causally related concepts is not dependent on causal directionality and therefore yields no differences in the retrieval speed between the predictive and diagnostic causal direction.**

Explicit and implicit access

Task specific access can be further differentiated between the *explicit* and *implicit* access to causal relations. So far, I have discussed the *explicit* access, when one is fully aware that a causal relation has to be evaluated and the relation is specifically retrieved from memory.

But, also *implicit* access to causal relations should be investigated. Since it addresses the associative strength between concepts it might be helpful for the differentiation between a *causal* and an *associative* asymmetry. Priming paradigms are designed to examine the *implicit* access to semantic memory. In *implicit* memory access the activation of a concept is assumed to spread to other connected concepts regardless of the type of relation that connects them (Collins & Loftus, 1975). Therefore the activation of a concept (prime) can influence the processing of subsequent concept (target), if these two share a connection. Only the strength of association between the concepts is regarded to influence the spreading speed, that is, concepts that are more strongly associated activate each other faster than concepts that are less associated.

Moreover, the stimulus onset asynchrony (SOA) between the prime and the target words is regarded to have an influence on the *type* of priming that takes place. That is, a short SOA (< 500 ms) is considered to investigate *associative* priming, which is the passive

spreading of activation between the concepts based on their *associative* connection. For example, activated (prime) concept "bread" speeds up the lexical decision about the (target) concept "butter", because they are associated.

In a long SOA condition evidence has been found that it could exhibit priming effects due to the semantic relation between the prime and the target concept. The majority of these semantic priming investigations assessed taxonomic relations such as category coordinates. But little evidence has been found that other types of relations cause semantic priming. Moss et. al. (1995) and others (Tyler, Moss, & Jennings, 1995; Schreuder, d'Arcais, & Glazenborg, 1984) were able to find priming for script and functional relations. However, their results were quite ambiguous, for example, they found priming for script relations only with an auditory stimulus presentation whereas functional relational priming was also found in a visual stimulus display.

So far, relational priming for causal relations has not been investigated. Given the variable results of other relational priming studies it seems to be quite unclear, if relational priming could be observed. Nevertheless, it would be interesting to address the possibility of relational priming for causal stimuli.

For my investigation this would mean that the causally related word pairs could be tested in a priming task, where the direction of the relation determines the prime and the target words. For example, in the predictive direction, the cause serves as the prime and the effect as the target. In the diagnostic direction it would be reversed.

This investigation method serves two purposes: 1) to reconfirm that the selected causal word pairs are indeed equally associated (short SOA condition) and 2) to find out if causal relations between concepts also yield *relational* priming like other relations (see Moss et. al. 1995) in a long SOA condition.

To summarize, testing causal concepts in a priming paradigm with a short SOA would help to discover differences between the predictive and diagnostic direction due to associative asymmetry and testing causal concepts with a long SOA might help to discover differences between the predictive and diagnostic direction due to causal asymmetries.

Therefore, the **causal** and the **associative view** make different predictions based on the duration of the SOA:

> **Associative view**: Priming is equally strong for both directions and for both SOAs, because the causal word pairs are equally associated and causal relations per se are represented as associations regardless of the causal direction.

Causal view: There are no priming differences for the short SOA condition, because the concepts are equally associated for both causal directions and the short SOA prohibits relational processing of the prime target relation. There might be priming differences in the long SOA condition because relational processes could contribute to the task and therefore relational priming for causal concepts might be observable. According to the **causal view** I will postulate the following hypothesis:

3. **In an *associative priming* condition (SOA < 500 ms) no priming asymmetry will be found for equally associated the causal concepts.**
    If causal *relational* priming exists it might be found in a long SOA (> 500 ms) condition: predictive causal relations show more priming than diagnostic causal relations.

Neuropsychological basis of task specific access

I discussed the task specific access to causal relations under the aspects of an associative and causal access process. I hypothesized that the retrieval times might be different for the task specifications and the direction of the causal direction. Another way to investigate the distinction between the **causal model** and the **associationistic view** of storage and retrieval is to differentiate between these two approaches on a neuropsychological level. In a functional magnetic resonance imaging (fMRI) study patterns of cerebral blood flow during these two specific access tasks would aid to further differentiation and support of the behavioral data. We[6] have conducted an explorative study, which I will motivate and specify the predictions more detailed in Experiment 6 in the experimental section.

---

[6] The fMRI study was carried out in collaboration of Prof. Dr. M. Lieberman and Prof. Dr. K. Holyoak at the University of California in Los Angeles.

4.   Experiments

The following section provides an overview of all experimental designs that I used to test my assumptions about the retrieval of causal relations from semantic memory. In my line of experiments I assessed different access processes to causal relations.

The presentation procedure of the stimuli remained the same over all 6 experiments: Pairs of words were presented one word after the other, but depending on the investigated access process the tasks were manipulated differently.

Experimental testing of Hypothesis 1: Asymmetries in causal access

The experiments have to differentiate between the retrieval of predictive and diagnostic causal relations. Therefore, the stimulus material needed to include causal relations in both directions for example, "spark-fire" for the predictive direction and "fire-spark" for diagnostic direction and unrelated concepts (e.g., door-pinball). Experiment 1 was the first attempt to investigate causal retrieval asymmetry and should be regarded as a pilot study to test the efficiency of the experimental design and stimulus material.

In Experiment 2 the causal stimuli were matched on strength of association between the predictive and diagnostic direction, to distinguish a causal from an associative influence on the retrieval process. The stimuli selection process rested on the University of South Florida Word Association Database (Nelson, McEvoy & Schreiber, 1998). Additionally, the word pairs were matched on conditional frequencies in both directions.

Participants were told that they should determine a *causal* relation between two concepts regardless of the order of presentation.

An additional study (Experiment 3) investigated another aspect of causal retrieval asymmetry: retrieval asymmetry of causal concepts due to an asymmetry based on a default predictive access preference. That means, the access to causal relations might be understood as an access of a *predictive* causal relation. Therefore, the access to causal relations was forced in either the predictive or diagnostic direction: The participants had to determine a causal relation for either the predictive or the diagnostic direction.

The retrieval speed for the two directions was measured in all three experiments.

Experimental testing of Hypothesis 2: Task specific access

Experiment 4 investigated the retrieval of causally related word pairs in semantic memory dependent on an *associative* or *causal* access. Two experimental groups were

given the same causal and filler stimuli as in Experiment 2, but their instruction differed. One group had to determine whether or not there was a *causal* relation between the two words, whereas the other groups had to determine, if the words were only *associated*. Thus, the access to causal relations was investigated in a *between-subjects* design where the task specific access (associative vs. causal) was the between subjects factor. The stimulus material contained causal relations in both directions (e.g., spark-fire vs. fire-spark) and associated, but non-causal concepts (e.g., emerald-ring vs. ring-emerald).

In Experiment 6a the same relational accesses were tested, but in a *within-subjects* design. That is, every participant had to retrieve *causal* and *associative* relations for *causally* related word pairs.

In both experiments the retrieval speed was measured.

Experimental testing of Hypothesis 3: Implicit access

Associative Priming

In Experiment 5, a priming paradigm, which should investigate the strength of association, was applied in order to account for possible differences between the forward (predictive) and backward (diagnostic) direction in the stimulus material selected from the word association database. The causal word pairs were implemented in a priming procedure: Subjects were presented either a causal (either cause or effect) or neutral (a string of asterisks) primes and had to pronounce the following (target) word. A short SOA (250 ms) was chosen to investigate the automatic associative activation. The naming latencies for the two directions and the priming conditions were measured.

Relational priming

Relational priming has been reported occasionally for different types of relations, but for causal concepts it has not been investigated yet. Therefore it might be interesting to find out, if causal concepts prime each other and in addition if causal directionality has an influence on the degree of priming. Relational priming has been reported for longer SOA conditions, because they allow the initiation of potential relational specific processes. In order to test causal relational priming the same priming paradigm as described above was used, except the SOA was extended to 1000 ms. Once again naming latencies for the two directions and the priming conditions was measured.

Explorative study: Neuropsychological basis of task specific access

These two different explicit access methods (causal vs. associative) might not only yield different results in the behavioral data, but also show differences in brain activation. Experiment 6b was undertaken to investigate differences between causal and associative access to semantic memory on a neural basis. It contained causally related and associated, but not causally related word pairs. The cerebral blood flow was measured of participants determining a causal or an associative relation between two causally related concepts.

4.1   Asymmetries in causal access

In the last section of the theory chapter I discussed the task specific access to causal relations. I introduced two different opinions about the representation and the retrieval of causal relations between concepts. The **associationistic view** postulates that causality is represented in the human long-term memory via association between concepts (Friedman, 1990) and the retrieval speed of causal relations depends on the strength of association between the concepts. Moreover, the connection between concepts in the forward direction of time is stronger therefore the connection between two causal concepts in the predictive direction (from cause to effect) is stronger than the connection in the opposite diagnostic direction (from effect to cause). This difference in association will yield differences in the retrieval speed between the two causal directions. That is, the direction forward in time (i.e., predictive direction) is retrieved faster than the diagnostic direction.

In contrast, the **causal model view** (Waldmann & Holyoak, 1992) postulates that causal direction is represented as abstract knowledge about causality in long-term memory. The predictive direction is more prominent than the diagnostic direction, based on occurrence of causal relations, the order of manipulation, and the ease of reasoning. Therefore access to causal knowledge in the predictive direction should be faster than the diagnostic access.

In this section I start with the empirical differentiation between the **associationistic** and the **causal view** of the retrieval of causal relations. In Experiment 1 I investigate the retrieval differences for causally related word pairs. However, these word pairs were not matched on strength of association, whereas the stimuli in Experiment 2 had an equal strength of association in both causal direction. These two experiments could differentiate between the influence of associative strength and causal asymmetry on the retrieval speed for causal relations. Experiment 3 was a supplementary investigation to control for an access strategy biased towards a predictive access. That is, participants might access causality in the predictive direction first and turn to the diagnostic direction afterwards.

4.1.1 Experiment 1

Experiment 1 was designed to test the explicit causal access to causal relations in semantic memory. According to **causal model view** the explicit causal access to causal relations in semantic memory underlies causal directionality and therefore the following prediction could be made:

The retrieval process for causal word pairs in the predictive order (from cause to effect) should be faster than the retrieval in the opposite diagnostic order (from effect to cause).

Methods

Participants and design

42 UCLA undergraduate students participated for course credit. Their vision was normal or corrected to normal.

The stimuli were presented in a within-subject design: every subject saw every word pair once. The order of the trials and the stimuli within a trial (predictive vs. diagnostic) were randomized and counterbalanced respectively. The trial type: predictive vs. diagnostic was the independent variable and the reaction time was the dependent variable.

Stimuli

The stimulus material consisted of 36 causal and 36 unrelated filler word pairs (see Appendix Table A1). The word pairs were matched on word length for each direction. The causal word pairs were created based on natural causal relations. Some of the causal relations (e.g., friction-charge) were collected from a German book series called "Was ist was" in which science phenomenon are explained for children and juveniles. The other causal relations were based on "everyday" causal perception (e.g., pepper-sneeze). The filler word pairs were random words (e.g., door-pinball) matched on word length with the causal words. The words were chosen from a regular English dictionary. They were presented in font "Arial Black" and size 24 on a white background. The experiment was programmed in Superlab ® and implemented on Macintosh Imacs with a 15" screen and a 1026 x 768 pixel resolution and 256 colors. The words were created as pct- files in Canvas 6.0 graphics software[7].

---

[7] For all experiments the stimuli were prepared in the same way.

Procedure

Participants were given a written instruction (see Appendix B.1). They were asked to decide, if there was a *causal* relation between two words presented on the computer screen. A causal relation was described as: "the event described by the first word *can cause* or *is caused by* the event described by the second word." After reading the instructions they had to repeat it in their own words to avoid misunderstanding about their task.[8]

On each trial subject saw a fixation cross in the center of the screen. After 1000 ms the cross disappeared and a blank screen was presented for 500 ms. Then the first word of the pair was presented for 1000 ms after that the second word was displayed thus the interstimulus interval (ISI) was 0 and the stimulus onset asynchrony (SOA) was 1000 ms. The second word remained on the screen until the participant pressed one of the response keys (Figure 12). If these words represented a causal relation, subjects had to press the letter "C" on the keyboard otherwise they had to press a key labeled "NC". Each subject answered 72 trials. Of the 36 causal trials 50 % were presented in the predictive direction (cause-effect) and 50% were presented in the diagnostic direction (effect-cause); the remaining 36 trials were the non-causal filler word pairs (i.e., there was no relationship between the two words). The response time for each trial was measured and recorded by the Superlab® software. The program also recorded the amount or errors for any given type of relation.
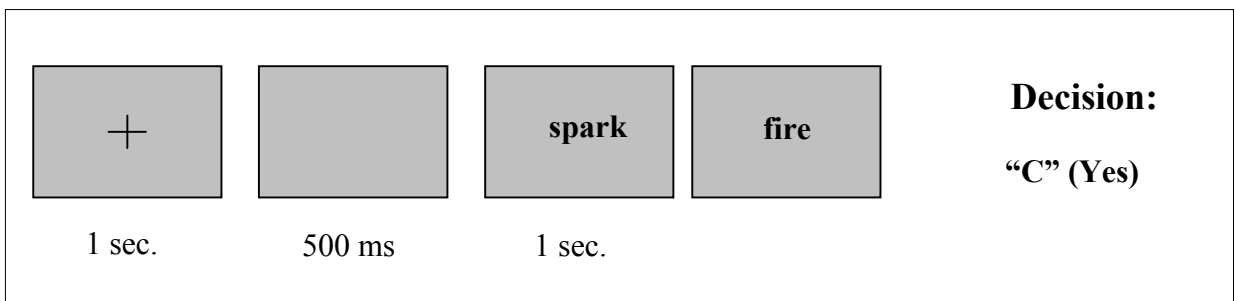
Figure 12: Example for a predictive trial.

Results

Reaction times

For the analysis of the reaction time only the correct answers for the causal trials were taken, the answer for the unrelated filler stimuli were not important for the distinction between the two causal directions. To control for outliers the mean and 2 SD of the raw data were calculated. Every data point 2 SDs above or below the mean was excluded.[9] The data

---

[8] For all experiments the instruction procedure remained the same.
[9] All reported data were corrected the same way.

of the causal trials show a difference of 107 ms in reaction time between the predictive and diagnostic causal relations (Figure 13).

Table 5:

Means and standard deviation of reaction time (in ms) for correct answers in the two causal conditions.

Condition	predictive	diagnostic
Means	1003	1110
SD	221	335

A paired sample t-test reveals that subjects were significantly faster, $t(41) = 2.79$; $p < .01$, in determining a causal relation if the first word presented was the cause; they were slower if the effect was presented first (Table 5). The effect size for this difference is small $d = .38$.
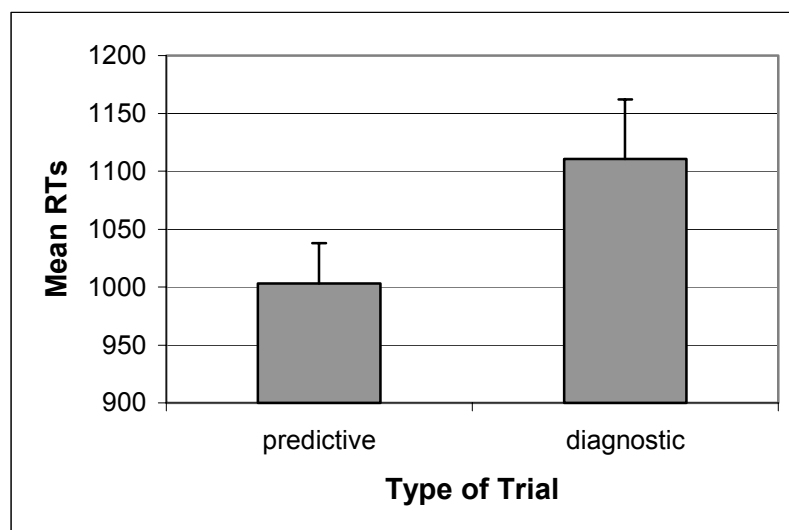
Figure 13: Mean response time (+ 1 SE) for the different directions collapsed over all subjects.

Error Analysis

The analysis of the number of error for the causal stimuli exhibits a marginally significant difference $t(41) = 2.01$, $p = .05$. The errors were defined as incorrect answers on causal trials. The mean numbers of errors are 3.02 for the predictive condition and 3.54 for the diagnostic condition. The error rate was not higher in the predictive condition producing faster reaction times. Hence, a speed-accuracy tradeoff could not be found.
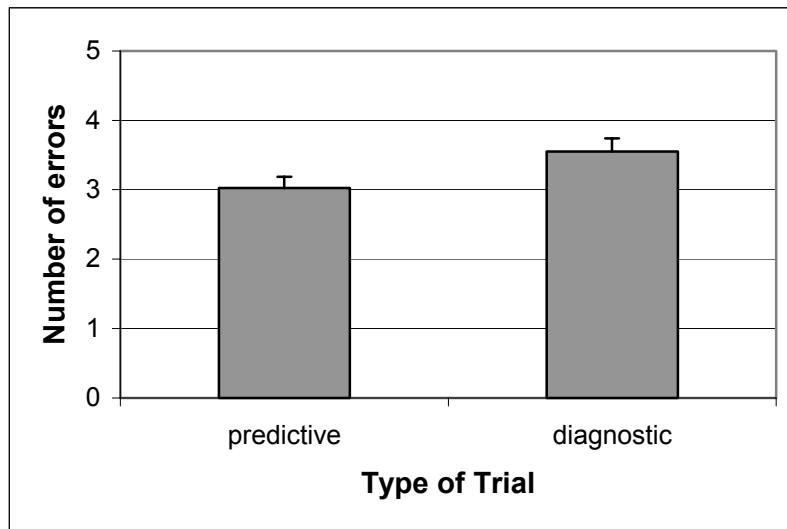
Figure 14: Mean number of errors (+ 1 SE) for the different direction collapsed over all subjects.

Conclusion

Given the differences in reaction time for predictive and diagnostic causal relations, these data suggest that causal events are easier retrieved from memory in a predictive order than in a diagnostic order. The results support the hypothesis that causal direction influences the retrieval process of semantic causal relations. Moreover, the number of errors also reflected the different levels of difficulty in the two conditions. The analysis of errors revealed no speed-accuracy tradeoff between the different conditions. Subjects performed equally well on all causal trials and did not make more errors in the faster condition.

These first preliminary data seem to support the **causal model view**, which would postulate a retrieval asymmetry for causal relations. However, in this experiment I did not control for the different strength of association between the words in the predictive and diagnostic direction. Also the filler words were only matched on word length, but not on strength of association. Theorists endorsing the **associative view** could argue that causal directionality may be reducible to associative relations. For example, Friedman (1990) has suggested that cause-effect relations may be associated more strongly than effect-cause relations. Accordingly, the differences in response time could be attributed to a stronger associative relatedness between the words in the predictive direction and not to causation per se. The associative relatedness is defined as the normative description of the probability that a presented word will call to mind a related second word (Postman & Keppel, 1970). It reflects the co-occurrence of two words in a proposition (McNamara, 1992) or the temporal

contiguity in verbal language (Plaut, 1995). Thus the cause might be stronger associated with the effect than vice versa, because causal relationships occur in this "natural" order not only in the physical world, but also in verbal language. That strong association for the predictive direction could cause a faster/stronger spreading of activation from causes to effects and their pre-activation. Then, if the second word corresponds to the pre-activated effects, the causal relation can be evaluated quicker compared to the opposite (diagnostic) direction. The pre-activation of the causes, given an effect, might be slower/weaker resulting in a longer evaluation process for the causal relation between the two words.

As mentioned in the chapter 3, the *typicality* of causal relations, could also have an influence on the retrieval process. Typicality in this case means that some causal relations are more salient in semantic memory in one direction than in the other, which could be attributed to the frequency with which these causal relations are perceived. The perceived conditional frequency of a given predictive causal relation might be higher than the perceived conditional frequency of the same causal relation in the diagnostic direction. For example, participants might regard the predictive direction of a causal relation such as "a bomb causes fire" as a more frequent relation than the diagnostic direction of the same events ("fire is caused by a bomb"). Consequently, the retrieval and evaluation of the causal relations might be based on the prominence, which could be more salient for the predictive direction of these causal relations in semantic memory.

To account for these arguments, a new set of stimuli for Experiment 2 and all the following experiments needed to be created.

4.1.2 Experiment 2

Given the potential criticism of the previous results, I decided to control for strength of association between the words and the perceived conditional frequency (i.e., typicality) of the two directions of causal relations. Moreover, to decrease the contrast between the causal and filler trials *associated* filler word pairs were selected and matched on their strength of association for each direction.

In order to get a new set of stimuli I collected word pairs from the University of South Florida "Word Association Norm" (Nelson, McEvoy, & Schreiber, 1998). The norming data was gained by asking participants to write down the first word that came to mind that was meaningfully related or strongly associated to a presented word on the blank shown next to each item. For example, if given (cue) CAR_____, one might write (target) ENGINE on the blank next to it. This procedure is called a discrete association task because each participant was asked to produce only a single associate to each word. Nelson et al. (1998) calculated the forward and backward strength between the cue and the target word for each of the 5019 normed words. The forward strength was calculated by the number of subjects that produced the same target word (for a given cue) divided by the number of participants in that group. The backward strength was calculated the same way, except that now the target word served as the cue word. The strength of association between the two words was represented by a number between 1.0 and 0.0 for each direction. I chose word pairs that possessed a causal relation from this database. These word pairs also had a low strength of association (0.00-0.2) in both directions (predictive/forward and diagnostic/backward). 250 word pairs were selected. The reason for choosing only low associated words was an elimination of the influence of strength of association as much as possible and a promotion of the semantic (i.e., causal) relationship during the retrieval process.

Then, to control for typicality effects depending on the presented direction, these 250 word pairs served as stimuli in a questionnaire. 80 students rated the word pairs according to the existence of a causal relation and its conditional frequency (see Appendix C). In the questionnaire causal and filler word pairs was presented. Students were asked to find out, if there was a causal relation between the presented words. If the students determined a causal relation, the next step for them was to imagine that the event described by the first of the two words occurred 100 times. Then they should rate the conditional frequency for the event described by the second word (e.g., "fire" occurs a 100

times, how often does "heat" occur?). They indicated their rating on a scale from 0 to 100 with increments by 10. The direction of the pairs was randomized within the questionnaire and counterbalanced over two questionnaires. I then selected only pairs that did not differ over 30 rating points between the two directions (predictive-diagnostic). The rating difference between the two directions summed up to 0 over all selected word pairs, that means, the overall frequency rating for predictive causal relations was the same as the overall frequency rating for the diagnostic causal relations (see Appendix C).

The new stimulus material would now control for the influence of an asymmetric strength of association and perceived conditional frequency. For these stimuli the **associationistic view** would predict that there should be no differences in the retrieval speed, because the retrieval speed would reflect the equal strength of association between the words. Consequently, if the reaction time still reveals an asymmetry between the two causal directions, then it seems more plausible to attribute it to causal asymmetry as opposed to associative asymmetry. This causality based retrieval asymmetry would support the **causal model view**.

I also exchanged the non-causal filler stimuli. The new filler stimuli had the same low strength of association as the causal stimuli. They were chosen from the same word association database. Their strength of association was equal for the forward and backward direction. By adding non-causal but associated filler stimuli, I wanted to make sure that participants base their causal judgment on causation and not just on association between the causal target stimuli. Accordingly, now all word pairs were associated, but only half of them also had a causal relation.

As in Experiment 1, I predicted different response times between predictive and diagnostic causal relations.

Methods

Participants and design

26 UCLA undergraduate students participated for course credit. Their vision was normal or corrected to normal.

The stimuli were presented in a within-subject design: every subject saw every word pair once. The order of the trials and the stimuli within a trial (predictive vs. diagnostic) were randomized and counterbalanced respectively. The trial type: predictive vs. diagnostic was the independent variable and the reaction time was the dependent variable.

Stimuli

The stimulus material consisted of 68 causally related (e.g., moon-tide) and 68 associated filler word pairs (e.g., ring-emerald). As mentioned above the causal and the associated filler word pairs shared the same low strength of association in each direction. They were selected from the USF "Word Association Norm" (Nelson, et. al., 1998) and causal words were rated according to their perceived conditional frequencies. The associated filler word pairs did not share a causal relation (see Appendix Table A2).

Procedure

The procedure was the same as in Experiment 1 (see Appendix B.2 for the instructions), except that subject were required to answer 20 practice trials (10 causal, 10 filler) at the beginning. They received feedback in the practice trials; during the experiment no feedback was given.

Results

Reaction time

There was a difference of 68 ms in the reaction time for the causal items. Once again subjects were faster on predictive trials than on diagnostic trials (Table 6 and Figure 15).

Table 6:

Means and standard deviation of reaction time (in ms) for correct answers in the two causal conditions.

Type of trial	predictive	diagnostic
Means	1016	1084
SD	175	245

A paired sample t-test reveals that the advantage for the predictive trials is statistically significant, $t(25) = 2.64$, $p = .01$ with a small to medium effect size $d = .34$.
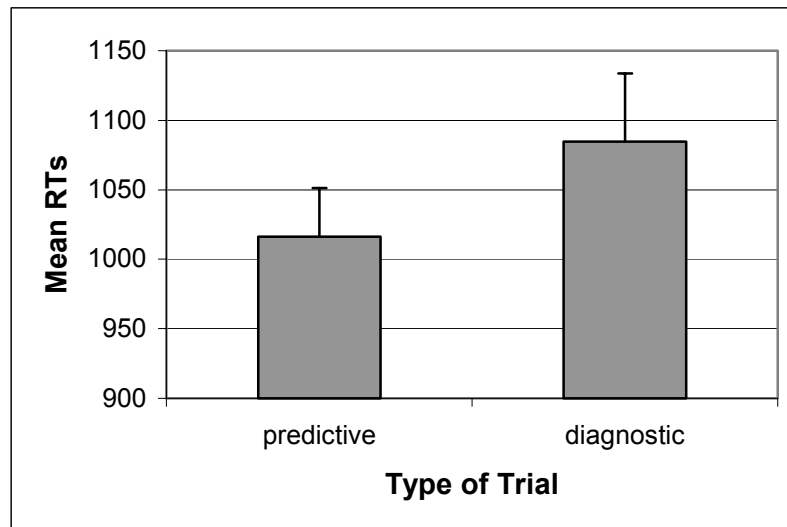
Figure 15: Mean response time (+ 1 SE) for the different directions collapsed over all subjects.

Error Analysis

Error rates for the two causal directions do not differ significantly, $t(25) = 1.74$, $p = .93$. The mean number of errors for the predictive condition is 5.03 and for the diagnostic condition it is 6.26 (Figure 16).
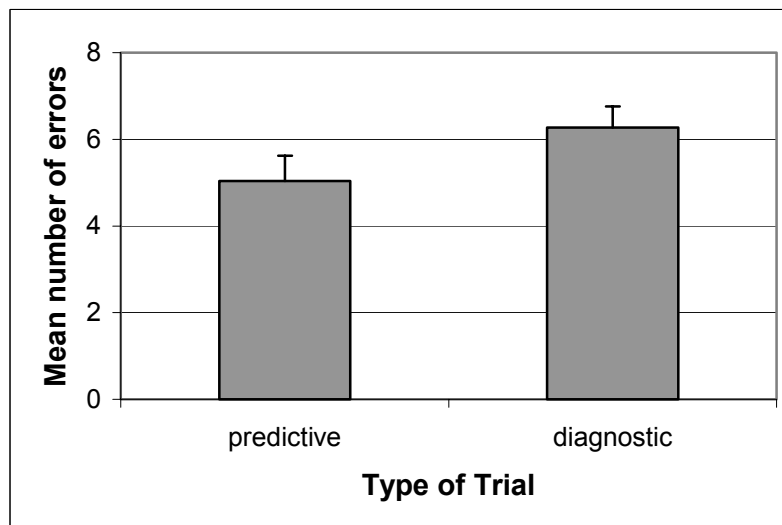
Figure 16: Mean number of errors (+ 1 SE) for the different directions collapsed over all subjects.

Conclusion

In the second experiment I controlled for different strength of associations between causes and effects and also for different conditional frequencies. The results from Experiment 1 could be replicated, even though the absolute differences between the predictive and the diagnostic direction decreased from 107 ms in Experiment 1 to 68 ms in

Experiment 2. This decrement could be attributed to *associative boost* (Moss et. al., 1995). Indeed, the stimuli in Experiment 1 seemed to have a higher strength of association in the forward (predictive) direction.[10] Therefore, this association boosted the response speed for the predictive direction. The new set of stimuli in Experiment 2 possessed only a weak, and moreover, symmetric strength of association in both directions. Hence an associative boost could be ruled out as an explanation for the reaction time difference between the predictive and diagnostic word pairs. These data suggest that the results of Experiment 1 were not only due to different strength of associations between the two directions, but also reflected causal directionality. The effect sizes of both experiments were almost identical, indicating that even though the strength of association was controlled for, the effect in Experiment 2 did not become less prominent.

The increasing reaction time for the associated filler trials in Experiment 2 could be attributed to the context effect (McCloskey & Glucksberg, 1979). The relation in the filler trials is less contrasting to the causal relation. Therefore participants needed longer to distinguish between the causal and the associated filler trials. The error rate for the filler trials also increased compared to Experiment 1, suggesting an increased task demand.

To summarize, Experiments 1 and 2 investigated the *causal* access to causally related word pairs. The stimuli were matched on strength of association in Experiment 2 to rule out retrieval differences based on different strength of association between the two directions. The results indicate that the retrieval asymmetry for the causal access arises, because of a causal asymmetry and not because of an associative asymmetry. Thus, there is a privileged causal direction in semantic memory that honors the natural order of causes and effects in the world. A comparison between the results of Experiment 1 and 2 shows that associative strength has an influence on the retrieval speed, that is, it makes it faster (associative boost).

These first two experiments support the **causal model view**, which would predict these retrieval speed differences. Especially Experiment 2 provides strong support, because an influence of associative strength on the retrieval process postulated by the **associationistic view** is ruled out by the equal associative strength between the words.

However, maybe the source of the retrieval asymmetry does not lie in the nature of the causal relation itself. An alternative explanation could argue that people may generally have the default strategy to look for predictive relations first when asked about the existence

---

[10] Their strength of association was assessed in the USF "Word Association Norm" post-hoc.

of a causal relation. This may be the general default strategy. Therefore, the slower response times for the diagnostic trials would not have been due to the influence of causal directionality, but to a biased access strategy for the predictive direction.

4.1.3 Experiment 3

To rule out that participants primarily interpret causal questions as requests to assess predictive relations, I explicitly specified a requested access direction in the task instructions. In the current experiment I gave participants the task to evaluate whether the items were either predictively or diagnostically related. In two different blocks participants had to judge if the first word was the cause of the second word (predictive block), or if the first word was the effect of the second word (diagnostic block). Again the filler items were weakly associated words. Item pairs in the predictive block were either predictively related or associated, whereas item pairs in the diagnostic block were either diagnostically related or associated. If subjects had used the predictive relation checking strategy in the experiments before, this strategy would now be futile in the diagnostic condition. It was explicitly stated that a successful processing of the task is only possible, if the causal relations were accessed via the explicitly forced direction. The **causal model view** and the **associationistic view** would make two different predictions. The **associative view** would predict no difference of the retrieval speeds between the two causal directions, because the causal relations have a symmetric strength of association. This prediction would lead to the conclusion that the retrieval asymmetries in Experiment 1 and 2 could also be based on a biased access strategy. In contrast, the **causal view** would predict that retrieval speed differs for the two causal directions replicating the previous results.

Methods

Participants and design

28 UCLA Undergraduates with normal or corrected to normal vision were tested. They got course credit for their participation.

The stimuli were presented in a within-subject design: every subject saw every word pair once. The order of the two blocks and the stimuli within a trial (predictive question vs. diagnostic question) were randomized and counterbalanced respectively. There were four

different counterbalanced versions. The type of block: predictive question vs. diagnostic question was the independent variable and the reaction time was the dependent variable.

Stimuli

      The stimulus materials were the same 64 causal and 64 associated filler word pairs, which were also used in Experiment 2. For each block a list of stimuli were created containing 32 causal word pairs in one of the two directions and 32 associated filler word pairs.

Procedure

      50 % of the participants were given the predictive instructions first and the diagnostic instructions second. The other 50% were given the instructions in the opposite order (see Appendix B.3). All subjects were instructed to fixate their eyes on the cross in the center of the screen. After 1 sec the cross disappeared and the screen was blank for 500 ms, then the first word was presented for 1000 ms followed by the second word. In the predictive condition they had to judge whether or not the first word was the *cause* of the second word by pressing the letter "c" on the keyboard. Whereas in the diagnostic condition they had to decide whether or not the word was the *effect* of the second word by pressing the letter "e" on the keyboard. For the non-causally associated filler word pairs they had to press "n". Each subject participated in both conditions. The order of the conditions was counterbalanced. Before each condition subjects were given 10 practice trials (5 predictive, 5 filler in the predictive and 5 diagnostic, 5 filler in the diagnostic condition) on which they got feedback.

Results

Reaction time

      Once again a difference between the predictive and the diagnostic access process could be found. Subjects were able to answer the predictive question faster than the diagnostic question (Figure 17 and Table 7).

Table 7:

Means and standard deviation of reaction times (in ms) for the correct answers in the two different blocks.

Type of Block	predictive	diagnostic
Means	817	886
SD	134	179

A paired sample t-test comparison of the reaction times for the predictive and diagnostic blocks reveals a significant effect. The reaction time is significantly shorter (by 69 ms) in the predictive condition than in the diagnostic condition, $t(27) = 2.35$ and $p = .02$. There is a medium effect size of $d = .44$.
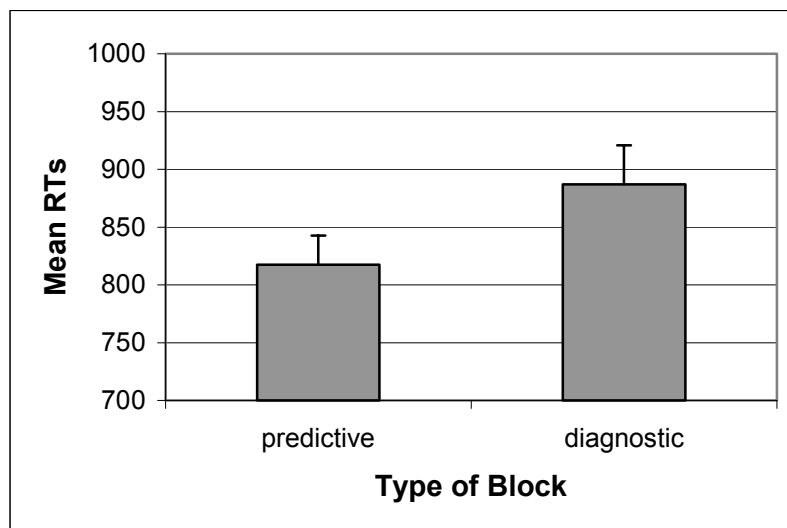
Figure 17: Mean reaction times (+ 1 SE) for the two different forced direction collapsed over all subjects.

Error Analysis

Participants made slightly more errors in the predictive block (Figure 18), but the error analysis reveals that the difference between the predictive and diagnostic blocks is not significant, $t(27) = 0.80$, $p = .43$.
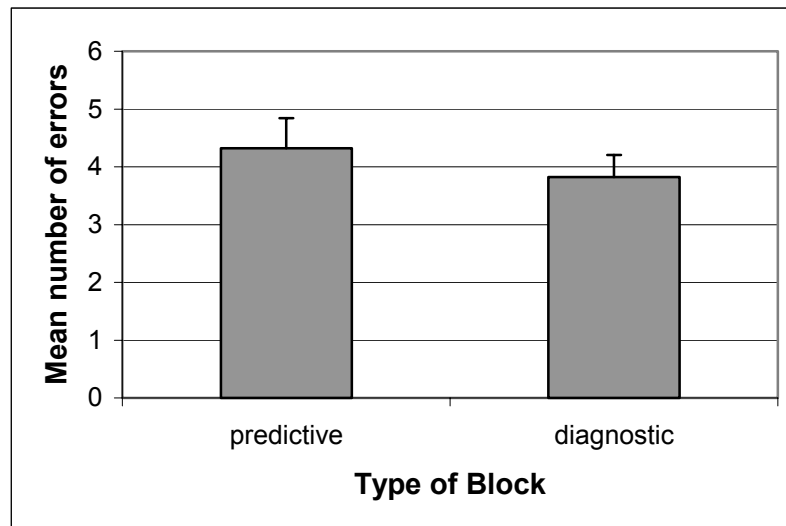
Figure 18: Mean number of errors (+ 1 SE) for the two different forced directions collapsed over all subjects.

Conclusion

Once again the assumptions of the **causal model view** and the **associative view** could be differentiated. The **associative view** would have predicted that the access to causally related but equally associated concepts would not yield a retrieval asymmetry. However, the data support the **causal model view,** which postulates a retrieval asymmetry for the access to causal relations. The results revealed that subjects were faster in determining the first word as the cause of the second word than vice versa. Moreover, the error rates showed that the participants did not favor speed over accuracy, because the difference in the error rates for the two blocks was not significant. Thus, the asymmetries found in Experiments 1 and 2 were not due to a biased access strategy adopted by the participants.

Given the obvious direction of the causal relation, subjects would have had the possibility to generate potential causes of an effect presented in the diagnostic block and vice versa in the predictive block. Applying this strategy, the reaction time would not differ between the two blocks.

On the other hand, if participants had the strategy to determine a causal relation between two words by checking the predictive direction first, their strategy would have failed for this experiment. They knew from the beginning which direction had to be examined. Consequently, they knew that in the diagnostic block the first word was either the effect of the second word or had no causal relationship at all. Subjects were significantly

slower in the diagnostic block compared to the predictive block. Thus, the access to causal relations in semantic memory, cannot be accomplished equally fast for either direction.

An alternative explanation for these data might be that the differences in reaction times are not due to causal asymmetry, but are rather accounted for by difference in access (i.e., processing) stages. Assuming the access to semantic memory is processed in various stages such as the intention to activate a relational concept, the activation of the concept of the relation in question, and the activation and retrieval of the concepts connected by the relation. One criticism would be that the access and the retrieval of causal knowledge do not depend on causal directionality, but the intention to activate a predetermined direction may be influenced by the degree of automation of this process. The intention to activate a predictive causal relation might be more automated than the intention to retrieve a diagnostic causal relation. Consequently, the data in Experiment 3 might be due to different degrees of automation of the intention to access causal relationships in semantic memory. A closer look at the response speed for trials at the beginning and trials at the end of the experiment might provide an answer to that criticism. The degree of automation for the intention to access causal relations in a diagnostic order could increase over the time during one experimental run and therefore result in faster reaction times on the last trials. I calculated a two-tailed t-test over the means of the first five and the last five trials to determine if subjects became faster. I found that the response times for trials at the beginning and for trials at the end of the diagnostic block did not differ significantly $t(27) = 0.37$, $p = .97$. This result lets conclude that the differences in reaction times were not caused by a different degree of automation of the intent to access the discriminative direction of causal relations.

Hence, the effect of causal asymmetry on semantic retrieval is not a consequence of preferences for specific strategies of access or intention activation differences, but is an intrinsic feature of the structure of causal relations in semantic memory. So far, the results support the **causal model view** once again.

4.2   Task specific access to causal semantic memory

In the previous experiments I tested the task specific causal access to causal relations. The results revealed a retrieval asymmetry for equally associated causal word pairs, if a causal relation had to be determined. I attributed these results, in support of the **causal model view**, to a representation of causal directionality in semantic memory. I also argued that the predictive causal direction seems to be more prominent in daily life resulting in a more efficient access of the predictive direction to causal knowledge and a faster processing of it compared to the diagnostic access.

The **associationistic view**, on the other hand, postulates an associative representation of causal relations between concepts and a stronger association of the predictive causal direction than the diagnostic direction. This associative difference would yield a retrieval asymmetry with a faster predictive access. However, I matched the association between the two causal concepts for each causal direction in Experiments 2 and 3, to control for that claim in the causal access task.

Nevertheless, in order to differentiate the assumptions of the **causal model** and the **associationistic view** directly, I need to contrast the *causal* and the *associative* access to the *same* causal relations.

In the associative access to causal relations only an association would need to be determined. The evaluation of an association between two causally related concepts might have different task specifications and demands being independent from causal or other relational constraints. To determine an association would only involve checking, if the presented concepts were related at all, regardless of the type of relation and specific relational properties. In contrast to the causal access, which would require an evaluation of the relational properties (i.e., causal directionality) as well. Consequently, the causal and the associative access specifications would address different features of the connection between concepts.

4.2.1 Experiment 4

In order to differentiate between the causal and associative relational access to stored knowledge, Experiment 4 was designed. It consisted of two different experimental conditions: the *causal condition* and the *associative condition*. Each condition contained causal stimuli and unrelated filler stimuli, but the instructions of the conditions differed (see Appendix A): Subjects in the *causal condition* had to determine, if there was a causal relation between the words, whereas subjects in the *associative condition* had to determine, if there was only an associative relationship between the two words and to avoid that participants based their judgment only on the causal relation between the concepts, I added word pairs that were only associated, but did not have a causal relationship to the *associative condition*. Thus, causation was not the only relation that would provide association, but also other relations like part-whole relationships (e.g., emerald-ring), or category coordinates (e.g., vehicle-bicycle) had to be evaluated based on association. Consequently, the task became more demanding and participants would make significantly more errors only relying on causation, especially in the associated, but not causally related trials.

For Experiment 4, I made the following predictions:

1.   Associative condition

For an *associative* access to causal relations, where the strength of association is equal for both causal directions, *no differences* in the retrieval speed between the predictive and diagnostic causal direction will be found.

2.   Causal condition

For the *causal* access to causal relations with equally associated concepts, a *retrieval asymmetry* will be found, that is, the retrieval of a predictive causal relations will be faster than the retrieval of an associative causal direction.

Methods

Participants and design

44 UCLA undergraduate students participated for course credit. Their vision was normal or corrected to normal. 22 participated in the *causal* and 22 participated in the *associative condition*.

All 44 subjects saw the causal word pairs and the unrelated filler word pairs, but only subject in the *associative condition* saw the associated, but not causally related word pairs. The direction of presentation was manipulated in both conditions. Thus, the direction of presentation and condition were the independent factors and the reaction time was the dependent variable.

Stimuli

For the *causal condition* the stimulus material consisted of 64 causal and 64 unrelated filler word pairs. For the *associative condition* there were 64 causal, 32 associative and 96 unrelated filler word pairs (see Appendix Tables A2 and A4). The causal and the associative word pairs were the same stimuli used in Experiment 2. The unrelated filler word pairs were made up of single words from causal relations from the questionnaire, which were not included in the causal stimuli selection because of too divergent ratings. The words were put together in a random order to avoid any semantic or associative relationship. In order to get a randomized pairing the word pairs were inserted into two columns of an Microsoft Excel® spreadsheet where a random number was generated for each of the pairs in a column next to it, then only one row of the word pairs was sorted by the random numbers. After that the word pairs were checked again to ensure that no relationship whatsoever existed between the two words.

As explained before, the *associative condition* did not only contain causally related, but also associated non-causal word pairs. This addition of associated word pairs resulted in an increase of trials in the *associative condition* (192 trials) compared to the *causal condition* (128 trials) and a confounding of the stimuli and condition. However, this item selection allows a comparison all of causal word pairs over the two conditions, without loosing stimulus material and a distinction between the "yes" and the "no" answers. If in the *causal condition* the filler words were associated, but not causally related, the correct answer for them would have been "no"; in the *associative condition* the answer would have been "yes". Therefore, the type of relations for the "yes" and the "no" answers was held constant between the two conditions.

Procedure

Causal Condition

Subjects were instructed to fixate their eyes on the screen. They had to press the letter "c" on the keyboard, if the words exhibit a causal relation, if not they had to press the

letter "n" on the keyboard (see Appendix B.4). They were given 20 practice trials (10 causal, 10 filler) with feedback.

Associative Condition

Participants had to press the letter "a" on the keyboard, if the words were related in some meaningful way, otherwise they had to press "n" (see Appendix B.4). They received 20 practice trials (10 causal, 10 filler) including feedback (Table 8).

Table 8:

Split-plot design of Experiment 3; Condition is the between subjects variable and the direction of presentation is the within-subject variable.

Number of subjects	Condition	Type of relation in the stimuli (word pairs)	Direction of presentation
22	*causal*	causal and unrelated fillers	predictive vs. diagnostic
22	*associative*	causal, unrelated filler, and associated, but non-causal	

Results

Reaction Times

For the overall analysis only the causal stimuli were considered, because only for them a directional difference was predicted. The analysis of the split plot design for the within-subject factor *direction of presentation* and the between subjects factor *condition* yields a significant effect for the factor *direction* $F(1,42) = 7.74$, $p < .01$, $MSE = 1232.14$, the effect size is $\varepsilon^2 = .17$. There is no significant effect for factor *condition* $F(1,42) = 0.93$, $p = .33$, $MSE = 39282.83$. The interaction for the two factors is significant $F(1,42) = 4.07$, $p = .05$, $MSE = 51749.65$ with a small effect size of $\varepsilon^2 = .08$ (see Table 9 and Figure 19).

Table 9:

Means and standard deviation of reaction times (in ms) for the correct answers in both conditions.

Condition	Causal		Associative	
Direction	*pred.*	*diagn.*	*pred.*	*diagn.*
Means	843	879	817	823
SD	140	160	129	136

The post-hoc with least significant difference method (*LSD*) comparison of the means of the reaction times showed a significant difference between the predictive and diagnostic causal word pairs in the *causal condition, LSD* <0.01. The difference for the causal word pairs in the *associative condition* was not significant, *LSD* = .58.
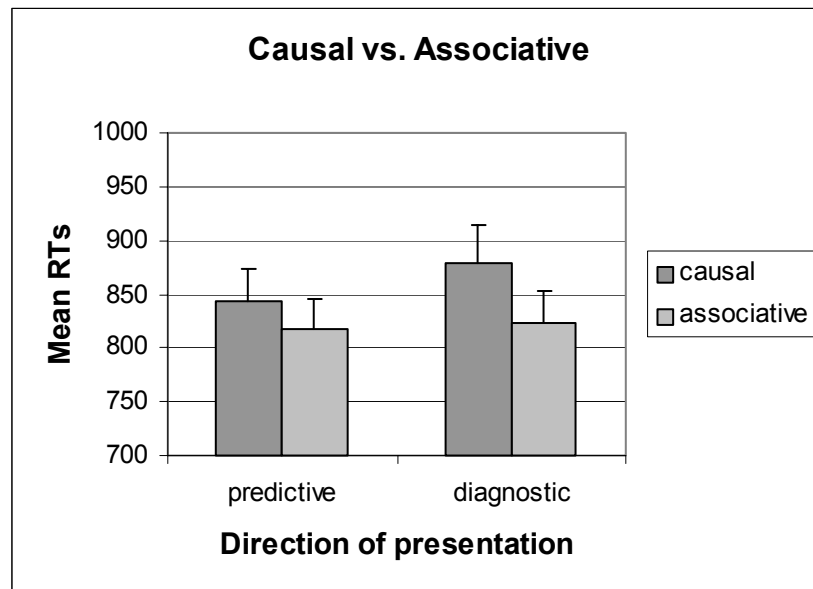
Figure 19: Mean response times (+ 1 SE) for causal stimuli in the associative and causal condition collapsed over all subjects.

Error analysis

The overall analysis for the errors of the causal stimuli exhibits a significant effect for the factor *direction* $F(1,42) = 4.57$, $p = .04$, $MSE = 3.78$, with an effect size of $\varepsilon^2 = .09$, but no significant effect is found for factor *condition* $F(1,42) = 0.16$, $p = .69$, $MSE = 5.85$. The interaction between the two factors is not significant as well $F(1,42) = 0.67$, $p = .42$, $MSE = 3.78$. Subjects made more errors in the overall diagnostic direction for both conditions together (Figure 20). The post-hoc comparison yields a significant difference for

the error rates in the *causal condition, LSD* = .04, but no difference in the *associative condition, LSD* = .35.
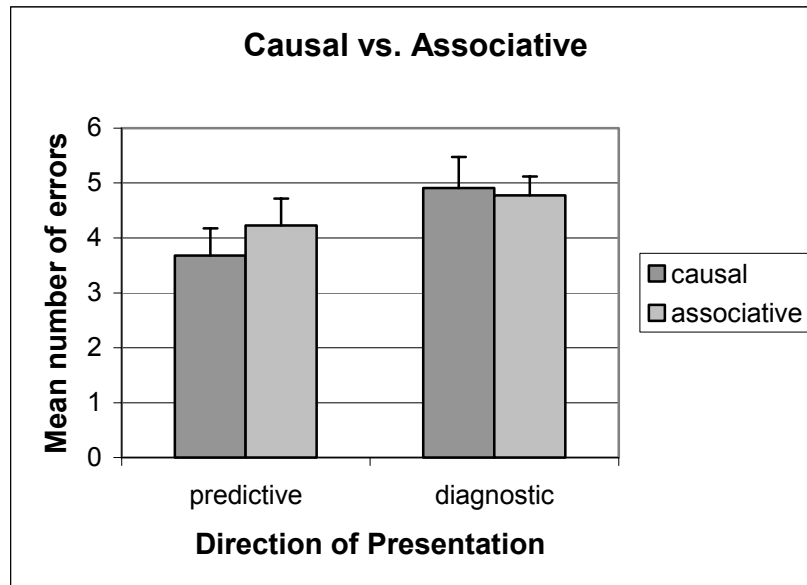
Figure 20: Mean number of errors (+1SE) for causal stimuli in the associative and causal condition collapsed over all subjects.

Conclusion

The overall analysis of the causal stimuli in a split plot design proved the interaction of the two factors *condition* and *direction of presentation* significant. If subjects had to determine a causal relation between the stimuli words, they produced slower response times in diagnostic trials than in predictive trials. For the determination of the associative relation they were equally fast in both directions. The overall response time for the causal and associative word pairs in the *associative condition* showed a tendency to be faster than in the *causal condition*, indicating that the task was also less demanding. The associative access to causal relations seems to be easier and to underlie fewer constraints, which need to be met to determine an association. The error rate also provides support for the two different access processes. In the *causal condition* the error rate was significantly higher for the diagnostic direction, which once again proves that there was no speed-accuracy trade off, and moreover, that the diagnostic direction was more difficult to process. In the *associative condition* the error rate remained the same over all related trials, leading to the conclusion that the *associative* access to causal relations was accomplished independently from the causal direction.

Consequently, the results support the assumption of the **causal model view** that different relations can be selectively accessed in a task-specific fashion. Causal directionality only had an influence if participants searched for the existence of a causal relation, whereas no such effect was found if they focused on the existence of an associative relation. In addition, the results show that causal asymmetry cannot be reduced to asymmetries in the degree of association claimed by the **associationistic view**. Even though, the causal word pairs had an equal strength of association for both causal directions, the retrieval asymmetry was still exhibited in the *causal* access task.

The results also show that people are capable of selectively retrieving relational knowledge. If we know, we have to retrieve a causal relation from memory, we activate abstract knowledge about causation and its characteristics, especially causal directionality. In contrast, if we have to retrieve an associative relationship from semantic memory, we only have to determine, if there is a connection between concepts. Directionality is not important, even though we have to judge a causal relation as an association, asymmetry is not taken into account, because it is not part of the constraints of association, it even contradicts the notion of association.

The stimulus material and the two conditions were confounded: The *associative condition* contained causally related and associated word pairs; the *causal condition* contained only causally related items. These associated, but not causally related word pairs were added to avoid that participants could base their associative judgment solely on the causal connection. This addition of stimuli to only one of the two experimental condition also allowed a comparison between all causally related stimuli word pair and a differentiation between the correct "Yes" and "No" answers.

Nevertheless, a more precise comparison between the causal and the associative access would be to investigate the task specific access to causal relations in a within-subjects design, where each person has to access causal relations in an associative and causal fashion. At this point I want to refer to Experiment 6a, in which the task specific access to causal relations is investigated in a within-subjects design.

Summary of experiments about the relational access to semantic memory.

So far the first four experiments have dealt with causal asymmetry deriving from causal directionality. The data supported the **causal model view** that causal relations are retrieved from semantic memory based on their directionality. If the access is executed in the predictive direction from cause to effect, the reaction time to judge a relationship

between two events as causal is faster than in the opposite direction (Experiments 1, to 3 and 4 *causal condition*). This difference has been replicated for causal word pairs that were matched on strength of association and conditional frequencies for both causal directions (Experiments 2, 3 and 4 *causal condition*). The non-causally related filler items were either not related at all (Experiments 1 and 4) or were weekly associated and matched on their strength of association in both directions (Experiments 2 and 3). In Experiment 4 the task specific relational access was investigated based on an *associative* and *causal* access. The results supported the notion of a task specific access: If a relation has to be judged as causal, causal direction plays a crucial role on relational retrieval (retrieval asymmetry in Experiment 4 *causal condition*), however, if a relation has to be judged as meaningful, causal directionality is not important, even though it is the only relation that provides the association between the two concepts (no retrieval asymmetry in Experiment 4 *associative condition*).

To summarize, the presented results support the postulates of the **causal model view** that a task specific access to causal relations yields an asymmetry of the retrieval speed depending on causal directionality. Supportive evidence for the **associationistic view** could not be found in the empirical investigation of an *explicit* task specific access.

4.3 *Implicit access to causal semantic memory*

As mentioned in chapter 3 there are two approaches to test semantic memory performance: an *explicit* and an *implicit* approach. In the explicit measurement knowledge stored in semantic memory is explicitly *probed* that is, subjects are aware which kind of memory content they have to retrieve. This process is regarded as an *active* search in memory.

In the previous experiments I explicitly probed the participants to access causal or associative relationships in semantic memory. Thus, I investigated the explicit access to causal concepts and relations. So far, I could find differences in response time for the retrieval of causal relations depending on causal directionality: Subjects were significantly faster in determining a causal relationship if the word pair was presented in their "natural" (i.e., predictive) order. However, it could be argued that these differences were not due to causal asymmetries, but rather due to different strength of association. Friedman (1990) argued that causes are stronger associated with their effects than vice versa. Therefore, in the second experiment the stimuli were chosen from an associative word norm database and matched on associative strength in each direction. Nevertheless, even though the word pairs are matched on strength of association, the explicit measure investigated only the task specific access and not the strength of association.

In contrast, the *implicit* measure of semantic memory is regarded to investigate, among other things, the associative strength between two concepts. In the implicit access to memory the activation of semantic memory contents is a *passive* and subconscious process. If a concept is activated, for example while reading a text, other concepts connected to it will get activation as well. The activation is said to spread from one concept to the other (Collins & Loftus, 1975). The speed and the amount of activation depend on how close the concepts are in semantic space and on how many concepts are connected to the activated concept. The closeness in semantic space can be determined by strength of association between two concepts. The influence of that activation on the processing of connected concepts is called *priming* (Neely, 1977). In the priming section of chapter 3, I reported two different types of priming: *associative* and *semantic* (i.e., *relational*) priming. Associative priming is regarded to rely on associative strength between two or more concepts and is typically investigated in short (< 500 ms) SOA settings; *relational* priming is regarded to involve deeper processing of the semantic relation between two or more concepts. Relational priming is investigated in long (> 500 ms) SOA conditions (Moss et al., 1995),

and has been reported for category coordinates (Shelton & Martin, 1992) and functional relationships (Moss et al., 1995). Causal relations so far have not been investigated in a relational priming paradigm, therefore, the question if causal relations influence priming has not been answered yet.

The investigation of associative priming for causally related word pairs would help to find differences based on the associative strength between the concepts, postulated by the **associative view**. That is, according to Friedman (1990) causes would prime their effects more than vice versa. Given that the selected causal stimuli are matched on associative strength for each direction, the degree of associative priming would be an indicator if the selection was precise.

Accordingly, the investigation of the implicit access to causal relations would serve two purposes: 1) Reconfirming that the selected stimuli have indeed an equal strength of association between the two directions and 2) Investigating the possibility of relational priming for causal relations.

4.3.1 Experiment 5

In this experiment I tested the causal word pairs in a priming paradigm. The causal word pairs were presented in the predictive and the diagnostic direction, but instead of determining a causal relationship between the two words, subject only had to pronounce the second word. Hence, no explicit access to stored semantic memory content was required.

For this experiment the following assumptions need to be made about priming in order to interpret possible results:

1. Priming is a very sensitive measure of the degree of association. If there is no association between the first and the second stimulus, priming would not occur.

2. Different SOAs can distinguish between associative and semantic priming. If indeed, there was a difference in the strength of association (i.e., higher in the predictive direction), more priming should occur for the predictive direction.

In order to account for associative and semantic priming I used two different SOA conditions: a short (250 ms) and a long (1000 ms) SOA condition.

Depending on the results different interpretations could be given:

1. There is neither associative nor relational priming for the causal word pairs:

There was neither an associative nor a semantic connection between the causal word pairs.

2. Priming is found in both SOA conditions:

a)   there is a *difference* between the predictive and diagnostic direction in both conditions.

There is indeed an asymmetry in the associative strength of the selected stimuli and therefore differences of the previous experiments might be a compound of strength of association between the two words and probably of semantic relations. However, a clear differentiation between associative and semantic relations would be difficult.

b)   there are *no differences* in the amount of priming between the predictive and diagnostic direction.

It could be concluded that the strength of association had no influence on the explicit retrieval time. Moreover, no difference in the long SOA condition would indicate that there is *no* relational priming for causally related concepts.

c)   there is a *difference* in the short SOA, but *no difference* in the long SOA condition depending on the presented direction.

The differences in the previous experiments could be attributed to the different strength of association between the two directions and not because of causal asymmetries.

d)   there is *no difference* in the *short* SOA condition, but there is a *difference* in the long SOA condition between the two directions.

The stimuli have an equal strength of association for both causal directions. The difference in the long SOA condition would indicate relational priming for causal concepts. The naming latency was influenced by causal directionality.

The four assumptions are predicted differently by the **causal model** and the **associationistic view** (Table 10).

Table 10:

Summary of assumptions and possible interpretation of results in Experiment 5.

	Priming condition		**Interpretation**
	short SOA (250 ms)	**long SOA (1000 ms)**	
Order of presentation	*predictive vs. diagnostic*	*predictive vs. diagnostic*	
a)	priming difference	priming difference	Different strength of association and semantic connection. → **Associative view**
b)	no priming difference	no priming difference	Equal strength of association and semantic connection, but no relational processing. → **Causal model view**
c)	priming difference	no priming difference	Different strength of association. → **Associative view**
d)	no priming difference	priming difference	Equal strength of association and relational processing. → **Causal model view**

Methods

Participants and Design

48 Students of the University of Göttingen participated for course credit. Their vision was normal or corrected to normal. 24 participated in the short SOA condition and 24 participated in the long SOA condition. They were randomly assigned to the different conditions. The experiment was a 2x2x2 split-plot design with factors *SOA*: 250 ms vs. 1000 ms being the between-subjects factor and factors *priming*: causal vs. neutral and *direction*: predictive vs. diagnostic being the within-subject variable.

Stimuli

The stimuli were the translated German version of the causal word pairs used in the other experiments (see Appendix Table A5). They were presented in "Arial Black" and font

24 on a white background. The experiment was programmed in Superlab ® and run on a Pentium II processor with a 17" screen and a 1026 x 768 pixel resolution and 256 colors.

The neutral prime was a string of asterisks the same font size and color.

There were 54 causal prime word pairs and 54 neutral prime word pairs. The number of stimuli was smaller in the German version, because some of the English words could not be translated one to one (e.g., bruise-> blauer Fleck).

Procedure

Participants were given written instructions (see Appendix B.5) indicating that they will see a fixation cross in the center of the screen and then either a word (causal prime) or a string of asterisks (neutral prime) for 250 ms (in the short SOA condition) or for (1000 ms in the long SOA condition), after that a second word will be presented. The participants had to pronounce the second word as quickly as possible (Figure 21). The neutral prime pairs and the causal prime pairs were presented randomly. The direction of the causal prime was counterbalanced, 50% of the trials were in predictive direction, thus subjects had to pronounce the effect word and the other half were in the diagnostic direction, subjects had to pronounce the cause word.
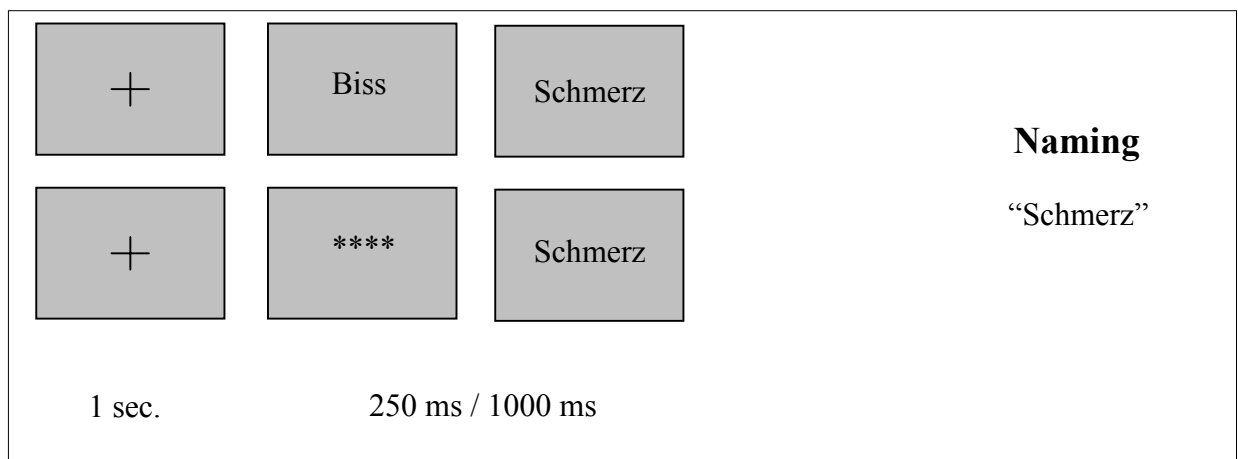
Figure 21: Two predictive trials. The upper row represents a causal prime trial, the lower row represents a neutral prime trial. The SOA was either 250 ms or 1000 ms.

The naming latencies were recorded via a headphone-microphone set connected to the computer's soundcard. Each subject responded to 108 trials. At the beginning subjects were given 10 practice trials (5 neutrally and 5 causally primed word pairs) to become accustomed with the task.

Results

Naming latencies

The data analysis yields a significant effect of the factor *priming*, $F(1, 46) = 8.33$, $p > .01$, *MSE* = 1117.95 with an effect size $\varepsilon^2 = .15$. Participants could pronounce causally primed words faster than neutrally primed words (see Table 11 and Figure 22).

Table 11:

Means and standard deviations for the two within-subject factors priming and direction collapsed over all subjects and the two SOAs.

Priming condition	Causal		Neutral	
Direction/2nd word	*cause*	*effect*	*cause*	*effect*
Means	580	585	592	601
SD	110	119	121	129

There is no significant effect of the factor *SOA*, $F(1, 46) = 1.64$, $p = .21$, *MSE* = 54396.43 and no interaction between the factors *SOA* and *priming*, $F(1, 46) = 0.29$, $p = .58$, *MSE* = 1117.95. The stimulus onset asynchrony has not effect on the naming latencies and on the amount of priming. The factor *direction* is significant, $F(1, 46) = 4.41$, $p = .04$, *MSE* = 538.62, $\varepsilon^2 = .08$. There is not interaction for the factor *direction* and *priming*, $F(1,46) = 0.23$, $p = .63$ and *MSE* = 465.21 and no three-way interaction of the factors: *priming*, *direction* and *SOA*, $F(1, 46) = 0.00$, $p = .98$, *MSE* = 465.21. The post-hoc comparison reveals a significant difference for the causal priming in the predictive word pairs, *LSD* < .01, and for the diagnostic word pairs, *LSD* = .02. Consequently, priming was equally strong in each direction independently of the two SOA conditions. The post-hoc comparison between the neutral predictive priming and the causal diagnostic priming is also significant, *LSD* < .01, which is responsible for the significant p-value of the factor *direction*.
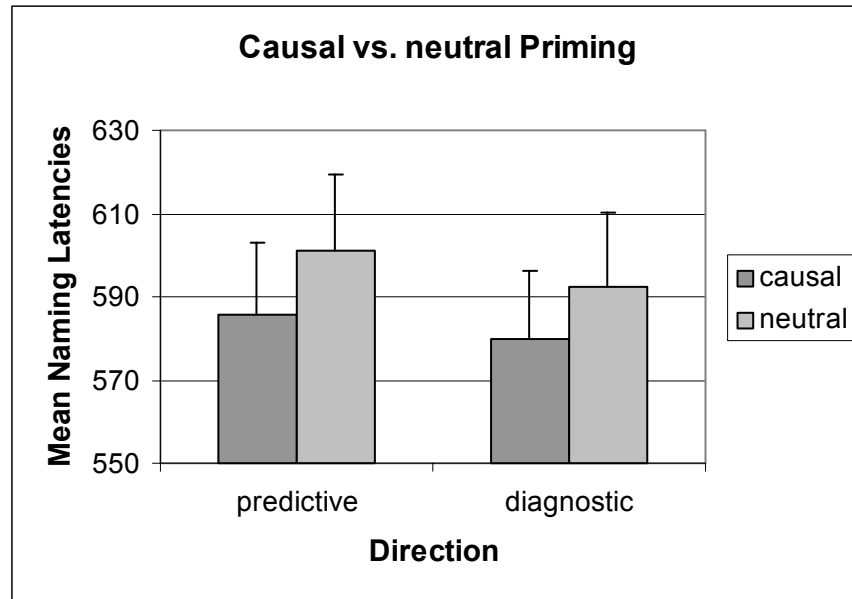
Figure 22: Mean naming latencies (+1 SE) for the factors priming and direction collapsed over all subjects and the two SOA conditions.

Conclusion

The results showed a priming effect for causally primed word pairs compared to neutrally primed word pairs in both SOA conditions. This means that there was *associative* priming in the short SOA condition, but no *relational* priming in the long SOA condition.

For the *associative* priming it can be concluded that the causally related word pair had an equal strength of association for the two causal directions. This supports the argument that the stimuli in Experiments 2 and following are matched on their strength of association, and moreover, that the results from the previous experiments can be attributed to causal asymmetries rather than associative strength. Consequently the priming results contradict Friedman's (1990) **associative view** that events along a timeline necessarily have a higher strength of association in the forward direction and therefore, causal relations have a higher strength of association in the predictive direction. In addition, the results also support the task specific access tested in Experiment 4. If subjects explicitly had to detect a causal relation, causal asymmetry became important, however, if they only had to derive an association the asymmetry vanished as postulated by the **causal model view**. In the priming paradigm subjects were not explicitly told to judge a relation between two concepts or to generate a relation, they only have to pronounce the second word. Thus, they were completely oblivious to the relationship between the two words passively activated along the connection of the two concepts in semantic memory.

In the long SOA condition investigating *relational* priming no differences between the two causal directions could be found. These data, however, seem to support the **associationistic view** and may lead to the preliminary conclusion that causal relations do not belong to the group of semantic relations that influence priming and thus could be reduced to associations. However, a pronunciation task is regarded to be less sensitive to *relational* priming compared to a lexical decision task where the probability of a target word being semantically related to the prime is confounded with the probability that the target is a non-word, given that it is unrelated to its prime (Neely, Keefe & Ross, 1989). That is, if the target word is related to its prime, it necessarily has to be word and cannot be a non-word. The *relational* priming results reported by Moss et al. (1995) and others (Shelton & Martin, 1992) have been found in lexical decision tasks. However, the same stimuli have not been tested in a pronunciation task. Given the similarity between causal relations and functional relation, for example a match has the *function* to *cause* fire, *relational* priming for causal relations might be more likely to occur in lexical decision task in which the relation between the prime and the target word facilitates the lexical decision.

To summarize, Experiment 5 supports the notion that the *implicit* access to semantic memory is independent from the *explicit* access. Moreover, it confirms that the selected causal word pairs are equally strong associated in both causal directions and therefore it is another piece of evidence that the differences in the previous experiments are due to a task specific access process and not to differences in the associative strength between the predictive and diagnostic direction. Consequently, the distinction of the **causal model** and the **associationistic view** seems also possible by using techniques addressing the *implicit* access to semantic relations.

After I covered the *explicit* and *implicit* access processes in behavioral experimental designs, I now want to turn an experimental paradigm that investigates the relational access to semantic memory on the basis of cerebral blood flow. The final functional magnetic resonance imaging (fMRI) study reported in my dissertation should be understood as pilot studies to further investigations. For the purpose of structure and transparency of my dissertation, I positioned it at the end of this chapter.

4.4   The neuropsychological basis of the task specific access- Pilot study

In this study, I wanted to examine the activation of different brain regions during the relational access to semantic memory (see Experiment 4). By applying a functional magnetic resonance imaging (fMRI) measure to the task specific relational access I wanted to explore other potential differentiations between the **causal** and the **associative view**.

The assumption of the fMRI measurement is that a stimulation of the brain changes the metabolism in specific brain areas responsible for the stimulus processing. As a result the magnetic property of the blood will change in that area. Hemoglobin is that substance in the blood responsible for this magnetic property. The property can change depending on the presence or absence of oxygen binding (Oxyhemoglobin and Desoxyhemoglobin). If an area is active more oxygen containing blood is transported to that region. The fMRI scanner induces a magnetic field around the head. This field can make the task specific changes in the blood oxygen level visible.

The idea to apply an imaging paradigm to my investigations was inspired by an article by Martin and Chao (2001) about semantic memory and activated brain regions. They reported that during semantic memory performance task the left hemisphere especially the anterior and inferior prefrontal regions are involved selectively in semantic processing. In addition Beeman (1991) found that the left hemisphere activates only a restricted array of information directly related to the activated concept, whereas the right hemisphere is important for the comprehension of discourse by activating a larger array of information including more remotely related information (Beeman, Friedman, Grafman, Perez, Diamond & Lindsay, 1994), the maintenance of thematic coherence (St. George, Kutas, Martinez, & Sereno, 1999), context information (Atchley, Burgess, Keeney, 1999) and the interpretation of context in text comprehension (Just, Carpenter, Keller, Eddy, & Thulbon, 1996). Causality can be understood as a context based on which relations between concepts can be interpreted. That is, causal relations are context dependent. As could be seen in Experiment 4 causal relations could be accessed differently depending on the framing of the task (causal vs. associative). In the *associative condition* the context on which the association could be determined was quite flexible. The concepts had to be related (i.e., associated) in some way, whereas in the *causal condition* the context was causality. Mason and Just (2001) found activation in the right dorsolateral prefrontal cortex (DLPC) during the comprehension of causality in sentence processing. Another study about inductive and deductive reasoning

discovered a logic-specific network of rule based decision making including the right frontal cortex, thalamic nuclei and basal ganglia (Parsons & Osherson, 2001). Parsons and Osherson (2001) could show that *deductive* reasoning activated the right hemisphere more, whereas *inductive* reasoning tasks activated the left hemisphere more (Table 12.

Table 12:

Examples of a deductive and an inductive reasoning task.

Deductive reasoning task	*Inductive* reasoning task
(1) If he is an electrician then he spent two years in night school.	(1) If he is a heart specialist then he either bicycles to work or swims regularly.
(2) He is an electrician and owns a computer.	(2) He is a heart specialist.
(3) He spent two years in night school.	(3) He bicycles to work.

This distributed activation supposedly reflected the integration of information in that specific reasoning process. Parsons and Osherson (2001) claim that the right frontal lobe is part of the semantic memory system where information for the *reasoning* task is maintained and manipulated. In contrast to the left hemisphere where a *syntactic* and *grammar* rule based network includes the left inferior frontal lobe, thalamic nuclei and basal ganglia. This left "sided" network is responsible for the syntactic integration of verbal information.

In an EEG study about processing temporal discourse while reading a sentence, the activation of the left frontal lobe was associated with verbal working memory (Münte, Schlitz, & Kutas, 1998). In this study sentences were presented. These sentences were identical, except, the first word differed. The sentences started either with the word "*before*" or "*after*". For example, "*Before* the scientist submitted the paper, the journal changed its policy" and "*After* the scientist submitted the paper, the journal changed its policy" Münte et al. (1998) found an increased left-frontal negativity after 300 ms for sentences starting with "*before*". They argued that these sentences required maintaining more information in working memory, because the temporal information given by the first word indicated the temporal order of the following clauses. The word "*before*" indicated that the last clause of the sentence was actually temporally prior to the first clause. The temporal order of the events was incongruent to the order of the clauses in the sentence. This incongruent condition posed a higher demand on information integration increasing verbal working memory load. This increment also increased the activation in parts of the brain, where verbal working memory is supposed to be

processed. This task shows some similarity to the predictive and diagnostic distinction of causal directionality, because the integration of a diagnostic causal relation also requires the processing of an incongruent condition: The effect is presented prior to the cause. Therefore, in the Münte et al. (1998) study sentences starting with "*after*" would correspond to the *predictive* causal direction, whereas sentences starting with "*before*" would correspond to the *diagnostic* causal direction.

These interesting studies motivated my fMRI investigation. Hence, in a first exploratory attempt to make visible the access to causal relations in the brain the pattern of cerebral activation is investigated during a causal retrieval task with a focus on the left and right inferior frontal lobes.

4.4.1 Experiments 6a and 6b

As in Experiment 4, I compared two tasks, an *associative* and a *causal* retrieval task. Cause-effect relationships stored in semantic memory had to be accessed via two different semantic relations: *causation* and *association*. The hurdle that needed to be taken was the application of the task used in Experiment 4 to a fRMI scanner. The completely randomized split-plot design had to be given up and transformed into a block design where each block contained only the same type of trial and without a between-subjects variable. In a single trial design with a random presentation order the required rest-interval of about 16 sec between each trial would have exceeded the reasonable time period for subjects in the scanner. This rest-interval ensures that the brain activity after processing a given trial returns to the baseline. In a block task, the rest-interval is only given between each block and not between the trials belonging to the same task category in a block. For example, the task for one block is to determine a causal relation, all the trials within this block needed to be processed with regards to causation. Thus, subject's brain is in "causation mode" for that block. For the next block the task might be different, thus the brain has to switch to a different mode. Consequently, it is important that there is a resting period between each block allowing brain activity to return to baseline between different tasks. In a randomized single trial presentation, the task switches between trials making the resting period between each trial necessary.

Experiment 6a was conducted as the pilot study for the fMRI experiment. No brain activities were measured in this experiment, and only the reaction time data were of

relevance. This pilot study was necessary to assure that the behavioral results found in the first two experiments could be replicated in a block design.

Methods Experiment 6a

Participants and design

 24 UCLA undergraduate students participated for course credit. Their vision was normal or corrected to normal.

 The stimuli were presented in a within-subject design: every subject saw every word pair once. The order of the blocks and the stimuli within a given block (predictive vs. diagnostic) were randomized and counterbalanced respectively. Thus, there were two independent factors: *direction of presentation* (predictive vs. diagnostic) and *condition* (causal vs. associative). The reaction time was the dependent measure.

 Moreover, given the results of Experiment 3 a significant contrast was predicted for the causal direction in the causal condition, but not for the associative condition.

Stimuli

 The stimulus material consisted of 64 causal, 16 highly associative, but non-causal and 30 unrelated filler word pairs. The filler word pairs shared no relationship. The causal word pairs were the same as in Experiment 2. The highly associated word pairs (strength > 0.3) were selected from the USF database (see Appendix Tables A2 and A3). The reason to choose highly associated word pairs was to test for potential differences in strength of association between the causal and associated word pairs.

Procedure

 Subjects were instructed that they would see 10 blocks each containing 11 different trials. Before each block they were prompted by either the word "CAUSAL?" or the word "ASSOCIATIVE?" If they were prompted with the word "CAUSAL?" then they had to decide whether or not there is a causal relation between the two words. If the prompt word was "ASSOCIATIVE?" then they had to decide whether or not these words were related in a meaningful way. In both cases participants had to press the letter "y" otherwise they pressed "n" (see Appendix B).

 After the practice trials, subjects started with the experimental blocks. They began with the first prompt (causal or associative). They saw 11 trials per block. If they were prompted to look for causal relations, 8 of these 11 trials were causal and 3 of these had no

semantic relation at all. There had to be filler items in the block to assure that participants really processed the prompted relation and were not just pressing the "y" key. However, the number of the unrelated filler had to be reasonably small to be disregarded in the analysis and to get sufficient target relations in one block. The order of the trials within a block was randomized. There were two different kinds of causal blocks: predictive (spark->fire) or diagnostic (fire->spark), but the trials within a given block had the same causal direction. After each block subjects received a 16 sec break. Then a new prompt was presented. If the prompt was to look for associations, then 8 word pairs contained a meaningful relationship (semantic association) and 3 filler word pairs exhibited no relationship at all. The relationships within an associative block were either causal (spark->fire) or only associative, but not causal (bread->butter). The causal trials within an associative block were also either predictive or diagnostic. Subjects saw 10 blocks with two blocks containing the same kind of trials: Blocks 1 and 3 contained causal word pairs in the predictive direction, Blocks 2 and 4 contained causal word pairs in the diagnostic direction and Block 5 contained associated word pairs (Figure 23). The order of the block presentation was completely randomized

Subjects also had to answer 26 (13 causal, 13 associative) practice trials first.
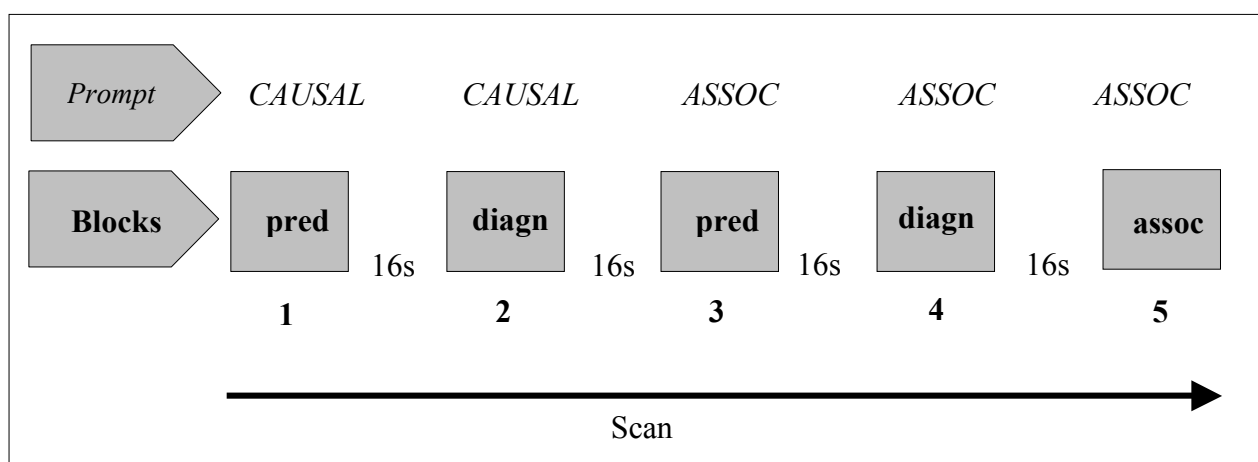
Figure 23. 5 Experimental Blocks showing the type of stimulus material and the type of prompting. The presentation order of the blocks was randomized for all subjects.

Results Experiment 6a

Reaction times

An overall repeated measurements design yield a statistically significant difference between all five types of block, $F(4,92) = 10.15$, $p < .01$, $MSE = 792.58$ with an effect size of $\varepsilon^2 = .31$. The planned contrast analysis shows that there is a significant difference between the causal word pairs in the causal blocks (1 and 2), $F(1,23) = 6.42$, $p = .02$, $MSE = 8696.80$.

The difference between the causal word pairs in the associative blocks (3 and 4) is not significant, $F(1,23) = 0.25$, $p = .62$, $MSE = 8623.48$ (Table 13 and Figure 24). The interaction between the *condition* (causal vs. associative) and the *direction* (predictive vs. diagnostic) was not significant, $F(1,23) = 2.09$, $p = .16$, $MSE = 13676.64$.

Table 13:

Means and standard deviation of reaction times (in ms) for correct answers in the five different conditions.

Prompt	CAUSAL		ASSOCIATIV		
Block	**pred.** **1**	**diagn.** **2**	**pred.** **3**	**diagn.** **4**	**assoc.** **5**
Means	872	940	865	878	762
SD	145	163	177	159	116
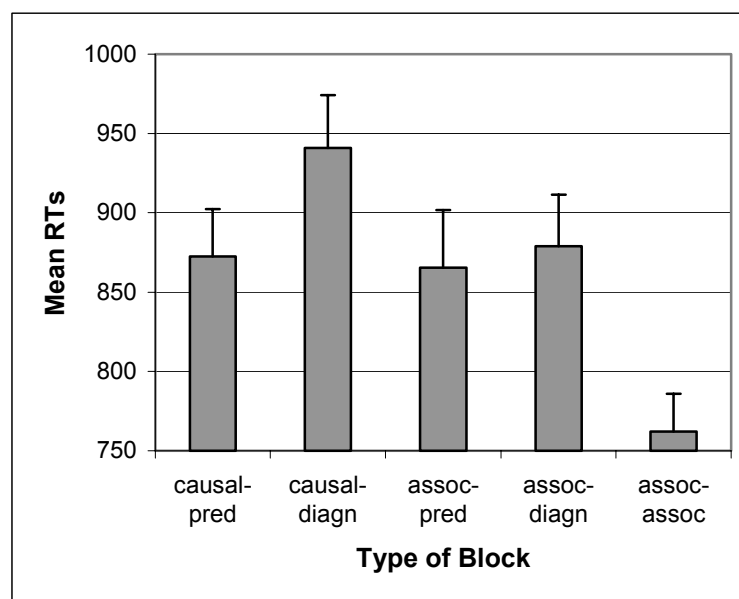
Figure 24: Mean reaction time (+1 SE) for the five different blocks collapsed over all subjects.

The contrast reveals that the reaction time for the predictive trials in the *causal condition* is significantly lower (872 ms) than for the diagnostic trials (940 ms). The reaction times for the causal stimuli in the *associative condition* did not differ for predictive and diagnostic trials.

Error Analysis

The overall analysis of errors shows a significant difference for the error rates in the five different type of blocks, $F(4,92) = 4.85$, $p < .01$, $MSE = 149.96$. The error rate was the

lowest for the highly associated word pairs. There is no significant difference found for the factors *direction of presentation*, $F(1,23) = .06$, $p = 81$, $MSE = 35.65$ and *condition,* $F(1,23) = 1.15$, $p = .29$, $MSE = 59.74$. The interaction of these two factors is not significant $F(1,23) = 0.01$, $p = .92$, $MSE = 29.74$ (see Figure 25).
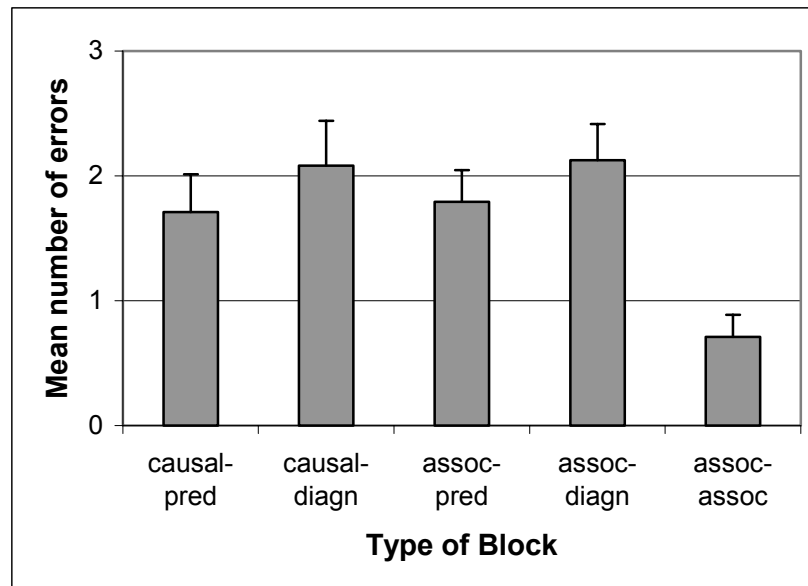
Figure 25: Mean number of errors (+1 SE) for the five different blocks collapsed over all subjects.

Conclusion Experiment 6a

The major advantage of this design compared to Experiment 4 was that there was no confounding between the stimuli and the condition allowing a direct comparison between the associative and the causal access to the *same* causal relations.

The results show a clear advantage for highly associated items over weakly associated items. The mean reaction time for the highly associated items was much faster than for the weakly associated word pairs. The causal, but weakly associated items produced a much slower RT even for the associated question. Once again the influence of strength of association was proven impressively. Moreover the lower demand in the *associative condition* and the block design of the experiment could have increased the prominence of the effect. The error rates support the lower demand argument. The associated stimuli in the *associative condition* were easier to judge.

In the *causal condition* once more the former pattern could be shown: The causal access to causal relation in semantic memory is faster if the word pairs are presented in the "natural" order from cause to effect. In the *associative condition* the causal stimuli did not elicit a reaction time difference for the relational judgment. Subjects were able to determine

the associative relationship between the cause-effect words equally fast for both directions. Consequently, it seems plausible to argue that the access to cause-effect relationships in semantic memory depends on the type of relation. If we have to judge if two concepts in semantic memory are connected via a causal relation (i.e., one causes the other or vice versa), the judgment is guided by the directionality of causal events. However, if we have to judge if there is only an associative connection (i.e., a meaningful connection) between cause-effect concepts, we will disregard the directionality of causal events being not part of the task. These results contradict the approach taken from the **associationistic view** postulating no influence on retrieval.

Given these results, this experimental paradigm was now to be tested under fMRI scanning conditions. The fMRI experiment was understood as a first pilot study to further neuropsychological experiments about causal knowledge. It had a more explorative character providing a first idea about the brain areas involved in the retrieval process of causal relations. As mentioned before it was hypothesized that the right frontal lobe has a significant function in the retrieval and integration of semantic memory contents, therefore the main focus was on the activation of the right inferior frontal lobe and possible hemispheric asymmetries depending on the integration of causal semantic knowledge.

Methods Experiment 6b

Participants and design

12 UCLA graduate students participated in the fMRI setting. They were paid $ 20. Their vision was normal or corrected to normal.

The same design was used as in Experiment 6a and the same contrasts were planned.

Stimuli

The stimulus material consisted of 64 causal, 16 associative, and 30 unrelated filler word pairs. The filler word pairs shared no relationship. The causal and associative word pairs were the same as in Experiment 2. Thus in contrast to Experiment 6a, the strength of association between the associative word pairs was as weak as for the causal word pairs. The reason to choose the weakly associated words for this experiment was to increase the demand and to decrease the contrast to the other stimuli in the *associative condition*.

Procedure and MRI acquisition

The subjects had to read and sign a consent form. Then, they were given the written experimental instructions. These instructions were similar to the one in Experiment 6a, but altered in order to fit the pattern of a response box, which was given to the subjects into the scanner (see Appendix B.6).

The scans were conducted in the 3 Tesla magnetic resonance imaging system (General Electric, Waukesha, WI) at the UCLA Brain mapping center. Axial high-resolution spin-echo scans of the whole brain were collected in the same plane as the same plane as the functional scans to aid group registration (26 slices; 4 mm slices, 1 mm gap; 128 x 128 matrix; TR: 4 s; TE: 54 ms). This T1 weighted structural scan was accomplished first to acquire an anatomical image of the brain. No functional data were collected in the structural scan. Subjects then had to process the practice trial, on which they received feedback about their performance. After the practice trials, the main experiment and the functional scanning began. For each of the 10 blocks T2 weighted functional scans were collected (16 slices; flip angle = 90°; 4 mm slices, 1 mm gap; 64 x 64 matrix; TR: 4 s; TE: 25 ms). Subjects started with a blank screen for 16 seconds, then the first prompt was presented for 4 seconds, after that the cross was presented for 1 second, followed by a 500 ms blank screen. The first word was shown for 1 second immediately followed by the second word for 1.5 seconds. Thus, a whole trial lasted 4 seconds. Each block consisted of 16 seconds blank, 4 seconds prompt, 11 x 4 seconds of trials. Accordingly, each block lasted 1:04 minutes. After 5:36 minutes, the first run of 5 blocks was finished and subjects were given a pause. After the pause the second run started with the second set of 5 blocks. In each run each type of block (1 to 5) was presented in a randomized order.

Image Post-Processing

The data from the fMRI scans were motion-corrected and warped into a site-specific atlas in Talairach space (Talairach & Tournoux, 1988), using AIR software (Woods, Grafton, Watson, Sicotte & Mazziotta, 1998). The data were smoothed with a Gaussian filter of 6 mm full-width half maximum. They were further analyzed in SPM99 (Wellcome Dept. of Cognitive Neurology, http://www.fil.ion.ucl.ac.uk/spm/) where all contrasts were run using an uncorrected threshold of $p = .001$ and a 'k' extent of 4 voxels with a voxel size of 2.0 x 2.0 x 2.0 mm. The planned contrast included the comparison of the *causal* and *associative condition*. The contrast analysis was based on subtraction of the associative blocks from the causal blocks and vice versa.

Results Experiment 6b- Behavioral Data

Reaction time

The analysis of the reaction time in the behavioral data shows a significant effect between the five blocks, $F(4,44) = 3.33$, $p < .01$, $MSE = 4539.59$ with an effect size of $\varepsilon^2 = .23$ (Figure 26 and Table 14).

Table 14:

Means and standard deviation of reaction times (in ms) for correct answers in the five different conditions.

Prompt	CAUSAL		ASSOCIATIVE		
Block	**pred.** **1**	**diagn.** **2**	**pred.** **3**	**diagn.** **4**	**assoc.** **5**
Means	923	925	870	911	845
SD	76	132	139	116	91

The planned contrasts show no difference between the two causal directions in the *causal condition*, $F(1,11) = 0.01$, $p = .93$, $MSE = 3505.34$ and in the *associative condition* $F(1,11) = 2.03$, $p = .18$, $MSE = 5082.49$. The interaction of the two factors is also not significant, $F(1,11) = 1.14$, $p = .31$, $MSE = 4042.35$.
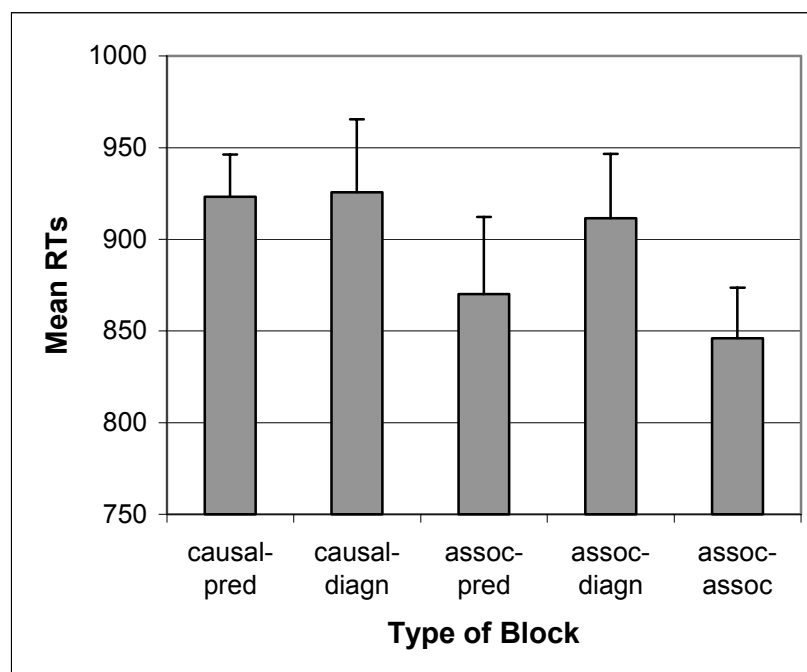
Figure 26: Mean reaction time (+ 1 SE) for the five different blocks collapsed over all subjects.

Error analysis

The overall error analysis reveals a significant difference between the five blocks*, F*(4,44) = 2.71, *p* = .04, *MSE* = 2.69 with an effect size ε^2 = .19 The factors *condition* is significant*, F*(1,11) = 8.98, *p* = .01, *MSE* = 2.83 and ε^2 = .45, but the factor *direction of presentation* does not yield any significance*, F*(1,11) = 0.6, *p* = .79, *MSE* = 2.67. The interaction of these two factors is not significant as well*, F*(1,11) = 0.18, *p* = .67, *MSE* = 2.74.
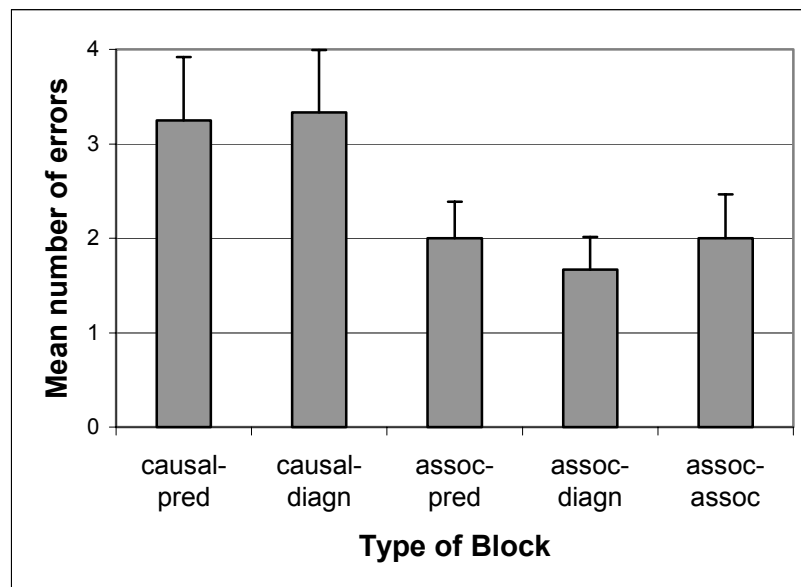
Figure 27: Mean number of errors (+1 SE) for the five different blocks collapsed over all subjects.

As can be seen in Figure 27, participants make fewer errors in the *associative conditions* compared to the *causal conditions*. The difference between the causal directions in the causal blocks is not significant, neither is the difference between the associative blocks.

Results Experiment 6b- fMRI Data

The analysis of the blood flow data was designed to locate regions active during the processing of causal relations. The results reveal a rather unilateral right frontal activation during the processing of the causal relations (Figure 28 and Table 15).
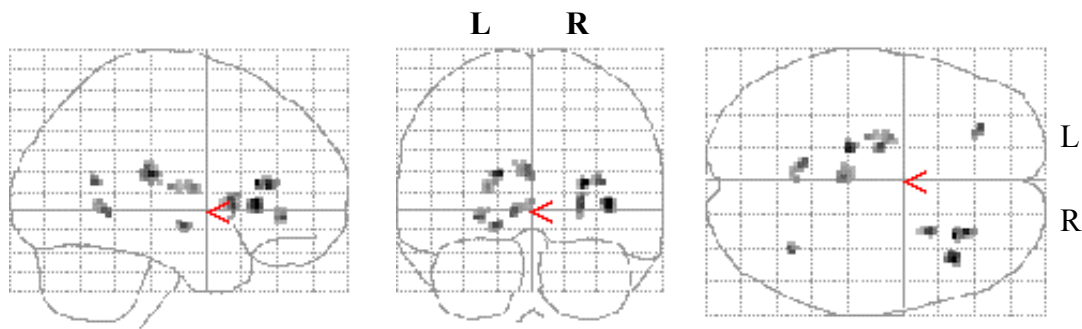
Figure 28: Brain activation during processing of causal relations.

Table 15 shows the activated areas for the subtraction of the associative activation from the causal activation. K extent describes the size of the activated cluster, for example, a K extent of 20 means that 20 voxels were activated forming a cluster. The Maxima (t) describe the t-values found in the contrast calculation for a given cluster. On the far right side of the table the position of the activation in the standardized brain is reported in Talairach coordinates: x-axes: right to left; y-axes: front to back (anterior to posterior), and z-axes: top to bottom (superior to inferior). The opposite contrast, the subtraction of the causal activation from the associative activation at the same threshold, did not reveal any significant activational differences.

Table 15:

Areas of activation obtained for a subtraction of causal minus associative relation processing. Anatomical location, extent of activation (voxels), maximum t value, and Talairach coordinates are provided.

Contrast: Causal-Associative

#	Location	K extent	Maxima (t)	x	y	z
1	Right inferior frontal gyrus	20	6.86	34	22	3
2	Right frontal lobe, sub-lobar	23	6.64	24	22	13
3	Left Thalamus, Pulvinar	12	6.37	-15	-30	14
4	Right lentiform nucleus, Putamen	32	5.95	22	10	2
5	Left limbic lobe, parahippocampal gyrus	11	5.80	-17	-15	-8
6	Left inferior frontal gyrus	15	5.77	-25	35	-2
7	Right parahippocampal gyrus	11	5.64	31	-58	9
8	Left lingual gyrus	16	5.53	-8	-54	-3
9	Left Thalamus	30	5.36	-3	-32	14
10	Left lentiform nucleus, Putamen	21	5.05	-20	-9	10
11	Left Thalamus, vent. post. lat. nucleus	21	4.96	-20	-21	9

The processing of the causal relations yields a bilateral frontal gyrus activation, but for the right hemisphere was located more lateral (near BA 13)[11] and for the left frontal lobe it was located more medial closer to BA (47). There was also an additional activation sub-lobar and for the lentiform nucleus in the right frontal cortex. The causal processing also showed activation in the right parahippocampal gyrus near BA (30) and the left parahippocampal gyrus near BA (37). Moreover, the left lingual gyrus and the nuclei of the left thalamus were activated.

Conclusion Experiment 6b- Behavioral Data

The analysis of the reaction time data showed no significant difference between the two causal conditions and between the associative-predictive and the associative-diagnostic blocks. These results compared with the analysis of the errors over all blocks can rule out a speed accuracy trade off: Subjects made fewer errors in the blocks where they also provided the fastest reaction times. The error analysis indicates that the *associative conditions* were less demanding than the *causal conditions*.

A closer look at the reaction time data shows that the response time for the associative-predictive trials was faster compared to the causal trials. Suggesting different task demands between the causal and associative access to semantic memory. The causal access is more difficult than the associative access, and therefore, the reaction time is slower and the error rate is higher.

The lack of a significant difference between the predictive and diagnostic of the causal blocks compared to the results in Experiment 6a could be attributed to the small number of subjects. Due to financial limits the number of subjects in a fMRI experiment is not very high. I was able to run 12 subjects. 24 subjects participated in the behavioral pilot Experiment (6a). Thus, the results in the fMRI study could be due to weak power of the experiment. As could be seen from previous computations the effect size was medium for all other experiments. Accordingly to detect a medium effect size would presuppose a higher number of subjects (provided the α and β levels were not altered).

Conclusion Experiment 3b- fMRI Data

In the following discussion I offer preliminary hypotheses about the neurobiology of causal relational processing. The nature of the conclusions is only provisional and the present

---

[11] BA stands for Brodmann Area.

results do not allow me to infer with certainty that either the specific functions of the activated areas, or even whether the particular activations are essential to the access of causal relations. Therefore, further studies will be required to collect more information.

The accessing of causal semantic relationship from semantic memory seems to activate a network including language areas and semantic working memory components. The left hemispheric activations involved the inferior frontal gyrus, close to BA 47, which is part of the left hemisphere's language sites (Petersen, Fox, Posner, Mintun, & Raichle, 1989; Price, 1998), and the lingual gyrus. The bilateral parahippocampal activation supports the argument that semantic declarative memory is accessed (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Wagner, Schacter, Rotte, Koutstaal, Maril, Dale, Rosen, & Buckner, 1998). Furthermore, the activation of thalamic nuclei and basal ganglia is supposed to mediate working memory, executive strategies and rule-based learning (Alexander, Gutcher, & DeLong, 1990; Middleton & Strick, 1991; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999). The connection between the thalamic nuclei and the right frontal lobe as been regarded as forming a logic-specific network (Parsons & Osherson, 2001). It is argued that the left hemispheric basal ganglia support the left hemisphere grammatical rule processing (Damasio & Damasio, 1992), whereas the right hemispheric basal ganglia support rule-based deduction in right frontal areas (Parsons & Osherson, 2001). Therefore, the judgment of causal relations may include parts of that network and semantic working memory. To determine a causal relation between two words would require to retrieve information about the properties of causality. The judgment is guided by knowledge about causation, just as the acquisition of causal knowledge is supposed to be guided by abstract knowledge about causation (Waldmann & Holoyak, 1992, Waldmann, 1996). Consequently, the information provided by the word pair needs to be integrated in a causal context, therefore, lexical information about the words had to be accessed taking place in the left hemispheric language areas. This lexical information needs to be processed in semantic working memory where the constraints of causation and the lexical information is integrated. That process would be accomplished via the thalamic-basal ganglia-right frontal lobe connection. The right hemispheric activation corresponds with studies where similar frontal regions were found to be involved for context interpretation (Bottini, Corcoran, Sterzi, Paulesu, Scenono, Carpa, Franckowiak & Frith, 1994; Shammi & Stuss, 1999) and the integration of causal information in a text comprehension task (Mason & Just, 2001). In contrast to Münte et al. (1998) only a small activation was found in the left frontal area suggesting the retrieval of a causal

relationship between two words does not require syntactic integration. In their study, Münte et al. (1998) used whole sentences, therefore grammar and syntax needed to be regarded to understand the sentence discourse. In my experiments only two words were presented, consequently grammar or syntax was irrelevant, but the meaning of the word had to be accessed in the lexicon, which could have resulted in an activation of the left hemispheric language areas.

Given these first results about the explicit retrieval of causal relations from semantic memory, the suggestion arises that the integration of a causal context is reflected by a different cerebral blood flow pattern than the integration of an associative context. Furthermore, the retrieval of causal relations seems to be connected to a neural circuit involved in deductive rule-based reasoning not relevant for the retrieval of associative relations. Therefore, it can be concluded that relational access to memory can have different qualities depending on the relation in question. For the differentiation of causal and associative access it can be concluded that these two processes are qualitatively different reflected by different cerebral activation patterns.

To summarize, the preliminary data of the fMRI investigation of task specific access support the postulates of the **causal model view**. The access to causation is different from association can also be differentiated by cerebral blood flow patterns.

4.5   Summary of results

Before I start discussing a theory of semantic memory and models to account for my data, I will summarize the results.

- Causal relations are retrieved from semantic memory more easily in their natural predictive cause-effect order than in the reverse diagnostic effect-cause order (Experiments 1 to 3).

- Furthermore different relations (*associative* vs. *causal*) can be selectively accessed in identical items in a task-specific fashion (Experiment 4). Causal directionality only effected retrieval if participants attempted to access causal relations but not when they focused on associative relations. Thus, causal relations cannot be reduced to associative relations in an *explicit* relational semantic memory access task.

- In an *implicit* semantic memory access task (priming), the passive activation of a causally related word following a prime word, was equally strong for causes and effects (Experiment 5). Causal directionality has no influence on the passive spreading of activation between associated concepts.

- The effect of causal asymmetry on semantic retrieval is not a consequence of preferences for specific strategies of access but is an intrinsic feature of the structure of semantic memory (Experiment 3).

- Experiment 6 could support the relational access view tested in Experiment 4. On the basis of brain activation a difference in location and intensity was found for the two different relational accesses.

The results of all experiments provide converging evidence and support for the **causal model view** and they contradict the proposals of the **associationistic view**.

5.   General Discussion

5.1   Causal asymmetries and retrieval processes

As introduced in chapter 3 causal relations have very specific properties: They share temporal and spatial contiguities, the connection between the cause and the effect depends on certain regularities like temporal order, and it is according to Hume perceived as necessary to infer causality. One prominent property of causation is causal asymmetry, which can be attributed to various factors:

- *Temporal order*: The cause precedes the effect in time

- *Probabilistic independence*: Causes of a given effect are probabilistically independent, but become dependent if the effect is hold constant whereas effects of a given cause can be dependent, but become independent, if the cause is hold constant.

- *Manipulation and action*: Causes can be used to manipulate their effects, but not vice versa.

These asymmetric features of causality also influence how human beings deal with causation. I reported results that showed that the acquisition of causal knowledge is influenced by the direction of causation (Waldmann & Holyoak, 1992); and I also discussed different storage aspects of causal knowledge. In my experiments I found that the *retrieval process* of causal relations from semantic memory is guided by *causal asymmetries*:

Causal relations are retrieved from semantic memory asymmetrically depending on the "natural" order of the connected concepts. They are accessed and retrieved faster, if the process is carried out in the predictive order from cause to effect; they are accessed and retrieved slower if the direction is diagnostic.

This retrieval asymmetry could only be found for an *explicit causal* access of causal relations, but not for an *associative* access of causal relations. *Explicit* access to associative relations is guided by the knowledge about what an association is: a bi-lateral and bi-directional connection of two or more concepts. As could be seen in the reported experiments the retrieval speed for an associative relation between causal word pairs is equally fast for both directions for equally associated word pairs (Experiment 4). *Implicit* measures of semantic memory supported this view. In an *implicit* process no relation needs to be generated, judged or retrieved, only a passive spreading of activation starting at one activated concept and traveling along the links to the other connected concepts is investigated. This

passive process depends on the *association* between the concepts and not on any other kind of relation.

The results of my experiments could show that the access to the database in semantic memory is quite flexible and economic. Depending on the task the same memory contents can be accessed via different relations: causation or association. Moreover, I could show that the results of the *explicit* and *implicit* measures investigating different aspects of memory access, contribute to the differentiation between a *causal* and an *associative* retrieval asymmetry.

The majority of these results (except the relational priming results) was predicted by the **causal mode view** and therefore support it. The predictions of the **associationistic view** were different and could be not supported.

5.1.1 Retrieval of causal relations in semantic memory models

The semantic memory models I discussed in section 3.2.1 make different assumptions about the structure of semantic memory, especially about the links between memory contents or the activation of concepts in semantic memory and the retrieval processes of causal relations. I discussed their limitations in representing causal relations and causal asymmetry.

The discussed models can only partly explain the data found in my studies. The *spreading activation theory* (Collins & Loftus, 1975) and the *distributed network model* (Rumelhart & McClelland, 1990) could explain the priming results of Experiment 5 and the results of the *associative* condition in Experiment 4. Both models have an associative character and thus are able to explain the associative tasks. Despite the existence of labeled links in *spreading activation theory*, the association between the concepts is regarded as symmetric. The stimulus material was equally associated for both causal directions, therefore no retrieval asymmetry would have been predicted by the semantic network models in an *associative* access task. However, the experiments also showed that the explicit *causal* access to causal relation is influenced by causal directionality. Potential explanations offered by network models are trivial: In the *spreading activation theory* causal asymmetry is reduced to the labeling of the link, which does not provide a satisfying explanation why the diagnostic direction is slower. In *distributed network models* reaction time depends on the activation of the output, but the model is not able to provide further explanations on how the output activation corresponds to reaction time.

Feature comparison models have difficulties in representing causal relations and causal asymmetry, accordingly they cannot account for the results found and therefore are not discussed any further.

Within the framework of *relational element theory* (Herrman & Chaffin, 1986) the representation of causal relation poses no problem. The representation of causal asymmetry, however, is more difficult, because the relational elements of causation incorporate a default causal directionality. This unidirectional relational element would yield a positive evaluation result, if the direction between the two concepts is predictive. A diagnostic direction would result in a negative evaluation. The results showed that participants were also able to evaluate diagnostic causal relations correctly, but slower. The *relational element theory* could only explain the results found in predictive causal trials.

In the associative trials the *relational element theory* would predict that the constraints of the relation are less strict, because association can be based on any type of relation and moreover, the directionality constraint would not exist for any association. The access of associative relations would be equally fast in both directions. Given these assumptions the relational element theory could account for the results that there was no retrieval asymmetry in the *associative* access.

For the results of the priming experiment the relational element theory would make no prediction, because priming is an *implicit* relational access and assumptions of the theory are restricted to *explicit* access.

5.2   Alternative models of relational access

So far I have discussed my results in the perspectives of the most prominent representatives of semantic memory models. The different explanations of these models for the data found in my studies have not been satisfying. None of the models could provide a clear and promising explanation for the behavioral data found in my experiments. Moreover, they were unable to offer straight forward processing assumptions about causal directionality without being arbitrary and post-hoc data explanation.

Consequently, different models need to be investigated to explain the empirical findings. In this final chapter I will introduce three models offering potential explanations for the results of my experiments. First, I will develop my idea of a hybrid model of task specific access to causal relations based on the *relational element theory* of Herrman and Chaffin (1986). Then, I will introduce and discuss two other tentative models that could also offer

potential explanations. These two models are *Explanatory Coherence* (ECHO) developed by Thagard (1989) and *Learning and Inference with Schema and Analogies* (LISA) developed by Hummel and Holyoak (1997). From these models I will take their basic processing assumption, which in itself cannot account sufficiently for the empirical findings supported by the **causal model view** and therefore I need to extend theses processes to fit the following properties:

- Causal stimuli having an equal strength of association in both directions (predictive vs. diagnostic) yield different retrieval times from semantic memory depending on the relational task. If a *causal* relation has to be determined between a causal word pair, the predictive verification process is faster than the diagnostic verification process. In contrast, if an *associative* relation has to be determined between a causal word pair, there is no difference in the verification process between both directions (Experiments 2, to 4 and 6).

> **The model should be able to represent different types of relationships between the same concepts**.

- The retrieval time for predictive causal relations is faster than for diagnostic causal relations.

> **The model needs to provide a satisfying assumption about the process responsible for that difference and in addition has to account for differences in other possible asymmetric relations.**

- Explicit and implicit access processes to semantic memory yield different results. In a priming paradigm causal stimuli show equal amounts of priming regardless of the causal direction (Experiment 5). Explicit access processes to causal knowledge exhibit directionality differences.

> **The model has to explain access differences depending on the nature of the access task.**

- Human beings are able to retrieve potential causes of a given effect from memory.

> **The model needs to represent the predictive and the diagnostic access to causation.**

5.2.1   A hybrid model of causal semantic memory retrieval

The model I propose is based on an *associative semantic network model* (Collins &
Loftus, 1975) with a relational demand specification introduced by the access task (compare
the *relational element theory* of Herrman and Chaffin, 1986). Consequently, it is a hybrid
between semantic networks and relational evaluation mechanisms, because the network and
the evaluation mechanism separately are unable account for all my experimental results.

Now, I first want to introduce the architecture of the model and then describe in detail
how it would account for the empirical findings in my experiments.

Architecture of the model and relational representation

The model is based on an *associative semantic network* and consists of *labeled nodes*
and *labeled unidirectional links* between the related nodes. The *nodes* correspond to stored
concepts and the *links* correspond to the *relational connections* between the nodes. The *type
of relation* between the concepts determines the *directionality* of the link. Asymmetric
relations such as causal or temporal relations are represented by a single directed link between
concepts, whereas symmetric relations such as covariation have two links, one for each
direction. Together all the links between two concepts constitute the association
(i.e., associative relation) between the concepts (Figure 29 and Figure 30). Association per se
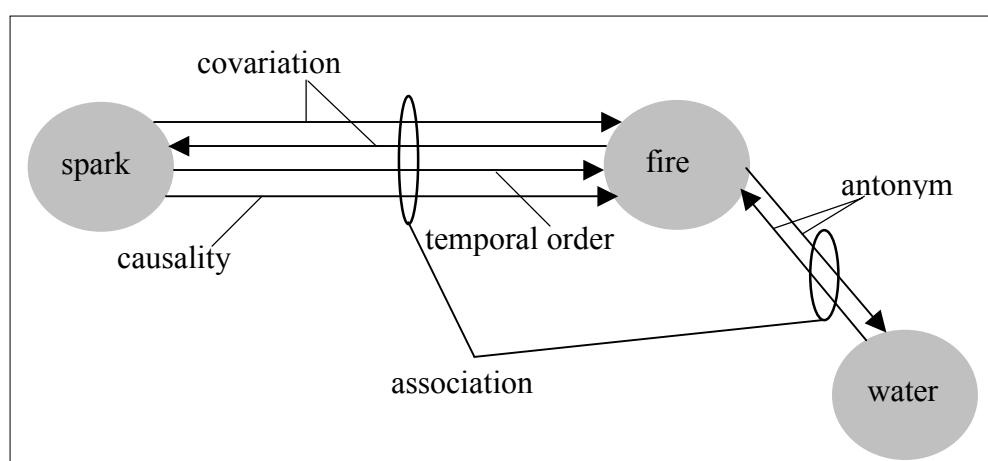is bi-directional, because it incorporates links that go either way.

**Figure 29: Semantic concepts and different types of relations. The arrows indicate the direction of the
relation and the direction of the spreading of activation. All links between two concepts form the
association between the concept.**

An association between the concepts can be based on any meaningful relation. Meaningful relations can for example be, covariation, antonym, causation, temporal order, etc. (Figure 29). They are acquired through explicit learning, individual experience, language, etc. Thus, an association could be compared to a fiber wire cable, each fiber carrying different kinds of information and different amount of information with a different speed. So "inside" the associative link could also be information about a causal relation.
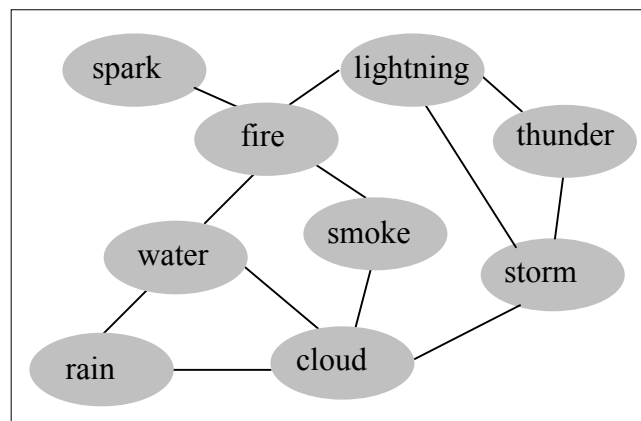
Figure 30: Larger associative semantic network with nodes representing the concept and links representing any meaningful relation between the concepts. For simplicity reasons the different links are represented by one associative link.

Given this structure the model is able to represent different types of relations between the same concepts, which is one of the property demands of the model. Moreover, this architecture would correspond to the **causal model view** postulating that various types of semantic relations including their specific features are represented in semantic memory.

Relational access in the hybrid model

After describing the architecture of the model, I now turn to its functions and an explanation of the relational retrieval process. I need to differentiate between the retrieval process of the *overall symmetric associative* relation and the retrieval process of a *specific asymmetric* type of relation such as causation. First, I introduce the access to the overall associative relation and then proceed to the access of asymmetric relations, where causality serves as an example of an asymmetric relation. The retrieval processes I am going to describe are based on the sequential presentation of concept pairs, that is, one concept is presented after the other and an evaluation process of the relational connection between the concepts. Retrieval of a relation depends on the order of the concept presentation and on the number of steps that need to be executed to evaluate the relation.

There is one important assumption the model makes about the evaluation process:

The relational checking mechanism can only process the relation in one direction and that is the direction specified by the type of relation. For example, the causal direction can be only checked in the predictive direction. Reflecting the prominence of the predictive causal direction influenced by training through daily life experience. (see section 3.3.2). This assumption provides the basis for the whole evaluation procedure.

Explicit retrieval of the overall association

In the hybrid model, all possible connections between two concepts form the association between them. This association is bi-directional. Therefore, the overall associative connection between two events is bi-directional. Evaluating a bi-directional relation does not underlie unidirectional constraints like the evaluation of a unidirectional relation such as causal or temporal relations. Consequently, regardless of the order of the presented concepts of the associative relation the evaluation process can be executed instantly after the first concept is presented. The described evaluation process is based on an *explicit* relational access and not on an *implicit* access like in a priming situation with short SOAs.

Evaluation process of an associative relation

In order to *explicitly* retrieve an *associative* relationship the associative connection between the two concepts presented needs to be evaluated. This evaluation process is accomplished by only the first step of a relational checking mechanism that can consist of up to three steps (the second and the third step become necessary only if an asymmetric relation has to be evaluated).

The order of presentation is represented in the evaluation process by two *numbered slots*. *Slot 1* stands for the *first concept* and *slot 2* stand for the *second concept*. The first step of this process also includes the assignment of the concepts to the slots in the order of their presentation. Immediately after the assignment of the first concept the relational checking mechanism can be started. The second concept is presented and assigned to *slot 2*. If a connection is found, the checking process terminates and a positive answer is given whereas, if no connection is found, the checking process terminates and a negative answer is given. This step is independent of the direction of the connection, because it only has to determine if there is a relational connection at all and not if the relation is directed. Figure 31 shows the retrieval of the associative relation between two causally related concepts and between two unrelated concepts.
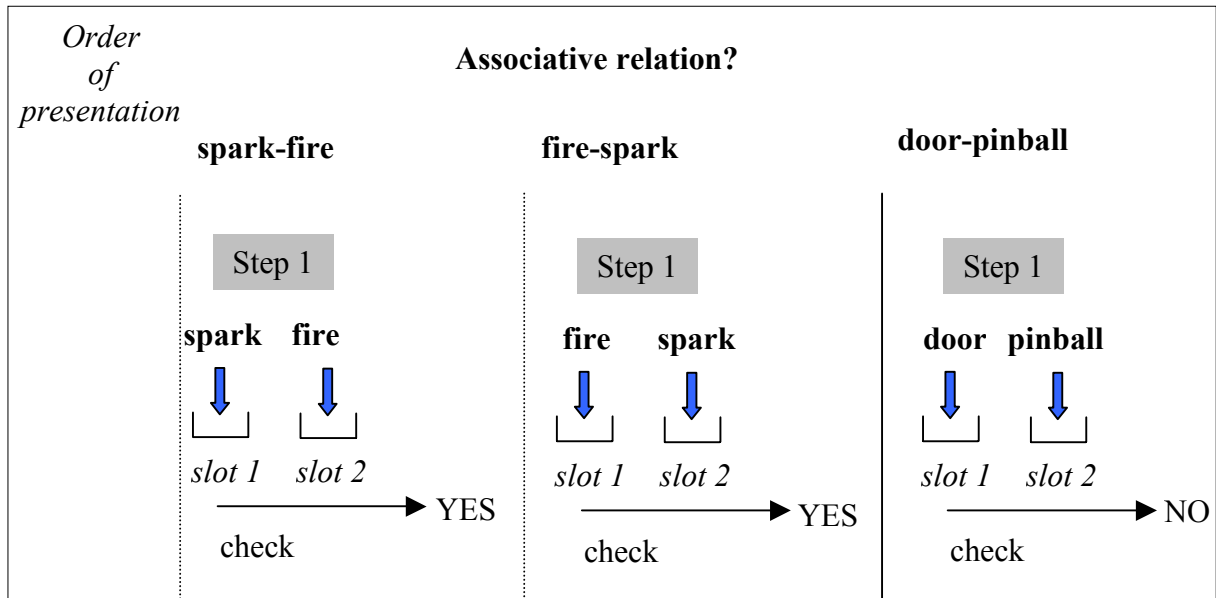
Figure 31: Three examples of step 1 of the relational checking process. For both representation orders and for associated and unassociated relations.

Explicit retrieval of an asymmetric relation

However, if the task demands an evaluation of an asymmetric relation between two concepts for both directions, the next two steps of the evaluation process become necessary. If a specific relation has to be accessed, the constraints of that relation are now superimposed on the overall associative connection. The checking process has to evaluate the relations linking the concepts for the specific constraints. The evaluation of an asymmetric (e.g., causal) relation requires to check the connections between concepts for that specific (causal) relation and moreover, to consider the direction of the relation and the order of concept presentation. Figure 32 shows the superimposing of causal constraints on the associative links between the concepts.
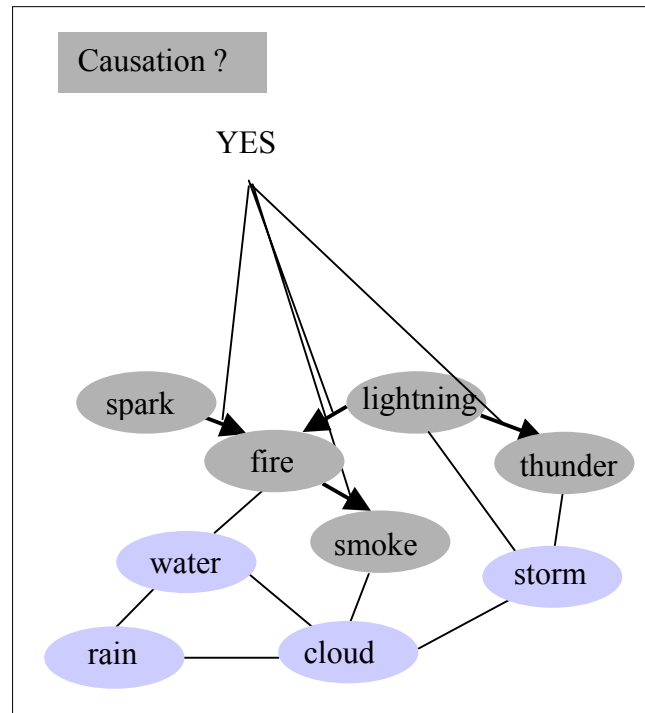
Figure 32: Causal access to semantic network. The constraints of causation is superimposed on the association between the concepts. The arrows represent the constraint of causal directionality

In an asymmetric relation the order of presentation can be with (forward order) or against the direction (backward order) of the relation. The evaluation process is executed in three steps. I already introduced the first step of the relational evaluation mechanism above. That is, the determination of the existence of a connection between two concepts. Now, I need to extend this evaluation process in order to explain the retrieval process for an asymmetric relation such as causation. There are different ways of accessing an asymmetric relation depending on the order of presentation and the retrieval task:

1. The task demands to retrieve an asymmetric relation between two words
   regardless of the order of their presentation. For example, determination of a
   causal relation between two concepts presented in the forward (i.e., predictive) *and*
   in the backward (i.e., diagnostic) order.

2. The order of the presentation of the two concepts is maintained and the asymmetric
   relation has to be *retrieved separately* for the *forward* and *backward order*. For
   example, a causal relation has to be determined for concepts presented either in the
   predictive *or* in the diagnostic direction.

Evaluation process of an asymmetric relation

In the following section I will describe the relational evaluation process for these two tasks separately. The relational evaluation process is based on the already mentioned assumption that the relational checking process can only be executed in the *predictive* direction.

Task 1:

To determine an asymmetric relation (e.g., causal relation) between two concepts regardless of the presentation order (of the two concepts) involves the following steps:

- **Step 1**: The first word is assigned to *slot 1*, the relational checking process starts, the second word is assigned to *slot 2*, if the evaluation yields a positive outcome (i.e., if a causal relation could be determined) checking mechanism is terminated and the response "Yes" is given. If the evaluation yields a negative outcome (i.e., there is no causal relation between the two words), then step 2 and 3 need to be executed.
- **Step 2**: The words are re-assigned in the reversed order, that is, the *first word* is assigned to *slot 2* and the *second word* to *slot 1*.
- **Step 3**: The relation is evaluated again. If the evaluation yields a positive outcome, the checking mechanism is terminated and the response "Yes" is given. However, if the evaluation yields a negative outcome the checking mechanism is terminated and the response is "No". Figure 33 depicts the process based on a causal example.

Determination of an asymmetric (causal) relation regardless of the presentation order

Forward order of presentation:

Step 1   **spark-fire**

spark   **fire**

slot 1   *slot 2*

check   YES

Backward order of presentation:

Step 1   **fire-spark**

fire   **spark**

slot 1   *slot 2*

check   NO

Step 2

spark   **fire**

slot 1   *slot 2*

Step 3

check   YES
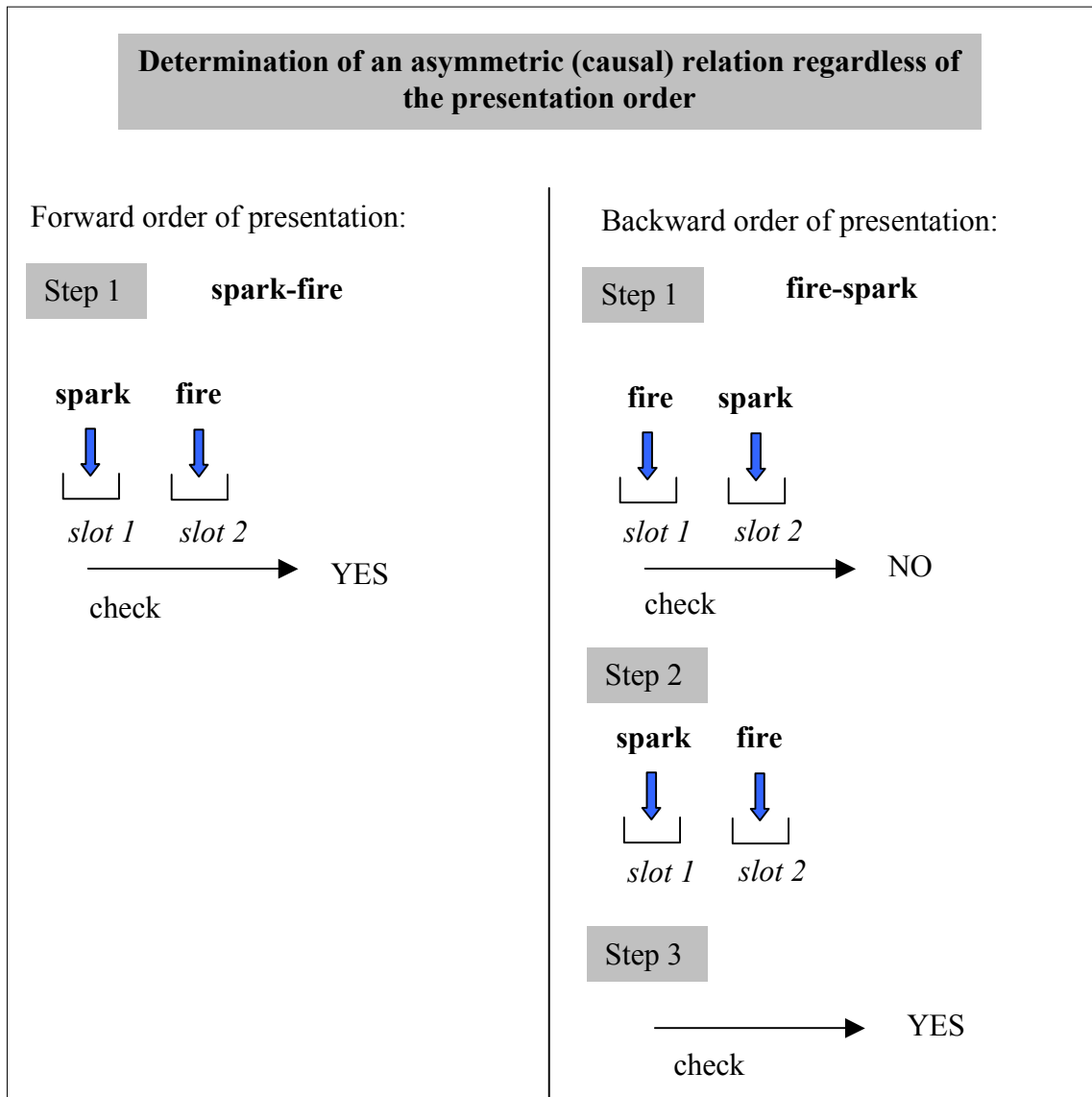
Figure 33: Steps of evaluation of an asymmetric relation based on a causal example.

As can be seen in Figure 33, the checking process can only be executed in the forward direction making necessary the re-assignment in the reversed order (see Step 2).

Task 2:

To determine an asymmetric (causal) relation for a specific order of presentation also involves one or three steps, depending on the order of presentation:

Forward order

If the presentation order (i.e., forward) reflects the relational direction (i.e., predictive) only step one is necessary (Figure 34 left hand side):

- **Step 1**: The first word is assigned to *slot 1*, the relational checking process starts, the second word is assigned to *slot 2*, if the evaluation yields a positive outcome (i.e., if a causal relation could be determined) checking mechanism is terminated and the response "Yes" is given. If evaluation yields a negative outcome the response "No" is provided.

Backward order

However, if the relation has to be determined for the backward order of presentation (i.e., diagnostic order) the evaluation process is similar to the one in Task 1, but the checking mechanism is put on hold until step three (Figure 34 right hand side):

- **Step 1**: The first word is assigned to *slot 1*, the relational checking process is put on hold, and the second word is assigned to *slot 2*.

- **Step 2**: The words are re-assigned in the reversed order, that is, the *first word* is assigned to *slot 2* and the *second word* to *slot 1*.

- **Step 3**: The relation is evaluated again. If the evaluation yields a positive outcome, the checking mechanism is terminated and the response "Yes" is provided. However, if the evaluation yields a negative outcome the checking mechanism is terminated and the response is "No".

The checking process does not start right away with the presentation of the first word but is delayed until the presentation of the second word and the re-assignment of the words. The process can only be executed in the forward direction therefore, it cannot start until the concepts are put into the fitting (i.e., forward) order. The process is depicted based on a causal example in Figure 34.
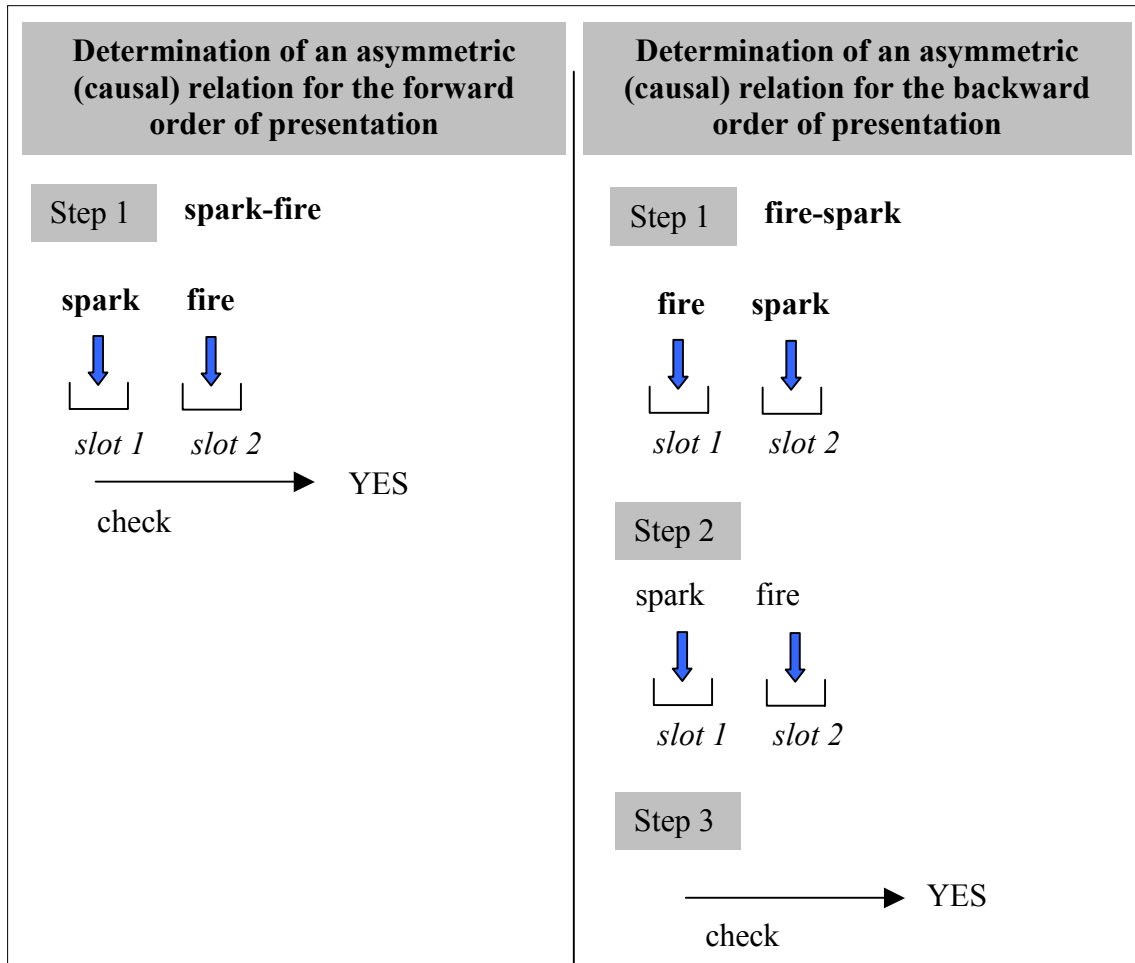
Figure 34: Evaluation process of an asymmetric relation for a specific order of presentation based on a causal example.

This three steps evaluation process reveals that it is not possible to instantly access the relation in the backward order in contrast to the forward order. The relation checking process can only be executed in the predictive direction. Any other direction would result in a delay of that process and in delayed response until enough information is provided to start the evaluation.

Summary of the relational evaluation process in the hybrid model and the empirical results

Depending on the task the relational retrieval process depends on the order of presentation and on the number of necessary steps of the checking mechanism. For an associative access the order of presentation can be neglected and only the first step of the checking mechanism needs to be executed. This process would explain why the retrieval

speed for the associative access to causal relations in Experiments 4 and 6a did not different significantly for the two causal directions.

In conditions of a *causal* access to causal relations the presentation order has an influence on the number of steps necessary during the relational checking mechanism, because the relational checking mechanism can only be executed in the predictive direction. In tasks where a *causal* relation has to be retrieved regardless of the order of presentation, the predictive presentation of the concepts can be evaluated in the first step of the checking mechanism. The evaluation of the same relation in the diagnostic direction involves all three steps (Table 16). *The checking process starts with the presentation of the first word.* These differences in the evaluation process would be able to account for the results found in Experiments 1, 2, 4 and 6. In these experiments the *causal* access to a causal relation revealed a retrieval asymmetry. The *causal* access to a predictive causal relation was faster than the *causal* access to a diagnostic causal relation.

In tasks where the causal access to a causal relation has to be retrieved for a specific causal direction, the predictive presentation of the concepts once again, can be evaluated during the first step of the checking mechanism. The evaluation of the same relation for the diagnostic order involves all three steps and moreover (Table 16), *the checking process is delayed until the third step*. The results found in Experiment 3 could be attributed to this response delay. The retrieval time for a predictive causal relation was faster compared to a diagnostic causal relation. The directional restriction of the relational checking process in only the predictive direction was responsible for the response delay like in Experiment 5.

Table 16:

Summary of the number of steps involving the retrieval of different types of relations.

Task specific access	Experiments	Number of steps and onset of checking mechanism	
associative	4	1 step, onset at the first step.	3 steps, onset at the first step.
causal	1, 2, 4 and 6	1 step, onset at the first step.	3 steps, onset at the first step.
	3	1 step, onset at the first step.	3 steps, onset at the *third* step.
implicit	5	No active relational checking mechanism, but passive spreading of activation depending on association.	

The results of Experiment 5 showed that an implicit access to causal relations exhibits associative priming, but no relational priming. Associative priming is based on the passive spreading of activation between the concepts and not on an active relational evaluation process. As can be seen in Figure 35, priming mechanisms can also be represented in the model. For example, if the node "fire" is activated, the activation spreads along the connections to the other links such as "water", "smoke", "spark", etc.
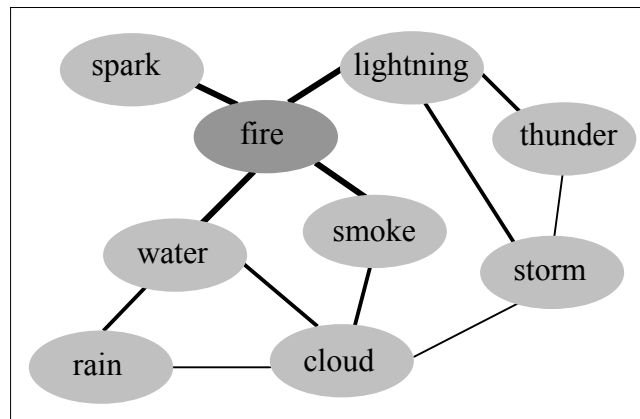
Figure 35: Spreading of activation between related concepts. The darker node is activated and the activation spreads to connected links. The thickness of the lines represents the degree of activation: thicker lines reflect more activation and thinner lines represent less activation.

To summarize, the hybrid model of relational access can fulfill all demands stated at the beginning of this chapter:

1. it can represent different types of relation between the same concepts,
2. it is able to account for retrieval asymmetry in asymmetric relations,
3. it can explain access differences depending on the nature of the task, and
4. it can represent the predictive and diagnostic access to causation.

All these demands are supported by the **causal model view** that postulates that there are different relations represented in long-term memory and that these relations cannot be reduced to association as proposed by the **associationistic view**. The access to these relations is task specific and yields an asymmetry in the retrieval time for causal or other asymmetric relations.

The hybrid model represents this retrieval asymmetry via the three steps of relational evaluation. One assumption of the hybrid model is that the checking mechanism can only be executed in the direction determined by the type of relation, that is, only in the predictive direction of a causal relation. The evaluation of a diagnostic causal relation involves reversing

the order of the concepts so that the checking process can be initiated. Given this assumption the empirical findings can be explained.

However, it could also be the case that the diagnostic direction can be accessed instantly and there are other factors that influence the retrieval speed. These factors could be independent from semantic relational evaluation but rather could be dependent on features of causality itself. For example, it is quite likely that an effect has more than one cause. Consequently, the evaluation of a diagnostic causal relation might also involve taking other alternative causes into account increasing the difficulty of determination of a diagnostic causal relation.

The next model I am going to describe deals with that issue. It is a constraint satisfaction model about explanatory coherence (ECHO) developed by Thagard (1989). It is important to note that ECHO is not a model about the representation and retrieval of semantic relations. Nevertheless, it might be helpful to explain why there is a retrieval asymmetry in the causal access to causal relations as found in the empirical results and as predicted by the **causal model view**.

5.2.2 ECHO

Constraint satisfaction models (Thagard, 1989, 2000) aim at capturing qualitative aspects of everyday reasoning. Their basic assumption is that people hold a set of interconnected beliefs. The beliefs pose constraints on each other, they either support each other, contradict each other, or are unrelated. *Coherence* between the beliefs can be achieved by processes, which attempt to honor these constraints. Thagard (1989) proposed seven principles of coherence:

1. **Symmetry**: Explanatory coherence is a symmetric relation.
2. **Explanation**: a) A hypothesis coheres with what it explains, which can either be evidence or another hypothesis; b) hypotheses that together explain some other proposition cohere with each other; and c) the more hypotheses it takes to explain something, the lower the degree of coherence.
3. **Analogy**: Similar hypotheses that explain similar pieces of evidence cohere.
4. **Data priority**: Propositions that describe the results of observations have a degree of acceptability on their own.
5. **Contradiction**: Contradictory propositions are incoherent with each other.
6. **Competition**: If P and Q explain a proposition, and if P and Q are not explanatorily connected, then P and Q incohere with each other.
7. **Acceptance**: The acceptability of a proposition in a system of propositions depends on its coherence with them.

Architecture and relational representation

Within a constraint satisfaction model beliefs are represented as nodes, which represent propositions (e.g., "A causes B"). The nodes are connected by symmetric relations. The numerical activation of the nodes indicates the strength of the belief in the proposition. A belief that is highly activated is held strongly, a belief that is negatively activated is rejected. The activation of a node depends on the activation of all other nodes with which it is connected. More precisely, the net input to a single node *j* from all other nodes *i* is defined as the weighted sum of the activation *a* of all related nodes (following Thagard, 1989, p.466, eq.5):

$$Net_j = \Sigma_i w_{ij} a_i(t) \quad (1)$$

The weights w represent the strength of the connection of the beliefs. Figure 36 shows the representation of a common-cause/common-effect combination in a constraint satisfaction model. The causal hypothesis node H1 on the top represents a structural causal hypothesis stating that the five events c1, c2, e1, e2, and x form a common-cause/common-effect situation in which c1 and c2 are the causes of a common effect x and e1 and e2 are the effects of event x in the common cause part of the structure.
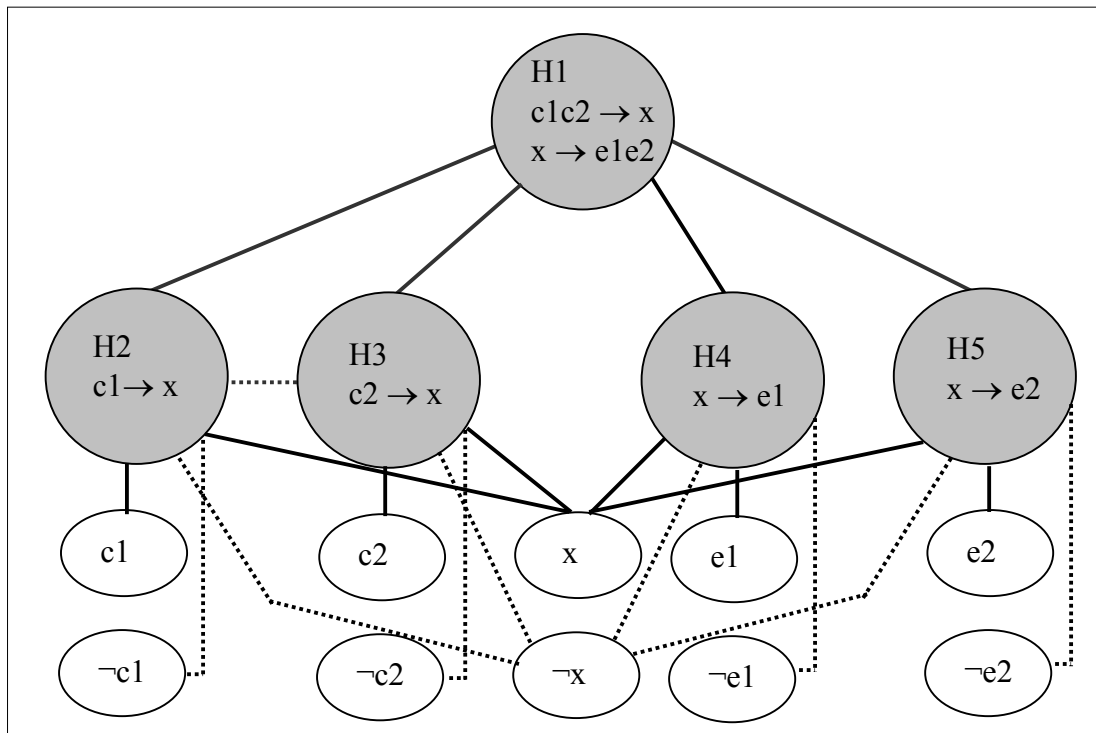
Figure 36: Constraint satisfaction model for the 5 propositions depicting a combination of a common-cause and a common-effect structure.

The nodes in the bottom level refer to patterns of events that can be observed. For simplicity reasons I only depicted the presence and the absence of the events. There are two different kinds of connections between the nodes. Solid lines indicate the excitatory links, dashed lines inhibitory links. The connections are defined the following way: For example, if the events x, e1 are present the observed pattern is in accordance with the hypothesis H4. This pattern might be observed, if x causes e1. Therefore the evidence nodes e1 and x are positively connected to H4, which in turn is then positively connected to H1. In general a hypothesis is positively connected to an evidence node, if the events mentioned in the hypothesis are present. If this is not the case, a negative default value is assigned. The dashed line between H2 and H3 is important because it represents an inhibitory link. Given a

common-effect situation, the two causes c1 and c2 compete with each other in explaining the effect according to the 6th principle of coherence (Thagard, 1989, 2000).

For example H1 could be a proposition such as "Fire (x) is caused by lightning (c1) and spark (c2) and fire (x) causes smoke (e1) and heat (e2)". Consequently, H2 would correspond to "lightning causes fire", H3 corresponds to "Spark causes fire", H4 refers to "Fire causes heat" and finally, H5 represents the proposition "Fire causes smoke". The two causes "lightning" and "spark" compete with each other in explaining the effect "fire"

Given this structure the ECHO model could represent causal relations. However, it is unable to represent an associative relation, because it is not designed for relational representation, only for modeling explanations.

Relational access in the ECHO model

The competition of two or more causes explaining the same effect would be ECHO's contribution to explain the retrieval asymmetry in the *causal* access to causal relations. If a concept representing the effect of a causal relation is presented, all potential causes become activated via the excitatory link from the structural hypothesis, but simultaneously inhibit each other, given their competition. This competition and the resulting mutual inhibition could influence the retrieval of a diagnostic causal relation. The evaluation of a diagnostic causal relation becomes more difficult and in turn could be processed slower.

On the other hand the evaluation of a predictive causal relation does not require taking alternative effects into account and moreover there is no competition between the effects of a common cause. The effects get simultaneous activation from the structural hypothesis and there is no mutual inhibition between them. Hence, the evaluation of a predictive causal relation is easier and could be faster.

The following simulation shows how the model could support the retrieval asymmetry for causal relations.

Simulation

Simulating the causal retrieval process of causal relations, I will have to assume that the predictive process reflects a common-effect situation, whereas the diagnostic process is simulated by a common-cause situation.

At the beginning of the simulations, the activation of the nodes representing hypothesis are set to a low default value. However, nodes representing empirical evidence are connected to a special activation node (which was not depicted in Figure 36) whose activation

remains constant at 1.0. The common-cause/common-effect combination was implemented using Microsoft Excel.[12] Default values were adopted from the literature if not indicated otherwise (Thagard, 1989). Initial activations were set to 0.01, inhibitory links between nodes to −0.05, and excitatory links to +0.05. The inhibitory link between H1 and H2 within the common-effect part was preset to a value of −0.20. To update the activation in each cycle of the simulation, first the net input net j to each node is computed using Equation 1. Second the activation of all nodes is updated using the following equation (Thagard, 1989, p.446, eq.4):

$$a_j(t+1) = a_j(t)(1-\theta)+net_j(max- a_j(t)) \text{ if } net_j >0$$

$$= a_j(t)(1-\theta)+ net_j(a_j(t)-min) \text{ otherwise. (2)}$$

In Equation 2, θ is a decay parameter that decrements the activity of each node in every cycle, min represents the minimum activation (-1), and max the maximum activation (+1).

The simulation for both retrieval conditions contained the same causal structure. The presentation of the first concept is simulated as the activation of the event x, which is "fire" in the causal example. The evidence node x (fire) is activated and the activation for all other connected nodes is updated until a stable equilibrium is reached, that is, the activation of all nodes does no longer change substantially, which was the case after the first 50 cycles (see Figure 37 and Figure 38). Then, the second concept is presented and therefore its evidence node will be activated. In the predictive causal example the evidence node (e1) for the event "smoke" is activated and the activation of all other nodes is updated until a stable state is reached. As can be seen in Figure 37 the two hypotheses H4 and H5 representing the two effects of the common cause are highly activated, whereas the hypothesis H2 and H3 have a low level of activation.

---

[12] The simulation was programmed by Dr. York Hagmayer whom I want to thank for his support.

Figure 37: Activation of the ECHO model during a predictive trial. Darker gray lines are the effects and lighter gray lines are the causes.

On the other hand, the retrieval of a diagnostic causal relation would yield a different activational pattern in the simulation (Figure 38). The first 50 cycles are identical to the predictive simulation, but the activation of the evidence node (c2) of the event "spark" results in an activation of H2, but in a rejection of H3 represented by the drop of activation below zero.
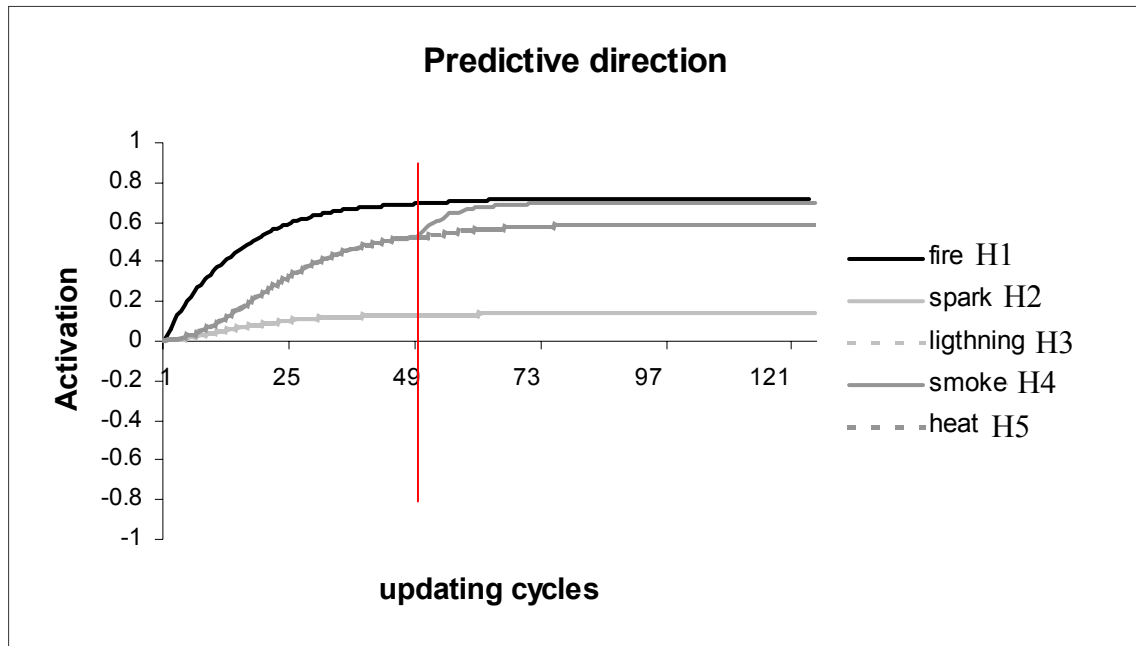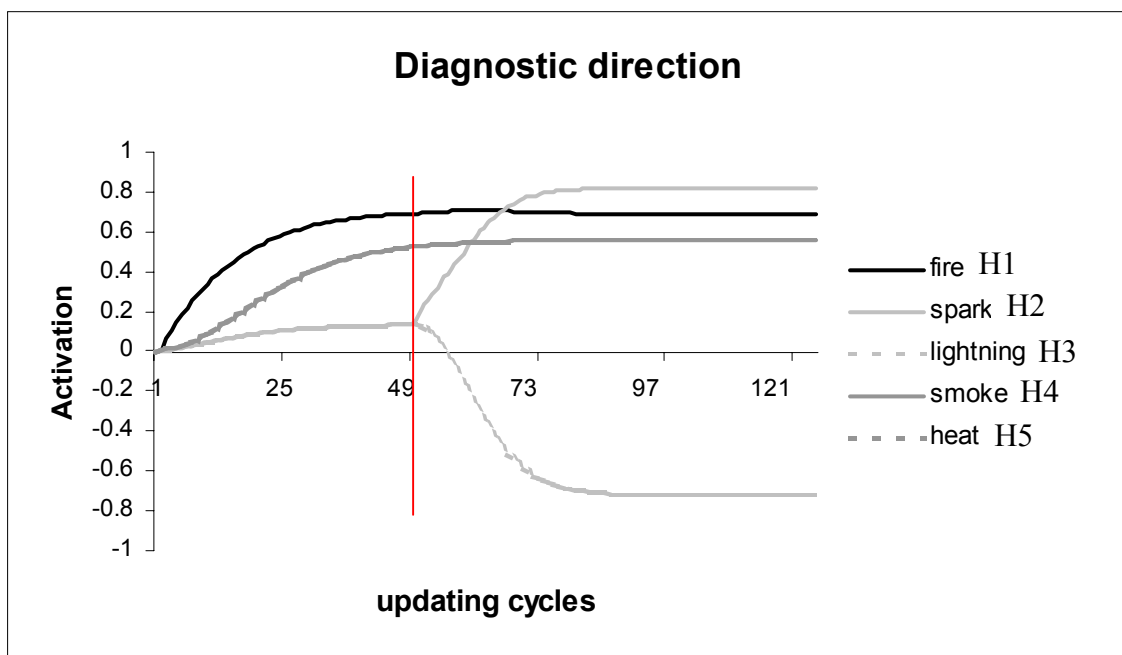

Figure 38: Activation of the ECHO model during a diagnostic trial. Darker gray lines are the effects and the lighter gray lines are the causes.

Summary of the relational evaluation process in ECHO and the empirical results

The simulation of the ECHO model showed that the competition of causes explaining the same effect resulted in a mutual inhibition. This inhibition could be responsible for the increased difficulty of the retrieval of a diagnostic causal relation. The increased demand might yield a slower retrieval speed for diagnostic relations observed in Experiments 1 to 4 and 6.

In contrast to the hybrid model, ECHO offers an explanation based on specific features of reasoning with causal hypotheses and not based on the retrieval mechanism of relational access. This different approach however, is incapable of representing or explaining the associative access to causal relations (Experiment 4), and the implicit access to semantic memory (Experiment 5). It is also not equipped to test all the predictions made by the **causal model** and the **associationistic views** (Table 17). Its account is limited to the prediction made by the **causal model view** about causal retrieval asymmetry.

Table 17:

Summary of ECHO's account of the empirical findings.

Task specific access	Experiments	Mechanism	
associative	4	The model cannot represent association and therefore cannot account for associative access to causal relations.	
causal		*predictive*	*diagnostic*
	1, to 4 and 6	No competition among effects	Competition among causes
implicit	5	No active relational checking mechanism, but passive spreading of activation depending on association	

To summarize, ECHO can only fulfill the demand of explaining the retrieval asymmetry for causal relations in a causal access process predicted by the **causal model view**. Moreover, ECHO would predict this difference by contrasting common-cause and common-effect structures of causal relations, but in an extreme case with only one cause and one effect the model would predict an equally fast the retrieval process fast for either direction, given the lack of competing causes inhibiting each other. This assumption would be contradictory to the

causal model view, because causal directionality is represented in long-term memory. It is an intrinsic feature of causality independent from the structure of causal dependencies.

The last model I am going to introduce returns to the representation of relation in the human mind. It was designed to simulate human analogical and relational reasoning. Like in ECHO, I will discuss only these parts and mechanisms of the model relevant for testing the distinction between the **causal model** and **associationistic view** and to account for my empirical findings.

5.2.3 LISA

The acronym LISA stands for "Learning and Inference with Schemas and Analogies". The model was developed by Hummel and Holyoak (1997) to provide a new idea about learning and reasoning. With this model they want to present a theory of how relational inferences and generalization could be accomplished within a cognitive architecture that is psychologically and neurologically realistic. They assume that the process of finding systematic correspondences between structured representations, which plays a central role in analogical reasoning is fundamental to all relational reasoning (Hofstadter, 2001). The power of relational reasoning resides in its capacity to generate inferences and generalizations that are constrained by the *roles* that elements play, rather than solely the properties of the elements themselves. Hummel and Holyoak propose a cognitive architecture called *symbolic connectionism* in order to reflect the human representational system and relational reasoning. This system is based on *distributed* representations of concept meanings, using temporal synchrony to bind *fillers* and *roles* into a relational structure. A *filler* can be an event, an object, a person, etc., whereas a *role* is a relation a *filler* can have to another *filler*. For example, "spark" and "fire" could be *fillers* and the *role* that binds them together is "cause of".

Given the complexity of the LISA model and its processing assumptions about inference of analogies, I will only refer to processes of the model that would be important for my purposes.

Architecture of the model an relational representation

Hummel and Holyoak (2001) argued that human reasoning is based on distributed *symbolic* representations that are generated rapidly. *Symbols* in this system are distributed representations that explicitly specify the semantic content of their referents. This cognitive architecture is a symbol system where the relations and their arguments are represented independently but can be bound together to compass *propositions*. A *proposition* is made up by a 4-level hierarchical structure depicted for a causal example in Figure 39.

Causation and Association in LISA

One of the demands on the model is the representation of association and causation. Figure 39 depicts how the model would possibly represent a causal relation between two concepts (e.g., ***spark*** and ***fire***). The top level describes the proposition (P1). For causality the

proposition refers to a causal relationship ***causes (spark,fire)***. The second level from above represents the sub-propositions, the connection of the object (*filler*) and its predicate (*role*): **spark** is the *cause* and **fire** is the *effect*. In the level below the objects (**spark**, **fire**) and their predicates (*cause*, *effect*) are represented separately. And finally, the last level contains semantic units that represent semantic properties of the objects and predicates. Properties of a cause could be "first" and "before", whereas the properties of an effect could be "second" and "after". All links between the levels are bi-directional and excitatory.
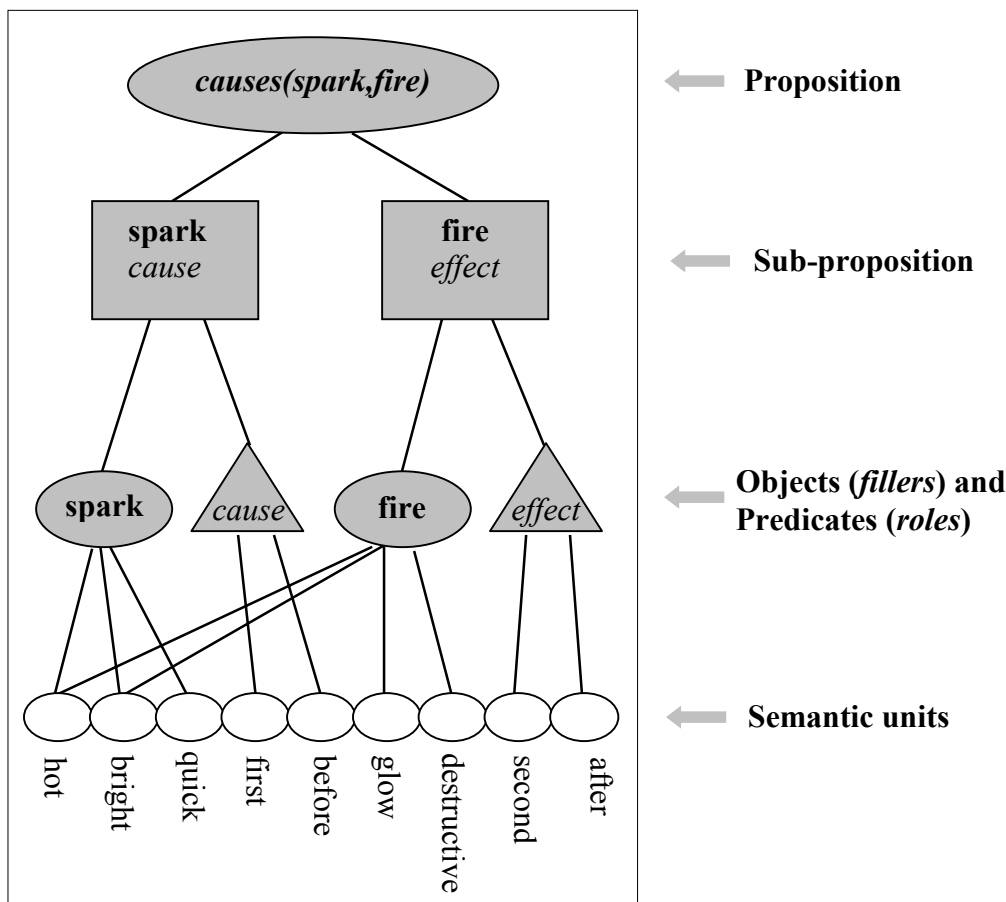
Figure 39: Representation of a causal relation in LISA.

The representation of an associative relation is quite similar, but the association-roles are connected to the same semantic units stating, for example, that associations have no (temporal) order or other directional constraints, and that they specify the occurrence of fillers in a given context (Figure 40).
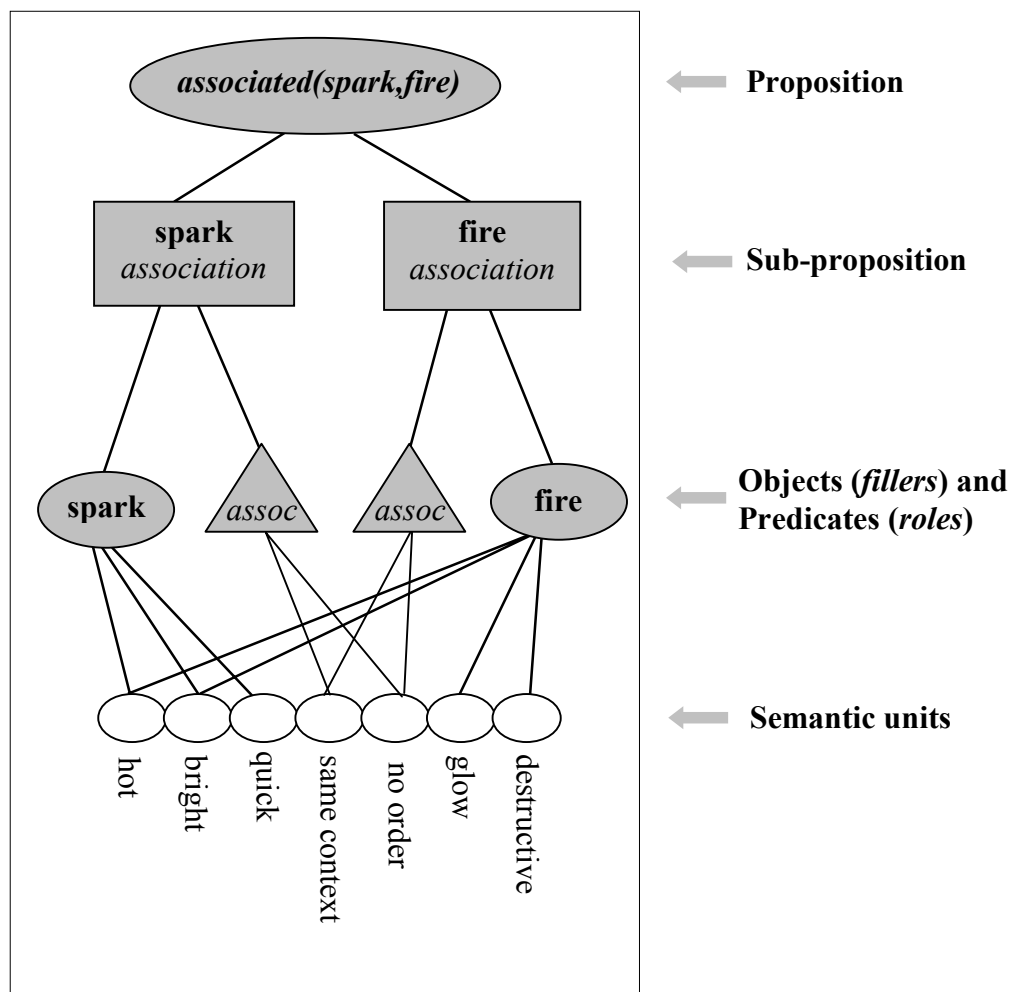
Figure 40: Representation of an association between two concepts

Relational access in the LISA model

Given this architecture the LISA model would make different predictions about the task specific access to relations. For the associative access to causally related concepts the model would predict no influence of the order of presentation, because as can be seen in Figure 40, even though there are two association *roles* they are connected to semantic units that explicitly specify that association is not influenced by directionality. Therefore an associative proposition such as ***associated(spark,fire)*** is equally fast evaluated regardless of the order of presentation of the concepts.

On the other hand the order of presentation is important for the causal access to a causal relation. In the predictive case the order of presentation mimics the causal directionality. LISA would predict that a causal proposition ***causes(spark,fire)*** could be evaluated without any problems, but the same proposition would be more difficult to evaluate

if the order of presentation is diagnostic. If the order of presentation is reversed for a causal relation that is, if the effect is presented before the cause, the model needs to be altered to correspond to that incongruent condition. An incongruent condition can be implemented by *cross-mapping* the predicates of the cause and the effect with semantic units. That is, the units "after" and "second" will be connected to the *cause* and the semantic units "first" and "before" will be connected to the *effect*, everything else being equal (Figure 41).
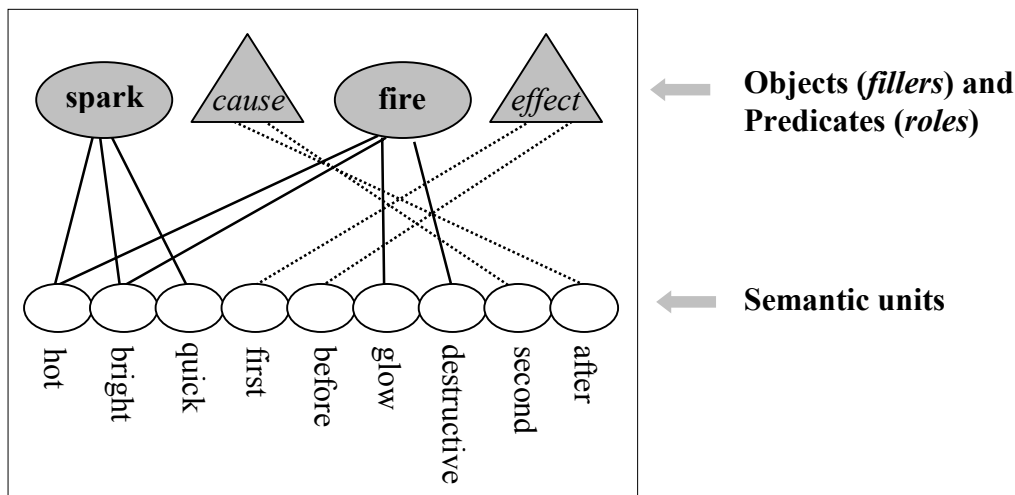
Figure 41: Cross-mapping of cause and effect roles to opposite semantic units. The dotted lines show the cross-mapping.

These processes not only allow the representation of different relations, but also the representation of causal asymmetry. It is notable that the asymmetry is based on the causal direction and not on associative strength as the **associationistic view** would claim. Therefore it differentiates between the **causal model** and the **associationistic view**. Moreover, LISA makes the same prediction as the **causal model view**.

A simulation of the causal access to causal relations was carried out to investigate the retrieval asymmetry.

Simulation

In collaboration with Prof. Dr. John Hummel at the University of California in Los Angeles the causal access to a causal relation was simulated in the LISA model. Table 18 shows the simulation parameter for the causal access. The parameters of the diagnostic simulation reflect the *cross-mapping* of the predicates to the opposite semantic units. The firing sequence reports the order of the activation of the predicates. It was identical of both conditions. The PR1 (cause) was activated before PR2 (effect).

Table 18:

Simulation parameters for the predictive and diagnostic causal retrieval

Predictive causal retrieval	Diagnostic causal retrieval
Proposition: P1(causes spark fire)	Proposition: P1(causes spark fire)
Predicates (*roles*): PR1 **Cause** first before relation causal cause PR2 **Effect** second after relation causal effect	Predicates (*roles*): **Cross-mapping!** PR1 **Cause** second after relation causal effect PR2 **Effect** first before relation causal cause
Objects (*filler*): O1 **Fire** hot bright glow destructive O2 **Spark** hot quick bright	Objects (*filler*): O1 **Fire** hot bright glow destructive O2 **Spark** hot quick bright
Firing sequence: PR1 / PR2	Firing sequence: PR1 / PR2

Processing and performance of the model are expressed in cycles. During each cycle the connection between all units (proposition, sub-propositions, fillers and roles) is updated.[13] Figure 42 shows the activation of the model if the concepts were presented in the predictive direction (left had side of the figure). The simulation involved a cyclic activation of the concepts cause (PR1) and effect (PR2), which in turn activated the causal proposition (P1). The y-axes represents the activational level.

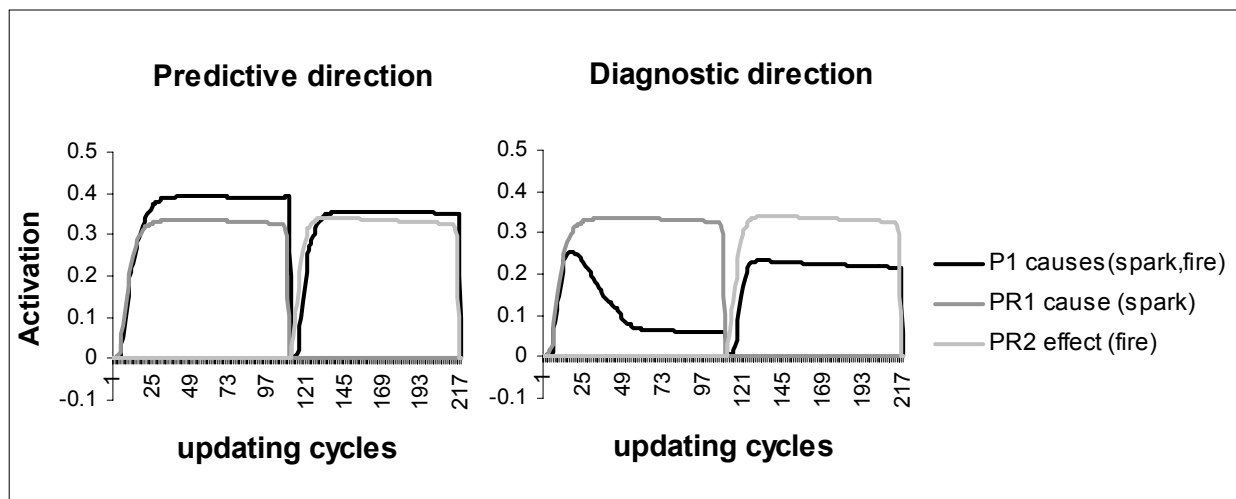**Figure 42: Simulation of a predictive and diagnostic presentation of a causal relation in LISA.**

---

[13] The exact documentation of the algorithms of the connection weights can be found in the Appendix of Hummel and Holyoak (2001).

As can be seen in Figure 42, the causal proposition P1 was immediately activated when the first concept PR1 was activated (i.e., presented). The first concept presented is assigned to the cause and the second concept is assigned to the effect.

In the simulation of the diagnostic presentation the activational pattern was different: The proposition *causes (spark, fire)* (P1) was suppressed during the presentation of the first concept PR1 and got activated slowly with the presentation of the second concept PR2 (right hand side of Figure 42). It is important to note that in LISA the order of presentation is simulated via the semantic units and not by the actual activation of the predicates (i.e., cause and effect). Consequently, as can be seen, the *cause* got activated first, but given its cross-mapping to semantic units of the *effect*, it is regarded as incongruent, and thus, the *proposition* (P1) is suppressed.

Summary of relational processing in LISA and empirical results

As described above the architecture of the LISA model is equipped to represent causal relations and causal asymmetries. Moreover, the task specific access to causal relations can be explained by the model. An *associative* access to a causal relation is independent of the causal direction, that is, an associative proposition would be equally activated regardless of the firing sequence of the predicates. Therefore the LISA model could account for the data found in the *associative condition* of Experiment 4 in which the retrieval speed was equally fast for both causal directions.

The results of the *causal* access could also be explained by the model. The retrieval asymmetry for the two causal directions would be based on an incongruent cross-mapping condition, in which the order of the presentation is incongruent to the semantic specification of the *roles*. As could be shown in the simulation, the causal proposition was suppressed, if the *roles* were cross-mapped to the opposite semantic units. This suppression would correspond to the diagnostic condition in the experiments, where the response time was slower compared to the predictive direction (Experiments 1 to 4 *causal condition*, and 6).

The LISA model would make no prediction and therefore cannot account for the results found in the priming experiment (Experiment 5), because it is a model about explicit relational reasoning and not about implicit relational access.

Table 19:

Summary of the empirical results explained by LISA

Task specific access	Experiments	Mapping	
associative	4	Two association roles connected to the same semantic units.	
causal		One cause and one effect role connected to different semantic units.	
		Predictive direction	*Diagnostic direction*
	1, to 4 and 6	congruent mapping	incongruent cross-mapping
implicit	5	No predictions and therefore no explanation.	

The LISA model is quite equipped to account for the empirical results of my experiments and for the assumptions made by the **causal model view**. In addition it is helpful to differentiate the **causal model view** from the **associationistic view**. However, like the ECHO model it is unable to offer an explanation for the priming results of Experiment 5.

5.3   Perspectives

In the last chapter, I discussed three different models in order to account for my empirical findings and to differentiate between the **causal model** and **associationistic view**. The criteria set by the empirical findings were partly met by the models.

ECHO was the weakest model in explaining the representation and task specific access to semantic relations. It offered an explanation for the retrieval asymmetry found in the causal access of causal relations, but not for the representation of other semantic relations, and the symmetric retrieval results for the associative access. However, it needs to be mentioned in favor of the model that it was not designed to represent semantic relations, and therefore, it is plausible that it is not capable of it.

The LISA model could account for the empirical findings in the explicit task specific access. It would predict a retrieval asymmetry for the causal access to causal relations and no retrieval asymmetry for the associative access to causal relations, like the **causal model**. But also like ECHO it is not a model about semantic retrieval processes. It was designed to account for relational reasoning processes. These processes underlie the explicit access to semantic memory and do no involve the implicit access. Consequently, LISA cannot make any predictions for the implicit access and about the empirical findings in priming paradigms.

The model I introduced first, was a hybrid model based on semantic network models and a relational checking process. The hybrid model was capable of accounting for all the empirical findings and it made the same prediction as the **causal model view**. Its disadvantage so far is that it has not been simulated to validate its performance. The implementation of the hybrid approach into a computer model would be next step to further differentiate between the **causal model** and **associationistic view**.

Another issue, I would like to address is the distinction between temporally and causally related events. As I tried to make clear causation should not be reduced to temporal relations and temporal order. Of course, temporal directionality is part of a causal relation, which is based on regularity, but in itself the temporal sequence of causal events cannot account for the influence of a cause on its effect. In their studies Van der Meer and colleagues (1992, 1999, 2002) try to account for causal directionality by considering it equivalent to temporal directionality. However, the stimuli, which they declare as temporal relations describe regularities and thus can also be considered as causal relationships (Hume, 1739). It seems impossible to differentiate temporal order from causal order in the

retrieval process of regularity based relations from semantic memory. The only way to separate these two relations clearly is to use episodic knowledge in a memory retrieval process, because only then the temporal sequence does not reflect regularity and would not be perceived as causality. Therefore, it still remains to be proven that the results Van der Meer and her colleagues found were due to distinct temporal relationships and not causation.

6. References

Alexander, G. E., Crutcher, M., & DeLong, M. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, 'prefrontal', and 'limbic' functions. *Progress in Brain Research, 85*, 119-146.

Aristoteles. (1929). *The Physics* (Vol. 1). P.H. Wickstead & F. M. Cornford (Eds.), Cambridge, MA: Harvard University Press.

Atchley, R. A., Burgess, K., & Keeney, M. (1999). The effect of time course and context on the facilitation of semantic features in the cerebral hemispheres. *Neuropsychology, 13*, 389-403.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10* (1), 82-93.

Beeman, M. J. (1991). Coherence inferencing and structure building in the cerebral hemispheres. *Dissertation Abstracts International, 51*(11-B), 5610.

Beeman, M. J., Friedman, R. D., Grafman, J., Perez, E., Diamond, S., & Lindsay, M. B. (1994). Summation priming and coarse semantic coding in the right hemisphere. *Journal of Cognitive Neuroscience, 6*, 26-45.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Scenono, P., Carpa, P, Franckowiak, r. S. J., & Frith, C. D. (1994). The role of the right Hemisphere in the interpretation of figurative aspects of language. A positron emission tomography study. *Brain, 117,* 1241-1253.

Buehner, M. J., & Cheng, P.W. (1997). Causal Induction: The power PC theory versus the Rescorla-Wagner model. *In the Proceedings of the Nineteenth Annual Conference of Cognitive Science Society* p. 55-60. Hillsdale: Erlbaum.

Chang, T. M. (1989). Semantic memory: facts and models. *Psychological Bulleting, 99,* 199-220.

Chapman, G. B, & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition, 18,* 537-545.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic model of causal induction. *Journal of Personality & Social Psychology, 58,* 545-567.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99,* 365-382.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104,* 367-405.

Chertkow, H., Bub, D. N., & Seidenberg, M. (1989). Priming and semantic memory loss in Alzheimer's disease. *Brain & Language, 36*, 420-446.

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

Damasio, A. R., & Damasio, H. (1992). Brain and language. *Scientific American, 267,* 88-95.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature, 380,* 499-505.

Dowe, P. (1992a). An empirical defense of the causal account of explanation. *International Studies in the Philosophy of Science, 6,* 123-128.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General, 120,* 339-357.

Faye, J. (2001). Backward Causation. In *Stanford Encyclopedia of Philosophy.* Online at http://plato.stanford.edu/ entries/causation-backward/.

Fischler, I. (1977). Associative facilitation without expectancy in a lexical decision task. *Journal of Experimental Psychology: Human Perception & Performance, 3,* 18-26.

Fleischman, S. (1990). *Tense and narrativity.* Austin: University of Texas Press.

Freyd, J. J. (1987). Dynamic mental representation. *Psychological Review, 94,* 427-438.

Friedman, W. J. (1990). *About time.* Cambridge/MA: MIT Press.

Friedman, M. J. (2002). Arrows of time in infancy: The representation of temporal-causal invariances. *Cognitive Psychology, 44,* 252-296.

Gabrieli, J. D. E., Fleischman, D. A., & Keane, M. M. (1995). Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science, 6,* 76-82.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 11,* 501-518.

Graf, P., Squire, L., R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10,* 164-178.

Hempel, C. (1965). *Aspects of scienctific explanation.* New York: The Free Press.

Hermann, D. J., & Chaffin, R. (1986). Comprehension of semantic relations as a function of the definition of relations. In F. Klix and H. Hagendorf (Eds.), *Human memory and cognitive capabilities. Mechanisms and performance.* Elsevier Science Publishers B. V. (North-Holland).

Hitchcock, C. (1997). Probabilistic Causation. In *Stanford Encyclopedia of Philosophy.* Online at http://plato.stanford.edu/ entries/causation-probabilistic/.

Hofstadter, D. R. (2001). Epilogue: Analogy as the core of cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 499-538). Cambridge, MA: MIT Press.

Hull, C. L. (1943). *Principles of Behavior, an Introduction to Behavior Theory.* Oxford, England: Appleton-Century.

Hume, D. (1739/1978). *A treatise of human nature.* Oxford: Clarendon Press.

Hummel, J. E., & Holyoak, K., J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427-466.

Hummel, J. E., & Holyoak, K., J. (2001). A symbolic-connectionist theory of relational inferences and generalization. *Psychological Review.* (in press).

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulbon, K. R. (1996). Brain activation modulated by sentence comprehension. *Science, 274,* 114-116.

Kintsch, W. (1992). How readers construct situation models for stories: The role of syntactic cues and causal inferences.. In A. F. Healy, S. M. Kosslyn, & R. M. shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William Estes* (pp. 261-278). Hillsdale, NJ: Erlbaum.

Klix, F. (1980a). General psychology and the investigation of cognitive process. *Zeitschrift für Psychologie, 188,* 117-139.

Klix, F., van der Meer, E., & Preus, M. (1984). Semantische Relationen: Erkennungsaufwand und psychophysiologische Reaktionstendenzen. In F. Klix (Ed.), *Gedächtnis - Wissen - Wissensnutzung.* Berlin: DVW.

Kluwe, R. H., Wolke, D., & Bunge, B. (1982). Zur kategorialen Organisation semantischer Information bei 10jaehrigen Kindern und bei Erwachsenen. *Sprache & Kognition, 1(1),* 15-26.

Krüger, F., & van der Meer, E. (1999). Zeitbezüge im semantischen Gedächtnis. Abstracts der 4. Fachtagung der Gesellschaft für Kognitionswissenschaften KogWiss99, Universität Bielefeld.

Lewis D, (1993). *Causal explanation.* In D.-H. Ruben (Ed.), *Explanation.* Oxford readings in Philosophy. Oxford University press.

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly 2/4,* 261-4. Reprinted in Sosa, E. and M. Tooley (Eds.). *Causation.* Oxford University Press.

Mackie, J. L. (1974). *The cement of the universe.* Oxford: Clarendon Press.

Martin, A. & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology, 11(2), Special Issue: Cognitive neuroscience.* 194-201.

Mason, R. A., & Just, M. A. (2001). *How the brain hemispheres process causal information in text: a multiple process theory of comprehending causality.* Abstracts of the Psychonomic Society 42nd annual meeting. p. 103.

McClelland, J. L., & Rumelhart, D. E., & the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstrcuture of cognition.* (Vol. 2). Cambridge, MA: The MIT Press.

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11,* 1-37.

McNamara T. P., & Sternberg, R. J. (1983). Mental models of word meaning. *Journal of Verbal Learning & Verbal Behavior, 22,* 449-474.

McNamara, T. P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review, 99,* 650-662.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 558-572.

Menzies, P., & Price, H. (1993). Causation as a Secondary Quality. *British Journal for the Philosophy of Science, 44,* 187-203.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227-234.

Michotte, A. E. (1963). *The perception of causality.* New York: Basic Books.

Middleton, F. A., & Strick, P. L. (1991). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science, 266,* 458-461.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21,* 863-883.

Moss, H. E., Tyler, L. K., & Jennings, F. (1997). When leopards lose their spots: Knowledge of visual properties in category-specific deficits for living things. *Cognitive Neuropsychology, 14,* 901-950.

Münte, T. F., Schlitz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature 395,* 71-73.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading of activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226-254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. In D. Besener & G. Humphreys (Eds.), *Basic processes in reading: visual word recognition.* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 1003-1019.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Ockham of, W. (1980). Quodlibeta, in *Opera Theologica* IX, St. Bonaventure: St. Bonaventure University Press.

Parsons, L. M., & Osherson, D. (2001). New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cerebral Cortex, 11,* 954-965.

Pavlov, I. (1927). *Conditioned Reflexes.* New York, NY, US: Oxford University Press.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. A. & Raichle, M. E. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience, 1,* 153-170.

Plaut, D. C (1995). Semantic and associative priming in a distributed attractor network. In *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 37-42). Hillsdale, NJ: Erlbaum.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13,* 564-574.

Postman, L., & Keppel, G. (1970). *Norms of word associations.* New York: Academic Press.

Price, C. (1998). The functional anatomy of word comprehension and production. *Trends of Cognitive Psychology, 33,* 43-63.

Quillian, M. R. (1968). Semantic memory. In M. Minksy (Ed.), *Semantic information processing.* Cambridge, MA: MIT Press.

Quillian, M. R. (1969). The teachable language comprehender: a simulation program and theory of language. *Communications of the ACM, 12,* 459-476.

Reichenbach, H. (1956). *The direction of time.* Berkeley: University of California Press.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations on the effectiveness of reinforcement. In A. H. Black and W. F. Prokasy (Eds.). *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world.* Princeton, NJ: Princeton University Press.

Schacter, R. L., & Buckner, D. L. (1998). On the relations among priming, conscious recollection, and intentional retrieval: Evidence from neuroimaging research. *Neurobiology of Learning & Memory, 70(1-2), 284-303. Special Issue: Proceedings of the sixth conference on the neurobiology of learning and memory.* Brain and memory: From genes to behavior.

Schreuder, R. F., D'Arcais, G. B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45,* 339-354.

Shammi, P., & Stuss, T. D. (1999). Humor appreciation: A role of the right frontal lobe. *Brain, 122,* 657-666.

Shank, R. C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into knowledge structures.* Hillsdale, NJ: Erlbaum.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bowwer (Ed.). *The Psychology of Learning and Motivation, 21* (pp. 229-261). New York: Academic Press.

Shelton, J. R., & Martin, R. C. (1992). How semantic is semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1191-1210.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model of semantic decision. *Psychologcial Review, 81,* 214-241.

Spellman, B. A., Holyoak, K. J., & Morrison, R. G. (2001). Analogical priming via semantic relations *Memory & Cognition, 29,* 383-393.

St George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: Engagement of the right hemisphere during discourse processing. *Brain, 122,*1317-1325.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereoaxic atals of the human brain.* New York: Thieme.

Thagard, P. (1989). Explaining disease: Correlations, causes, and mechanisms. In *F. C. Keil, R. & A. Wilson, (Eds.), Explanation and cognition.* (pp. 255-276) Cambridge, MA, US: The MIT Press.

Thagard, P. (2000). *Coherence in thought and action.* Cambridge, MA, US: The MIT Press.

Thorndike, E. L. (1911). *Animal intelligence: experimental studies.* New York: Macmillan.

Tulving, E. (1972). Episodic and semantic memory. In: E. Tulving, W. Donaldson, *Organization of memory.* Oxford, England: Academic Press.

Tulving, E., Hayman, C. A., & Macdonald, C. A. (1991). Long-lasting perceptual priming and semantic learning in amnesia: A case experiment. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17,* 595-617.

Tyler, L. K., & Moss, H. E. (1997). Functional properties of concepts: Studies of normal and brain-damaged patients. *Cognitive Neuropsychology, 14,* 511-545.

Tyler, L. K., Moss, H. E., & Jennings, F. (1995). Abstract word deficits in aphasia: Evidence from semantic priming. *Neuropsychology, 9,* 354-363.

Van der Meer, E., & Schmidt, B. (1992). Finale, kausale und temporale Inferenzen - Analyse ihres kognitiven Hintergrundes. *Zeitschrift für Psychologie, 200,* 303-320.

Van der Meer, E., Beyer, R., Heinze, B., & Badel, I. (2002). Temporal order relations in language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 770-779.

von Wright, G.(1971). *Explanation and Understanding.* Ithica, New York: Cornell University Press.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B. R., & Buckner, R.L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science, 281,* 1188-1191.

Waldmann, M. R. (1996). Knowledge based causal induction. In D. R. Shanks, K. J. Holyoak and D. L. Medin (Eds.). *The Psychology of Learning and Motivation, 34: Causal learning* (pp. 47-88). San Diego: Academic Press.

Waldmann, M. R., & Holyoak, K. J. (1992). Predicitve and diagnsotic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121,* 222-236.

Wasserman, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science, 4,* 298-302.

White, P. A. (1988). Causal processing origins and developement. *Psycholgical Bulletin, 104,* 36-52.

Woods, R. P., Grafton, S. T., Watson, J. D., Sicotte, N. L., & Mazziotta, J. C. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Jounral of Computer assisted Tomography, 22,* 153-165.

Wu, M., & Cheng, P. W (1999). Why causation need to follow from statistical association: boundary conditions for the evaluation of generative and preventive causal powers. *Psychological Science, 10,* 92-97.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1196-1207.

A. Appendix

Table A1:

Unnormed causal and filler word pairs of Experiment 1. FSA is the forward strength of association and BSA is backward strength of association. The strength of association was determined retrospectively.

Causal word pair		FSA	BSA	Filler word pairs	
abuse	trauma	0.00	0.00	angel	barrel
alcohol	accident	0.00	0.00	basin	academy
bacteria	ulcer	0.00	0.00	basket	kite
burn	pain	0.00	0.03	beauty	compass
clouds	rain	0.03	0.37	bedtime	gallon
crash	injury	0.00	0.00	bubble	velvet
drought	famine	0.02	0.00	colony	sleeve
drug	relief	0.00	0.00	couch	dough
fall	bruise	0.02	0.00	dancer	clock
grief	suicide	0.00	0.00	diabetes	penny
mutation	cancer	nn	0.00	drill	guest
pepper	sneeze	0.01	0.04	envelope	lettuce
pollen	allergy	0.00	nn	faculty	empire
pressure	blast	0.00	0.00	ferry	apple
quake	damage	0.00	nn	flush	chowder
salt	thirst	0.00	0.00	gate	cord
scratch	bleeding	nn	0.01	helmet	arch
slope	slide	0.00	0.00	highway	anchor
smoke	cough	0.00	0.04	horn	pickle
spill	stain	0.01	0.00	investor	priming
stress	fatigue	0.02	0.00	king	ceiling
stroke	amnesia	nn	0.00	lamb	bearing
sweat	smell	0.00	0.07	leopard	river
sweets	cavity	0.00	0.00	miracle	ginger
traffic	noise	0.01	0.03	mousepad	justice
UVlight	tanning	nn	nn	needle	pianist
vacuum	suction	nn	nn	paint	gamble
virus	fever	0.00	0.02	painting	grammar
vitamin	health	0.00	0.18	pants	bandage
water	cooling	nn	0.08	pigeon	adult
weaving	fabric	0.00	0.00	point	king
wind	erosion	0.00	0.00	postman	therapy
Mean of strength of association		0.03	0.00		

Table A2:

Normed causal and associative filler word pairs for Experiments 2 to 4 and 6. FSA and BSA indicate the forward and the backward strength of association respectively. FR describes the difference of the frequency ratings. Positive values indicate a higher perceived frequency for the predictive direction, whereas negative values indicate a higher perceived frequency for the diagnostic direction.

Causal word pairs		FR	FSA	BSA	Associative word pairs		FSA	BSA
absence	withdrawal	-12	0.01	nn	acrobat	athletes	0.02	0
acid	corrosion	5	0	nn	agency	firm	0.02	0
alcohol	accident	-4	0	0	ambulance	rush	0.01	0
attack	defense	1	0.05	0	antelope	gazelle	0.03	0.03
bacteria	infection	0	0.01	0	atlas	dictionary	0.01	0
bang	deafness	-11	0	nn	basketball	teams	0	0.02
beat	bruise	8	0	0.05	bedroom	furniture	0.01	0
betrayal	distrust	13	0	0	caffeine	mountain	0.01	0
birthrate	population	1	nn	0	car	plane	0.01	0.03
carcinogen	tumor	2	nn	0	chipmunks	acorn	0	0.01
chromosome	gender	2	0.03	0	claw	dogs	0.01	0
compliment	blush	-4	0	0	cocktails	fruits	0	0.01
crime	arrest	-20	0	0.02	computer	apple	0.01	0.02
crush	damage	3	0	0.01	control	volume	0	0.01
dairy	diarrhea	-7	0	nn	dagger	fight	0.01	0
diet	hunger	16	0.01	0	decency	respect	0.03	0
disease	injection	13	0	0.01	elephant	zebra	0	0.01
drought	famine	-2	0	0.02	elevator	floor	0.01	0
drug	relief	23	0	0	email	attachment	nn	0
education	career	-7	0	0.03	engine	roar	0.02	0
espionage	treason	-4	0	0	envy	admire	0.03	0.03
eyedrops	dilation	-20	nn	0	family	sibling	0	0.03
fertilizer	growth	15	nn	0	forecast	weather	nn	0
fracture	cast	12	0	0	girl	maid	0	0.01
frequency	pitch	-3	0	0	glands	pituitary	0.03	nn
frowning	wrinkles	24	0.02	0	glass	window	0.01	0.02
gang	riot	-6	0	0	grab	pull	0.03	0.11
gases	explosion	-24	0	0	graduation	gown	0.02	0
genes	baldness	-10	0	0	graph	numbers	0.02	0
gold	wealth	7	0.02	0	ground	potatoes	0	0.02
hormones	mood	4	0	0	harbour	seaman	0	nn
humidity	sweat	27	0.02	0	insurance	estimate	0	0.01
illness	treatment	12	0	0.04	kill	theft	0	0.01
invitation	visit	10	0	0	kindness	sympathy	0	0.02

Causal word pairs		FR	FSA	BSA	Associative word pairs		FSA	BSA
joke	amusement	-1	0	nn	latin	medicine	0.02	0
lamp	heat	-3	0	0	lettuce	vegetables	0.08	0
lesion	scar	7	nn	0	lime	corona	0.02	nn
lightning	fire	4	0	0	lover	girlfriend	0.03	0.02
magnet	attraction	6	0.1	0.04	money	groceries	0	0.02
moon	tide	-3	0	0.02	mother	wife	0	0.03
movie	nightmare	0	0	0.02	newspaper	gossip	0	0.01
mutation	cancer	-19	nn	0	office	employment	0	0.02
nuts	allergy	-9	0	0	ounce	gallon	0.02	0
order	delivery	-9	0	0	painting	wall	0.05	0.02
pain	aggression	10	0	nn	paper	envelope	0	0.03
panic	escape	-3	0	0	patty	hamburger	nn	0.02
period	cramps	4	0	0.07	planter	farmer	nn	0.02
pollution	asthma	-7	nn	0	power	voltage	0.01	0.02
pressure	bursting	-22	0	0	propeller	helicopter	nn	0.03
sadness	crying	-20	0.05	0.13	protestants	baptist	0	0.01
salt	thirst	21	0	0	ring	emerald	0	0.03
scratch	blood	-7	0.01	0	round	screw	0	0.01
shampoo	tears	-6	0	0	sandwich	tomatoes	0	0
shock	scream	3	0	0	security	force	0.01	0
spice	flavor	-24	0.05	0.02	session	course	0.01	0
spill	stain	-7	0	0	shape	curve	0.01	0.02
sprain	swell	18	0	nn	shrimp	ocean	0.03	0
stress	fatigue	6	0	0.02	soup	cracker	0.03	0.04
study	pass	6	0	0	spray	roach	0.03	0.03
sunlight	freckles	5	0	0.02	story	passage	0	0
sweets	cavity	-27	0	0	terms	meaning	0.01	0
training	fitness	12	0	0	test	hypothesis	0	0.03
trash	stink	20	0	0	towers	skyscraper	0.02	0.05
trauma	coma	-9	0.01	0	tuba	saxophone	0.03	0.01
UVlight	tanning	-9	nn	nn	umbrella	tote	0	0.01
vacuum	suction	14	0.2	nn	uniforms	officers	0	0.01
virus	epidemic	-24	0	nn	vehicle	bicycle	0.02	0.01
wind	erosion	10	0	nn	vessel	vein	0.01	0.02
Sum of rating differences		0						
Mean strength of association			0.01	0.01	Mean strength of association		0.01	0.01

Table A3:

Normed associative stimuli of Experiment 6a.

Associated word pairs		FSA	BSA
checkers	chess	0.26	0.22
cigarette	smoke	0.45	0.32
corals	reef	0.38	0.40
cuss	swear	0.35	0.19
fist	knuckles	0.05	0.26
garden	landscape	0.00	0.01
halloween	pumpkin	0.14	0.03
leadership	team	0.03	0.00
nouns	adjectives	0.04	0.33
patient	doctor	0.37	0.03
positive	negative	0.63	0.60
pyramid	egypt	0.35	0.34
razor	blade	0.29	0.24
reunion	gathering	0.00	0.02
robber	thief	0.36	0.22
town	city	0.53	0.31
Mean strength of association		0.26	0.22

Table A4:

Unrelated filler word pairs of Experiment 4.

Unrelated word pairs		Unrelated word pairs		Unrelated word pairs	
ambulance	window	girl	agriculture	patty	kitchen
ankle	farming	glands	sailor	phone	switch
archer	phonebook	glass	rush	planter	power
basin	academy	grab	screw	point	queen
basket	kite	grass	fist	posters	hamburger
beauty	compass	gymnastic	vegetables	potatoes	insult
bedtime	tomatoes	harbor	garage	printer	angel
brush	dices	insurance	icecream	radiation	jockey
bubble	velvet	investor	priming	report	bike
caffeine	sky	kill	clock	revolting	roach
chef	fear	king	ceiling	rise	bank
clown	map	lamb	bearing	roles	cats
conductor	groceries	landscape	maths	roof	seaman
consulate	door	latin	disgust	round	roar
cookie	nose	lead	curve	salad	respect
couch	dough	leather	pull	savage	airport
dancer	liquid	lemon	soccer	security	floor
deer	pencil	leopard	river	shape	aluminum
diabetes	penny	lettuce	bars	ship	e-mail
diamond	gear	medicine	passage	smock	plug
disk	ground	mile	apron	spray	theft
doorbell	architect	miracle	ginger	square	indian
drill	guest	money	piano	store	session
eagle	child	mouse	light	story	dinner
eggs	liar	mousepad	justice	survivor	cup
elbow	pistol	mouth	actor	tea	graph
elevator	force	needle	currency	therapy	barrel
engine	glove	office	mirror	traffic	armrest
fabric	soup	onions	sphere	tree	maid
fairy	officers	page	tuxedo	truck	zebra
gate	cord	pants	bandage	voltage	mountain
gentlemen	chapter	parents	weather	water	boxer

Table A5:

German stimuli of Experiment 5 (Priming).

Causal word pairs					
Alkohol	Unfall	Keim	Entzündung	Salz	Durst
Ausrutschen	Fallen	Knall	Taubheit	Säure	Ätzung
Ausschütten	Fleck	Kochen	Dampf	Schlag	Beule
Bakterien	Infektion	Komet	Krater	Schneesturm	Chaos
Baldrian	Schlaf	Kratzer	Bluten	Schwüle	Schweiss
Betrug	Misstrauen	Krebs	Metastasen	Sonnenlicht	Bräune
Biss	Schmerz	Lähmung	Hinken	Stirnrunzeln	Falten
Blitzschlag	Feuer	Lampe	Wärme	Stress	Erschöpfung
Dämpfe	Schwindel	Langeweile	Gähnen	Süssigkeiten	Karies
Dünger	Wachstum	Laser	Verbrennung	Training	Fitness
Ehebruch	Scheidung	Läuse	Juckreiz	Trockenheit	Hungersnot
Erdbeben	Zerstörung	Medikament	Linderung	Übelkeit	Erbrechen
Flamme	Rauch	Menthol	Frische	Vakuum	Sog
Flut	Panik	Mond	Gezeiten	Verkehr	Lärm
Gene	Glatze	Müll	Gestank	Verstauchung	Schwellung
Gewürz	Geschmack	Nüsse	Allergie	Virus	Fieber
Gold	Wohlstand	Pfeffer	Niesen	Wind	Erosion
Hormone	Stimmung	Reibung	Ladung	Witz	Lachen

B. Appendix

B.1 Experiment 1

One of the most important tasks every organism has to face is to learn about the causal texture of the world. Knowledge about cause-effect relations allows us to predict the future or explain present events. Knowing for example, that a specific mushroom causes stomach disease allows us to predict what will happen when somebody eats this mushroom or explain why somebody suffers from stomach pain. Causal relations exist in a large number of domains (e.g., medicine, physics, biology, etc.)

On each trial you are going to see a cross in the middle of the screen. Please fixate your eyes on that cross.
After the cross disappears, a word will flash in the middle of the screen for 2 seconds. Then a second word will appear. Your task is to judge whether there is a causal relation between these words or not, i.e. if the event described by the first word causes or is caused by the event described by the other word.
Please press the "C" button on the keyboard, if there is a causal relation between these words.
Please press the "NC" button on the keyboard, if there is no causal relation between these words.

Be as quick and as accurate as possible with your decision.

After your judgment a new trial will start with the cross.

B.2 Experiment 2

One of the most important tasks every organism has to face is to learn about the causal texture of the world. Knowledge about cause-effect relations allows us to predict the future or explain present events. Knowing for example, that a specific mushroom causes stomach disease allows us to predict what will happen when somebody eats this mushroom or explain why somebody suffers from stomach pain. Causal relations exist in a large number of domains (e.g., medicine, physics, biology, etc.)

On each trial you are going to see a cross in the center of the screen. Please fixate your eyes on that cross.
After the cross disappears, a word will flash in the center of the screen for 1 second. Then a second word will appear. Your task is to judge whether there is a causal relation or just an association between these words.

Causal Relation:

If the event described by the **first word** *CAUSES* **or** *IS CAUSED BY* the event described by the **second word**, please press "**C**".

$+$		**fire**	**heat**	
$+$		**heat**	**fire**	**C**

Association:

If the events described are related in some ways, but do not cause each other, please press "**N**"

$+$		**cat**	**mouse**	**N**

You will start with practice trials. You will get feedback, if your answer was wrong and/or if you were to slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately.
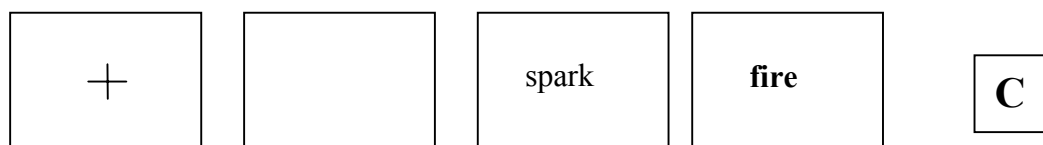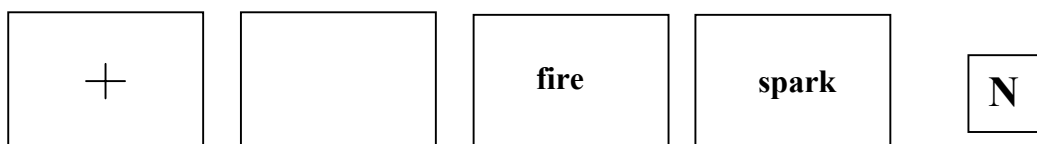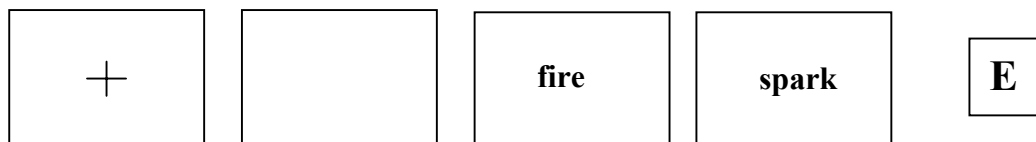After your judgment a new trial will start with the cross.

B.3 Experiment 3

Predictive block

On each trial you are going to see a cross in the center of the screen. Please fixate your eyes on that cross.

After the cross disappears, a word will flash in the center of the screen for 1 second. Then a second word will appear. Your task is to judge whether or not the first word is the cause of the second word.

If the event described by the **first word** *IS THE CAUSE OF* the event described by the **second word**, please press "C".

+		spark	**fire**	C

If the event described by the **first word** *IS NOT THE CAUSE OF* the event described by the **second word**, please press "N".

+		**fire**	**spark**	N

You will start with practice trials. You will get feedback, if your answer was wrong and/or if you were too slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately.

After your judgment a new trial will start with the cross.

Diagnostic block

On each trial you are going to see a cross in the center of the screen. Please fixate your eyes on that cross.

After the cross disappears, a word will flash in the center of the screen for 1 second. Then a second word will appear. Your task is to judge whether or not the first word is the effect of the second word.

If the event described by the **first word** *IS THE EFFECT OF* the event described by the **second word**, please press "C".

+		fire	spark	E

If the event described by the **first word** *IS NOT THE EFFECT OF* the event described by the **second word**, please press "N".

+		spark	fire	N

You will start with practice trials. You will get feedback, if your answer was wrong and/or if you were too slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately.

After your judgment a new trial will start with the cross.

B.4 Experiment 4

Causal condition

On each trial you are going to see a cross in the center of the screen. Please fixate your eyes on that cross.
After the cross disappears, a word will flash in the center of the screen for 1 second. Then a second word will appear. Your task is to judge whether or not there is a causal relation between these two words.

Causal relationship:
If the event described by the **first word** *CAUSES* **or is CAUSED by** the event described by the **second word**, please press "C".

+		fire	heat	C
+		heat	fire	C

If there is **no relationship** between the two words, please press "N".

+		ball	door	N

You will start with practice trials. You will get feedback, if your answer was wrong and/or if you were too slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately.
After your judgment a new trial will start with the cross.

Associative condition

On each trial you are going to see a cross in the center of the screen. Please fixate your eyes on that cross.

After the cross disappears, a word will flash in the center of the screen for 1 second. Then a

second word will appear. Your task is to judge whether or not there is an association between

these two words.

Association:
If the words are related in some meaningful way, please press "A".
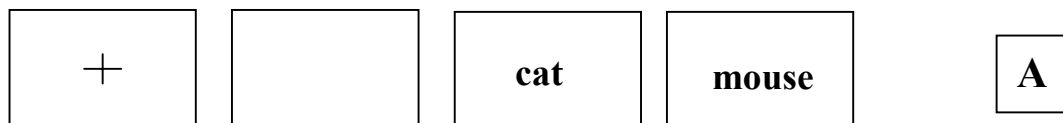
+		cat	mouse	A

If there is **no relationship** between the two words, please press "N".
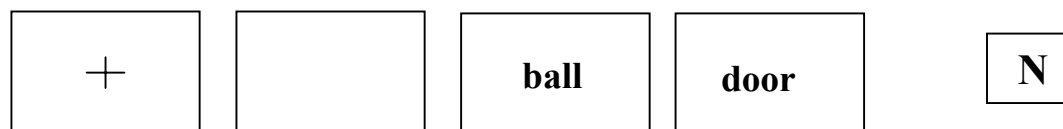
+		ball	door	N

You will start with practice trials. You will get feedback, if your answer was wrong and/or if you were too slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately.

After your judgment a new trial will start with the cross.

B.5 Experiment 5

Liebe Versuchsperson,

In diesem Experiment geht es darum, dass Du Wörter, die Dir auf dem Bildschirm präsentiert werden, so schnell wie möglich benennst.

Es wird Dir am Bildschirm ein Kreuz für 1 Sekunde gezeigt, bitte richte Deine Augen auf das Kreuz.

Nach dem Kreuz erscheint an der gleichen Stelle ********* oder ein Wort.

Danach erscheint ein weiteres Wort. Bitte spreche dieses zweite Wort so schnell wie möglich laut aus.

+	Haus	Auto	„Auto"
+	*********	Auto	„Auto"

Im Anschluß geht es wieder mit dem Kreuz weiter.

Da die Durchgänge sehr schnell sind, konzentriere Dich bitte so gut wie möglich. Am Anfang bekommst du einige Trainingsdurchgänge

Vielen Dank für Deine Teilnahme.

B.6 Experiment 6a

You are going to see 10 blocks of trials. Before each block you will be prompted with one of

the following two words " **Causal?"** and "**Associative?**"

If you were prompted with the word "**Causal**", then your task is to decide whether or not there is a causal relation between the two words.
If you were prompted with the word "**Associative**?" then you have to decide whether or not the words are associated with each other. If the words are **not related**, please press "**N**".
In all other ways the trials will be identical.

On each trial you will to see a cross in the center of the screen.
Please fixate your eyes on that cross.
After the cross disappears, a word will flash in the center of the screen for 1 second. Then a second word will appear.
Examples:

Causal Relation:
If the event described by the **first word** *CAUSES* **or** *IS CAUSED BY* the event described by the **second word**, please press "**X**".

+		**fire**	**heat**	**X**
+		**heat**	**fire**	

Association:
If the events described are related in some meaningful way, please press "**X**"

+		**cat**	**mouse**	**X**

No relation:

+		**ball**	**door**	**N**

There will be a break between each block for 16 seconds.
You will start with practice trials. You will get feedback, if your answer was correct, wrong or too slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately.
After your judgment a new trial will start with the cross.

Experiment 6b (fMRI scanner)

You are going to see 10 blocks of trials. Before each block you will be prompted with one of

the following two words " **Causal?"** or "**Associative?**".

If you were prompted with the word "**Causal?**", then your task is to decide whether or not there is a causal relation between the two words.
If you were prompted with the word "**Associative?**" then you have to decide whether or not the words are associated with each other. If the words are **not related**, please press the **right button** (*middle finger*) on your response box.
In all other ways the trials will be identical.

On each trial you will to see a cross in the center of the screen.
Please fixate your eyes on that cross.
After the cross disappears, a word will flash in the center of the screen for 1 second. Then a second word will appear.
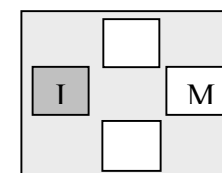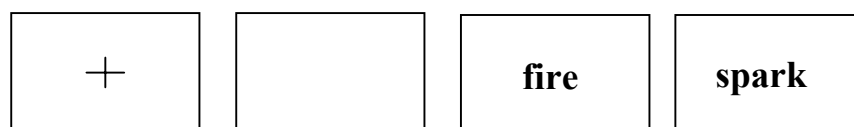*Examples:*

 Causal Relation:
If the event described by the **first word** *CAUSES* **or** *IS CAUSED BY* the event described by the **second word**, please press the **left button** (*index finger*) on your response box.

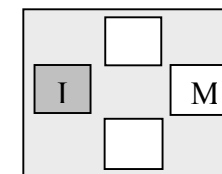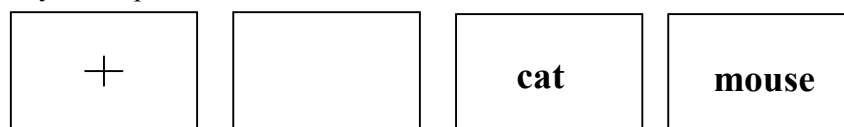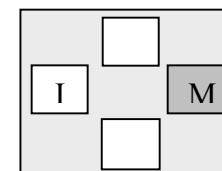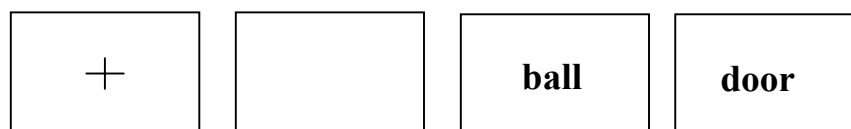| | | **spark** | **fire** | |

or

| | | **fire** | **spark** | |
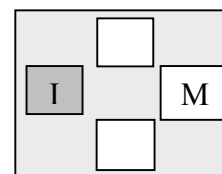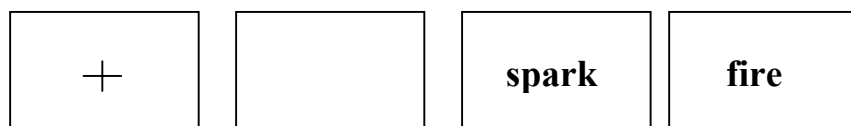
Association:
If the events described are related in some meaningful way, please press the **left button** (*index finger*) on your response box

| | | **cat** | **mouse** | |

No relation: Please press the **right button** (*middle finger*) on your response box

| | | **ball** | **door** | |

There will be a break between each block for 16 seconds.
You will start with practice trials where you will get feedback, if your answer was correct, wrong or too slow, therefore please respond as quickly as possible, but be also sure to answer the question accurately. After that a new trial will start with the cross.

C. Questionnaire

You will see a list of pairs of words. Between these two words there is a **causal relation** i.e. the event described by the word in column **A can cause or can be caused by** the event described by the word in column **B**.

If there is **no causal relation** between these two words, then please check "**NO**" in the causal relation box.

If there is a **causal relation** between these two words, then there are two things to do:

1. Check "**YES**" in the causal relation box
2. Imagine that the event described by the word in column **A occurs 100 times**. **How often does** then the event described by the word in column **B also occur**?

Here are some examples:

A	B	Causal Relation?		If event in column A occurs 100 times, how often does then event in column B occur?										
		NO	*YES*	*0*	*10*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	*90*	*100*
fire	heat		X											X
nightmare	horror movie		X		X									
door	pinball	X		X										

If you think that there is a causal relation between fire and heat (fire causes heat), check the "YES" box. Imagine that the event fire did occur 100 times, how often did then heat also occur? If you think that heat occurred 100 times then check the "100" box.

If you think the event in column B occurs e.g. 5 or 65 times, then put the "X" on the line between "0" and "10" or "60" and "70" respectively.

Lebenslauf

Daniela B. Sellner
Geboren am 11. 7. 1973 in Griesbach

1979-1983	Grundschule Ruhstorf a. d. Rott
1983-1992	Maristengymnasium Fürstenzell
1992	Abitur
1992-1993	Studium der Biologie an der Universität Regensburg.
1993-1999	Studium der Psychologie an der Universität Regensburg.
1996	Vordiplom
1996-1997	Studium der Psychologie an der Vanderbilt University in den USA.
1997	Praktika in der Ergonomieabteilung der Firma Siemens und am MPI für psychologische Forschung.
1999	Diplom in Psychologie Titel der Arbeit: ‚Abhängigkeit asymmetrischer Hemisphärenaktivierung von räumlich-visuell gesteuerter Aufmerksamkeit'
1999-2002	Wissenschaftliche Mitarbeiterin am Lehrstuhl für Handlungspsychologie und Forschungsmethoden an der Universität Göttingen.
2000-2001	Visiting Scholar an der University of California in Los Angeles.