ELSEVIER

Cognition 82 (2001) 27–58

COGNITION

www.elsevier.com/locate/cognit

Estimating causal strength: the role of structural knowledge and processing effort

Michael R. Waldmann*, York Hagmayer

Department of Psychology, University of Göttingen, Gosslerstrasse 14, 37073 Göttingen, Germany

Received 12 February 1998; received in revised form 16 August 2000; accepted 18 May 2001

Abstract

The strength of causal relations typically must be inferred on the basis of statistical relations between observable events. This article focuses on the problem that there are multiple ways of extracting statistical information from a set of events. In causal structures involving a potential cause, an effect and a third related event, the assumed causal role of this third event crucially determines whether it is appropriate to control for this event when making causal assessments between the potential cause and the effect. Three experiments show that prior assumptions about the causal roles of the learning events affect the way contingencies are assessed with otherwise identical learning input. However, prior assumptions about causal roles is only one factor influencing contingency estimation. The experiments also demonstrate that processing effort affects the way statistical information is processed. These findings provide further evidence for the interaction between bottom-up and top-down influences in the acquisition of causal knowledge. They show that, apart from covariation information or knowledge about mechanisms, abstract assumptions about causal structures also may affect the learning process.

Keywords: Causal reasoning; Causal models; Structural knowledge; Processing effort

1. Introduction

Causal learning is central for our survival. Causal knowledge allows us to anticipate harmful or gratifying events, and to plan actions to achieve goals. Despite the

* Corresponding author. Tel.: +49-551-393784; fax: +49-551-393656.

E-mail addresses: michael.waldmann@bio.uni-goettingen.de (M.R. Waldmann), york.hagmayer@-bio.uni-goettingen.de (Y. Hagmayer).

0010-0277/01/$ - see front matter
PII: S0010-0277(01)00141-X

fact that a large number of philosophers and psychologists agree on the importance of research on causality, no unitary concept has evolved so far. The main question of how we distinguish causal relations from accidental sequences of events remains highly debated.

2. Competing theories of causal learning

2.1. Covariation view

In the past 30 years, philosophers and psychologists have become increasingly interested in probabilistic notions of causality. Our knowledge about causal relations, such as "Smoking causes lung cancer", is often based on the observation of *covariations* between causes and effects. A number of philosophers have proposed a notion of causality that reduces causal relations to observable statistical laws (e.g. Eells, 1991; Salmon, 1971; Suppes, 1970). Roughly, it has been proposed that *causes alter the probabilities of their effects*. This idea has been adopted by psychologists who propose that causal induction involves the acquisition of knowledge about *contingencies* between causes and effects (e.g. Jenkins & Ward, 1965; Wasserman, Chatlosh, & Neunaber, 1983). Formally, an (unconditional) contingency can be defined as the difference between the conditional probability of a target effect E given the presence of a potential causal factor C and its probability given the absence of the factor (i.e. $P(E|C) - P(E| \sim C)$). If this difference is greater than 0, the contingency suggests that C is a facilitatory or generative cause, and if it is smaller than 0 then C may be inhibitory or preventive for the occurrence of the effect. Some researchers have assumed that these probabilities are estimated on the basis of stored frequency information (Cheng & Novick, 1992).[1]

A different, currently popular approach models causal learning as acquisition of associative weights (e.g. Shanks & Dickinson, 1987; Wasserman, Elek, Chatlosh, & Baker, 1993). However, Chapman and Robbins (1990) have proven that, at least in situations with one single cause and one effect, the asymptotic associative weight obtained by the Rescorla–Wagner learning rule (Rescorla & Wagner, 1972) corresponds to the results of applying the contingency rule (see also Cheng, 1997, for analyses of the relation between contingency theories and the Rescorla–Wagner rule). Thus, according to the covariation view causal induction is bottom-up learning about statistical contingencies between causes and effects.

One fundamental problem of the covariation view is that a set of events can be statistically analyzed in a large number of ways. The world is rife with correlations. Thus, attempts to attend to all statistical relations that can be estimated within sets of events would clearly surpass the capacity of human learners. Furthermore, not all statistical covariations within a set of events are causally meaningful. Spurious

[1] More recently, Cheng (1997) has proposed a different measure of causal power in which contingencies are normalized with respect to the base rates of the effect. Our experiments which focus on the role of causal co-factors are indifferent to the question of whether participants use contingencies or the more sophisticated causal power measure suggested by Cheng (1997).

correlations between causally unrelated events are a classic example of how covariations may be misleading. For example, even though there may have been a statistical correlation in the late 1960s between the length of hair of male students and being accepted at Berkeley, this correlation does not indicate a direct causal relation. Alternative ways of measuring statistical relations within a set of events may yield opposite assessments so that a principled method is needed to distinguish causally meaningful statistical indicators from indicators that cannot be directly linked to causal influences.

2.2. Mechanism view

The fact that causality is only indirectly reflected in covariation information has led to the alternative view that causal learning is mainly guided by prior knowledge about causal mechanisms (e.g. Ahn, Kalish, Medin, & Gelman, 1995; Bullock, Gelman, & Baillargeon, 1982; Shultz, 1982). According to this view, knowledge about causal mechanisms that connect causes and effects primarily underlies the learning process. Ahn et al. (1995), for example, gave participants who were requested to assess a causal relation a choice between covariation information or information indicating mechanisms. The experiments demonstrated a clear preference for information that provides evidence for or against mechanisms. The Koslowski (1996) experiments also point to the great relevance of mechanism information. According to her view, covariation information and prior knowledge about mechanisms interact. Identical covariations are viewed as indicators of greater causal strength when they are deemed plausible. For example, mentioning the plausible finding that increasing the number of visiting hours leads to higher recovery rates of patients yielded higher ratings of causal relatedness than the implausible finding that patients with heavy drapes on their windows recover faster (Koslowski, 1996, Experiment 5). Furthermore, only plausible covariations tend to be attended to. The plausibility is derived from prior knowledge about possible mechanisms.

The basic problem of the mechanism view is that it does not explain how knowledge about mechanisms is acquired in the first place (see Cheng, 1993). It may be true that learners tend to use specific prior knowledge when it is applicable to a new situation. However, a complete theory of causal learning has to address the question of how people learn about causal relations when no prior learning about similar situations has taken place. Causal model theory (see Waldmann, 1996) provides a third type of theory that attempts to reconcile the covariation approach with the view that prior assumptions about the causal domain guide the induction process even when no specific knowledge about mechanisms is available.

2.3. Causal model view

In the past few years a novel approach of causal learning has emerged that attempts to solve the methodological problems of the covariation view without postulating that an elaborate knowledge base about causal relations is already in place. According to this view, learners enter the learning task with *abstract* assumptions about hypothetical causal structures if more concrete knowledge is not available. A number of

philosophers and computer scientists have developed formal theories (e.g. Bayesian network theories) in which graphical representations of causal structures guide learning and inferences (e.g. Pearl, 1988, 1996, 2000; Spirtes, Glymour, & Scheines, 1993). Structural knowledge about hypothetical *causal models* encodes assumptions about the causal status of events (causes vs. effects) and about hypothetical direct and indirect causal relations. It does not embody knowledge about mechanisms of causes or effects although this information may be used if it is available. One important function of causal models is that they indicate which covariations are relevant for the estimation of causal power. Unlike knowledge about specific mechanisms, abstract structural assumptions do not imply the sign (inhibitory vs. generative cause) or strength of a causal relation. The structure selects the relevant statistical indicator. The sign and strength of the target causal relation are then fully dependent on the statistical structure of the learning input.

2.3.1. Causal priority

One example of how contingency assessments may be guided by prior structural assumptions is the fact that statistical relations are typically insensitive to one of the most fundamental properties of causal relations, *causal priority*: we know that causes precede their effects, and not the other way around. By contrast, contingencies do not encode directionality information. In situations in which the presence of a cause increases the probability of observing its effect, it is ceteris paribus true that the presence of the effect increases the probability of observing its cause. Thus, simple probabilistic relations do not represent the fundamental property of causal priority. In more complex situations with multiple causes and effects, the interpretation of the observed statistical evidence is crucially dependent on the assumed pattern of causal directionality. For example, three correlated events may be causally interconnected in various ways. Two events may play the role of alternative causes of the third event (common-effect model) which implies that either of the causes may generate the effect individually, or that the two causes may interact. In a different causal scenario the two events may be characterized as alternative effects of the third event, their common cause (common-cause model). In this situation, the causal model implies that the two effects are independently generated by their common cause, and that there should be a spurious covariation among the two effects. These different structural implications are a consequence of the specific pattern of causal directionality linking the events within a causal model.

Waldmann and Holyoak (1992) have therefore proposed a causal model theory which postulates an interaction between hypothesized causal models and the learning input. The main idea is that the structure of the causal models directs the interpretation of the learning input which in turn may modify the initial causal model. The impact of prior assumptions about causal directionality has been demonstrated in a number of experiments. In general, these experiments have shown that identical learning inputs were treated differently depending on the structure of the causal model invoked for the interpretation of the learning experiences (Waldmann, 1996, 2000, in press; Waldmann & Holyoak, 1990, 1992; Waldmann, Holyoak, & Fratianne, 1995). Waldmann (2000), for example, presented participants with learn-

ing inputs in which they received information about substances in the blood of fictitious patients as cues and had to learn to diagnose a novel disease. Learning was affected by initial instructions in which these substances were initially either characterized as causes of the disease (predictive learning) or as effects of the disease (diagnostic learning). In particular, cue competition among the causes (i.e. blocking) could only be shown when the cues were described as causes but not when they were characterized as effects. This finding can be explained on the basis of the view that learners used causal models when processing the learning input.

2.3.2. Causal co-factors

Causal priority is only one aspect of abstract prior causal knowledge that constrains the assessment of contingencies. The present article focuses on the role of prior knowledge in the choice of a method of estimating causal strength. In Section 2.1, unconditional contingencies were mentioned as empirical indicators of probabilistic causality. However, unconditional contingencies may even be observed between events that are not directly causally related, or the contingencies may not express the true causal relation between the two events. Problems always occur when there are additional factors which are correlated with the two observed events. These additional factors may be the cause for spurious correlations, or they may alter the observed statistical relation and therefore convey a wrong impression of the true causal relation. Philosophers (e.g. Cartwright, 1983; Eells, 1991; Salmon, 1980) and cognitive psychologists (Cheng, 1993; Cheng & Novick, 1992; Spellman, 1996a,b; Waldmann, 1996; Waldmann & Holyoak, 1992) have therefore proposed *conditional contingencies* as indicators of causality. Conditional contingencies refer to contingencies between two events, the potential cause C and the potential effect E, conditional upon alternative causal factors K_i being kept constant, i.e. as $p(E|C.K_1.K_2\ldots K_n) - p(E|\sim C.K_1.K_2\ldots K_n)$. An isolated period denotes an "and", and each K_i a choice between the presence or the absence of the co-factor.

For example, suppose we want to test the hypothesis that smoking (C) causes heart disease (E). Furthermore, we assume that smoking is correlated with coffee drinking (K_1), which may also be a cause of heart disease. To test our hypothesis, we should assess the conditional contingencies between smoking and heart disease in the subpopulation of coffee drinkers (K_1) and people who do not drink coffee ($\sim K_1$). If we then discover that smoking equally leads to heart disease in both subpopulations, we may conclude that smoking is an independent cause of this disease.

Even though the exact mechanism of taking co-factors into account is highly debated, a large number of studies have been presented in the past few years that show that participants are capable of taking co-factors into account when assessing the causal relation between a cause and an effect (e.g. Baker, Mercier, Vallée-Tourangeau, Frank, & Pan, 1993; Price & Yates, 1993). Spellman (1996a,b), for example, presented participants of an experiment with the task of assessing the causal impact of two unknown colored liquids on the growth of plants. In the learning materials the unconditional contingencies were kept constant, whereas the conditional contingencies varied, or vice versa. (These situations are examples

of Simpson's paradox which will be discussed below.) The main finding was that participants focused on subgroups instead of the total distribution. Thus, they tended to estimate conditional contingencies in which the potential impact of the co-factor was kept constant instead of unconditional contingencies.

Conditional contingencies, however, are not foolproof indicators of causality either. As Cartwright (1983) points out, this method only yields correct results when the subgroups are properly selected. Conditionalizing on the wrong variables may lead to erroneous contingency estimates. An instance of this problem is known in the philosophical and statistical literature as *Simpson's paradox*, first discovered by Pearson in 1899 (see Cartwright, 1983; Eells, 1991; Pearl, 2000; Simpson, 1951). Simpson's paradox describes the fact that a given contingency between two events which holds in a given population can disappear or even be reversed in all subpopulations when the population is partitioned in certain ways. Cartwright (1983) cites a study on the graduate admissions of Berkeley that demonstrates the problem (Bickel, Hammel, & O'Connell, 1977). The graduate school of Berkeley was accused of discriminating against women. And indeed, at first sight the probabilities seemed to support the causal hypothesis that being a woman causes one to be rejected at Berkeley: the probability of admissions was higher for male students than for female students. However, the researchers looked at the data more carefully. When the admissions were analyzed separately for each department, one by one, the correlation between gender and admission completely disappeared. The reason for this was that women tended to apply to departments with higher rejection rates. Department by department women were accepted in the same ratio as male applicants, whereas across all the departments proportionally fewer women were admitted.

Table 1 from Eells (1991, p. 63) gives an example of how this can happen. In this example, Department 1 accepts 90% of the female and of the male applicants. Department 2 only accepts 20% of the female and of the male applicants. Thus, within each department male and female applicants are accepted in the same proportions. However, more female applicants apply to Department 2 which is harder to get in. Therefore, overall, across all departments, more than three times as many male applicants (83%) are admitted than female applicants (27%).

This example may lead to the methodological suggestion that it is always a good idea to partition into subsets of events, and compute conditional contingencies. However, this strategy may also lead to false assessments. The reason why in the Berkeley admissions case the analysis should be based on the department level is that the departments are *causally relevant* for the effect under investigation. The

Table 1
Fictitious relative frequencies of admission to graduate program (after Eells, 1991, p. 63)

	Department 1 (%)	Department 2 (%)	Total (%)
Male	81/90 (90)	2/10 (20)	83/100 (83)
Female	9/10 (90)	18/90 (20)	27/100 (27)
Contingency	0	0	0.56

departments decide about the admissions, and not the whole university.[2] If, by contrast, it had been shown that the contingencies reverse when the applicants were partitioned on the basis of their roller skating skills, this would not count as an argument against sex discrimination (Cartwright, 1983). Only partitions by causally relevant variables are relevant for evaluating causal laws. If causally irrelevant variables were also considered, almost any contingency can be obtained by choosing the right partition of the event space.

Simpson's paradox has often been used to demonstrate how possible *confounds* can create a spurious covariation that misrepresents the underlying causal relation. An early example from medicine comes from a Gedankenexperiment by Blyth (1972). In his scenario a physician compares an old and, alternatively, a new treatment to two populations: one of urban and affluent patients, the other of rural and poor patients. The new treatment seems to have disastrous effects when compared to the old treatment. Yet the physician has the very definite impression that the new treatment works better. He is right, because the key to Simpson's paradox here is that the urban patients are more likely to suffer from a more severe variant of the disease, and the new therapy has been administered (in a not intentionally planned way) mostly to the urban population.

The practical relevance of this scenario has been demonstrated by a re-analysis of a study about the medical consequences of smoking (Appleton, French, & Vanderpump, 1996). In this study a sample of 1314 women in the mixed (urban and rural) district of Whickham (Newcastle-upon-Tyne) was examined in 1972, and 20 years later in 1992. Unexpectedly, smokers turned out to have a longer life expectancy than non-smokers. However, a closer look at the data revealed that in this study younger women, who generally have a greater life expectancy, were more likely to be smokers than the older women. In this situation, the causal impact of smoking can only be assessed if the causally relevant age confound is held constant.

Pearl (2000) has recently presented a formally rigorous analysis of the relationship between confounding and causal inference. In one of his examples a certain drug seems to reduce the recovery rate in the studied population while simultaneously increasing the recovery rates of the subgroups of males and females. Fig. 1 displays the causal structure underlying this situation. In the data underlying this situation males, who recover regardless of the drug more often than females, are also more likely to use the drug than females. Thus, the gender factor confounds the target relation between the drug and the recovery rate because it is causally related to both events. The assessment of the strength of the causal link between treatment and

[2] One reviewer presented the interesting possibility that in the Berkeley admission example a cynical administration may try to introduce gender discrimination by imposing high standards on departments to which more women are likely to apply. In this scenario, unlike in the original example, the act of setting the standards for the different departments is the crucial causal intervention to be assessed, which, along with the non-discriminatory policies of the naive individual departments, creates gender discrimination. Thus, it is appropriate to consult the overall distribution of admission rates because this information reflects the relevant causal factor, the number of admissions in each department, which is intentionally regulated by the administration on the basis of the gender of applicants.

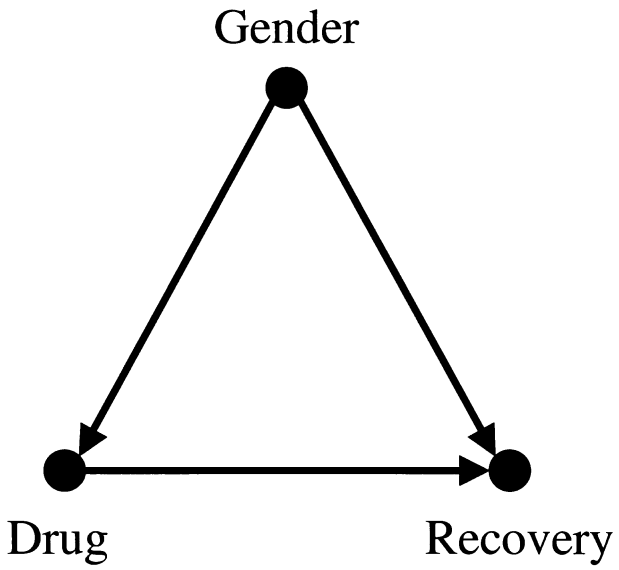

Fig. 1. A causal model underlying a variant of Simpson's paradox discussed by Pearl (2000, p. 178) (see text for further explanations).

recovery rate is only possible in this non-experimental situation when the second causally relevant factor, gender, is held constant.

Another instructive example of Simpson's paradox comes from an analysis of a decision making situation. Shafir (1993) has constructed an imaginary scenario in which a student who is facing an exam wants to decide whether to go on a vacation afterwards or to stay at home. He has taken 50 exams, half of which were followed by a vacation. Overall he recalls having been happy more frequently when staying at home, which seems to suggest to rather stay at home in the future. However, partitioning the exams into the ones he passed and the ones he failed reveals a different pattern. In each of these subgroups he recalls having been happy more often when he went on a vacation than when he stayed at home. The reason for the paradox is that the student went on a vacation more often after failing the exam than after passing, and that after failing an exam he often felt unhappy. As in the previous examples, a response to the causal question of how the act of traveling will affect his well-being is only provided by an analysis of the subgroups in which success in the exam has been held constant.

2.3.3. No causes in, no causes out

What all these examples show is that causal induction is crucially dependent on prior causal knowledge. This has been neatly described by the title "no causes in, no causes out" of a chapter of Cartwright (1989) which was borrowed here. New causal relations may be induced using contingency estimates. However, the contingencies only reflect causal relations when the observations are partitioned on the basis of

causally relevant rather than irrelevant variables. The causal relevance of these partitioning variables has to be established prior to the new induction task. Thus, Simpson's paradox exemplifies the basic assumption of causal model theory that the interpretation of the learning input is based on prior assumptions about general properties of the causal situation. According to this theory, prior knowledge and bottom-up learning interact. Causal assessments are based on contingencies displayed in the learning input. However, the way contingency estimates are derived from the data is dependent on prior assumptions about the causal structure underlying the observed events. These prior assumptions need not include specific knowledge about mechanisms. Assumptions about the causal role are sufficient. Experiments 1 and 2 investigate the role of prior assumptions about causal relevance.

Prior assumptions about potential alternative causes are only one example of knowledge-based constraints on contingency assessments. Eells (1991, chapters 3 and 4) has extended the analysis by Cartwright (1983). He has shown that it is also appropriate to hold constant independent events that are not direct alternative causes of the target effect but interact with the cause in producing the effect. For example, specific physiological features may not by themselves increase the probability of heart disease but they may affect the capacity of smoking to cause the disease. In this example, it is also appropriate to control for these physiological parameters even though these factors are not direct alternative causes of the disease. The term "causal relevance" should therefore be used in the broad sense of alternative causes of the target effect or independent events that interact with the target cause.

2.3.4. Causal structures

So far causal structures have been discussed in which multiple causes converge on a common effect. However, there are other possible causal structures linking three events.

Fig. 2 displays four different causal structures linking three events. (These structures were investigated in Experiment 3.) In all four structures a potential cause C is directly linked to a potential effect E. The structures differ with respect to the causal role of the event K. The decision about whether a co-factor is causally relevant is only one factor determining whether this event should be kept constant or not. Correlated events may play other causal roles in complex causal networks.

Fig. 2A shows a common-effect structure in which C and K both converge on the effect E. For example, C may represent smoking, E heart disease, and K may be an alternative potential cause of heart disease, such as coffee drinking, or it may interact with smoking in producing heart disease (e.g. certain physiological characteristics). The discussion in the last section has demonstrated that in such situations it is appropriate to hold K constant when assessing the relationship between C and E.

Fig. 2B,C displays two variants of *causal chains*. In both chains, C directly causes E but in Fig. 2B ("chain-1") the third event K is located at the end of the chain as an effect of E, and in Fig. 2C ("chain-2") it is located at the beginning of the chain as a cause of C. For example, C and E may again represent smoking and heart disease. In Fig. 1B K may represent shortness of breath, a possible effect of heart disease,

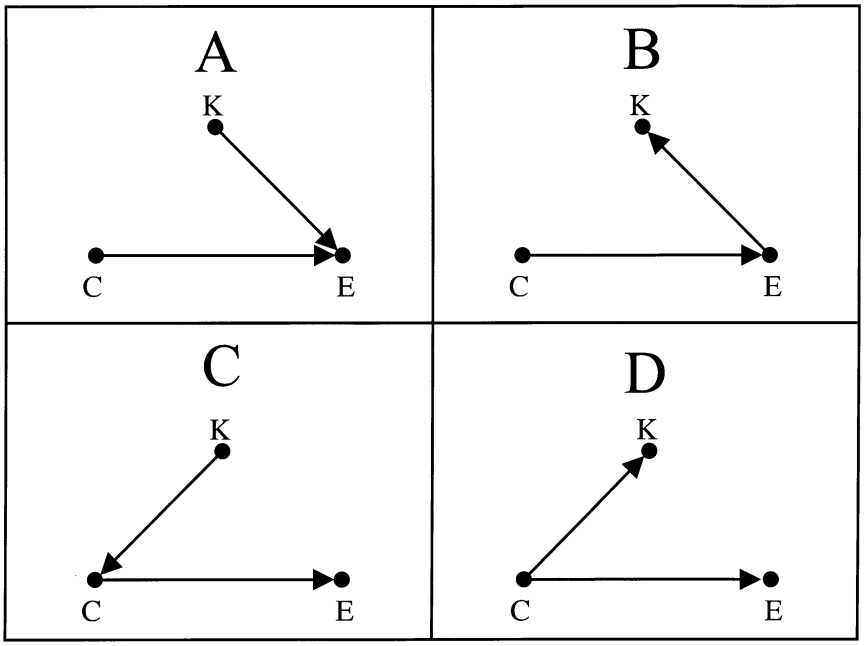

Fig. 2. Four causal structures with three linked events, a common-effect structure (A), two types of causal chains ("chain-1" in B, "chain-2" in C), and a common-cause structure (D). Event *C* represents a potential cause, event *E* an effect, and *K* a correlated third event that is differentially linked to the other events in the four structures.

whereas in Fig. 1C *K* may stand for tobacco commercials, a possible cause of smoking. Which events should be held constant when assessing causal relations in such chains?

In general, one *must not* hold constant *intermediate* events that mediate a causal relation between a target cause and an effect when assessing the indirect causal relation between these two events (see Cartwright, 1989; Eells, 1991). For example, when assessing the indirect causal relation between *K* and *E* in the chain displayed in Fig. 2C it is inappropriate to hold fixed the causal intermediate *C*. This can easily be understood with an example in which the only way of *K* to cause *E* is through the intermediate event *C*. In this case *K* may increase the probability of *E* by increasing the probability of *C*. However, if the value of *C* is fixed at some level, *K* cannot raise the probability of *E*. Holding *C* fixed in such a situation would *screen off K* from *E*. Hence, the conditional contingency between *K* and *E* would be zero, and the true indirect causal relation between these two events would not be reflected in the contingencies. With more complex structures in which further links connect *K* and *E*, the values may not go to zero but the contingency between these two events would still be distorted when the intermediate *C* is held fixed.

For similar reasons, one also must not hold fixed events that are causal *subsequents* of the target causal relation (see Eells, 1991). Fig. 2B displays an example of

such a situation (chain-1). A causal chain is shown in which C causes E, and E in turn causes K. When assessing the causal contingency between C and E it would be inappropriate to hold fixed the terminal effect K. Holding fixed this event would increase or decrease the probability of E which in turn may dampen the contingency between C and E. In summary, causal intermediates, effects of these intermediates, and effects of the target effect *must not* be held constant. Holding these events fixed may dampen the observed contingency between the target cause and the target effect. This way the relation between the observed statistical contingencies and the underlying causal relation will be distorted.

The situation changes for causal chains in which K is the initial cause producing C which in turn causes E (chain-2). Under restricted conditions it may be unnecessary to hold K fixed in such a structure when assessing the causal relation between C and E but in other situations it is necessary to hold the initial cause fixed. Holding this event fixed is unnecessary when the chain satisfies the *Markov condition*. In Markov chains the contingency between C and E is independent of the initial cause K. The Markov condition states for the causal chain depicted in Fig. 2C that the conditional probability of E given C in the presence of K equals the conditional probability of E given C in the absence of K (i.e. $P(E|C.K) = P(E|C. \sim K)$). Similarly the conditional probability of E provided C is absent should also stay the same regardless of whether K is present or absent (i.e. $P(E| \sim C.K) = P(E| \sim C. \sim K)$). Thus, K has no influence on E independent of C. C perfectly screens off K from E, that is once K has produced event C the occurrence of event E is solely dependent on the capacity of C to produce the effect, and K does not play any further role. In such a situation, it does not make a difference whether K is held constant or not.

However, Markov chains are only a special case of possible chains. Eells (1991) discusses an example in which K does not cause E but causes C *and* interacts with C in the production of E. In this example, K stands for a physiological condition, C represents smoking, and E lung cancer. The physiological condition K itself does not cause lung cancer but it increases the likelihood of smoking. Furthermore, the physiological condition K interacts with smoking in the production of cancer. Smoking may prevent lung cancer when the condition is present, and cause it when it is absent. In such a situation, the causal impact of smoking can only be assessed when the presence or absence of the physiological condition is controlled for. This example shows that sometimes it is appropriate to hold fixed the initial cause of a causal chain. However, most researchers who work with normative models, such as Bayesian networks, assume that causal chains honor the Markov condition unless there is clear empirical evidence disproving this assumption (e.g. Glymour & Cooper, 1999). One goal of Experiment 3 was to investigate whether lay people share this simplifying default assumption with the experts.

Fig. 2D shows a common-cause structure in which a cause C causes both event E and event K. To use once again the smoking example, C may represent smoking and E and K the two independent effects heart disease and pollution due to exhaled smoke. The standard assumption for such structures is that the common cause C screens off E from K. Thus, conditional on C the two effects are independent of each other. Once the status of C is known, the predictability of either effect cannot be

improved by knowing whether the collateral effect is present or absent. A famous example by the philosopher Reichenbach (1956) describes a situation in which two actors contract a stomach disease at the same time. However, the coincidence is rendered statistically independent once the common cause, food poisoning, is known.

In such common-cause structures it is not necessary to hold fixed the collateral effect (e.g. *K* in Fig. 2D) when assessing the causal contingency between the cause and the other effect (e.g. *C* and *E* in Fig. 2D). In fact, similar to the situation in which *K* is the initial cause of a Markov chain (Fig. 2C), it should not make a difference whether the collateral effect is held fixed or not. However, there may be situations in which the two effects interact even when the common cause is partialed out (see Eells, 1991). These interactions may be due to additional unknown causal connections linking the two effects in the underlying causal network. If that was the case, holding fixed one of the interacting effects would actually distort the contingencies between the cause and the other effect for the same reasons that subsequent effects may dampen contingencies. In such situations, the collateral effect (e.g. *K*) *must not* be held fixed. In summary, with common-cause structures it is generally advisable not to hold fixed one of the effects when the causal relation between the common cause and some other effect is to be assessed.

The previous analyses can be summarized in the following rule (see Eells, 1991, p. 206; Pearl, 1996): when assessing a causal relation between an event *C* and an event *E*, we *must* hold fixed all non-subsequent causal factors that are causally relevant to *E* (i.e. alternative causes or interacting events), and we *must not* hold fixed subsequent causal factors that are directly or indirectly caused by *C*.

This analysis shows again the interdependence between prior assumptions and the estimation of causal strength. Prior assumptions about causal structures guide the choice of a statistical indicator of causal strength (see also Pearl, 2000, for guidelines on the analysis of more complex causal models). Experiment 3 investigates whether human learners conform to these normative rules in situations with three covarying causal events.

2.4. Processing effort

There is a problem with the strategy to hold fixed potential co-factors to assess causal relations on the basis of contingencies. This strategy only guarantees success when *all* co-factors are controlled. Otherwise one can never be sure that some unknown factor may distort the relation between the underlying causal relation and the observed contingencies. One problem with this prescription is, of course, that it is unlikely that we will ever know all co-factors. But even if we did, it may be practically impossible to follow the strategy of holding every causal factor constant. As the philosopher Dupré (1993) has argued, if a specific effect has n alternative causes then 2^n possible combinations of the presence and absence of these factors would have to be inspected to assess a specific causal relation. Thus, a situation with 31 factors may already exceed the potential of the human population to fill each cell

with only one person. Dupré suggests that in science we should abandon the methodological goal to come up with a *complete* causal story (see also Cherniak, 1986).

Completeness may even be less of a plausible goal for everyday causal cognition. Causal knowledge is typically acquired under capacity and time constraints that require pragmatic solutions. We do not need complete knowledge; what we do need is knowledge that allows us to make satisfactory predictions and plan successful actions. Partial causal knowledge does not necessarily lead to failure. We may very well successfully act or predict even on the basis of partially incorrect causal knowledge. Also we need to weigh our processing resources against the potential gain of more differentiated knowledge. When knowledge is easily acquired we may attempt to obtain a more complete account, whereas in situations in which our resources are more heavily demanded a more sketchy representation may suffice. Herbert Simon (1956, 1982) has coined the term "satisficing" to describe knowledge or cognitive strategies that may not be perfect but satisfactory in the contexts in which this knowledge is to be used.

Schaller and O'Brian (1992) conducted experiments that demonstrate the role of processing load in group stereotype formation. In their experiments they presented participants with information about the performance of two groups, Group A and Group B, in solving anagram tasks. Overall Group B outperformed Group A but this apparent superiority was reversed when the length of the anagrams was taken into account. Group A tended to show lower performance because it preferably attempted to solve the more difficult, longer anagrams than Group B. Thus, the statistical structure formed a variant of Simpson's paradox. The results show that the participants of the experiment ignored the co-factor, difficulty of anagram, unless they were explicitly instructed to organize the data according to this factor, or when they were given more time for processing the information. Thus, these results support the hypothesis that processing effort is an important constraint (see also Fiedler, Walther, Freytag, & Stryczek, in press). Sperber, Cara, and Girotto (1995) also have recently postulated a trade-off between processing effort and task requirements in logical reasoning (see also Gigerenzer & Goldstein, 1996).

3. Overview of experiments

The goal of the experiments that will be presented in this article is to show that prior knowledge and processing effort guide the strategies of causal strength estimation. A number of previous studies have investigated causal structures in which multiple causes converge on a common effect, and have demonstrated that learners take potential causal co-factors into account (e.g. Baker et al., 1993; Price & Yates, 1993; Spellman, 1996a,b). In contrast, the aim of the present research is to explore the conditions in which participants *use or ignore* co-factors and other types of correlated events. The experiments use causal structures that embody a variant of Simpson's paradox. Simpson's paradox is a powerful methodological tool to address these issues as it permits a clear empirical dissociation between judgments that are based on the overall distribution as opposed to judgments based on the analysis of

subgroups. In general, participants received identical learning inputs in the contrasting conditions of the experiments and were asked identical questions about the observed events. However, prior assumptions about causal roles were manipulated by means of differential initial instructions. Both knowledge-based constraints and processing constraints were systematically explored. Concerning knowledge-based constraints, Experiments 1 and 2 investigated whether prior assumptions about the causal relevance of a co-factor affect the strategies of contingency assessment. These experiments focused on common-effect structures with potential co-factors that were either described as causally relevant or causally irrelevant. Experiment 3 explored other potential causal roles of correlated events including common-cause structures and causal chains. The impact of processing load was studied by manipulating the presentation mode of the learning input. The learning input was either presented in list format (Experiments 1 and 3) or in a trial-by-trial presentation (Experiment 2). The subgroups were either presented in a blocked arrangement (Experiments 1 and 2) or in an intermixed fashion (Experiments 2 and 3). These manipulations are motivated by the assumption that intermixed presentation of the subgroups, especially in the context of trial-by-trial learning, makes it harder to analyze the statistical structure of subgroups. Thus, we expected that co-factors would only be taken into account when the processing load was low enough to permit conditionalization. In the cognitively more demanding situations we expected participants to resort to less demanding strategies of data processing.

4. Experiment 1

The Cartwright (1983) analysis of the Berkeley admission problem has demonstrated that in situations with multiple potential causes of a common effect, conditional contingencies should be assessed which control for causally relevant co-factors. Causally irrelevant co-factors should be ignored. Controlling for causally irrelevant co-factors may lead to distortions of the underlying causal relations. Because the causal relevance of the co-factors has to be established prior to the induction task, this requirement is a clear example of how prior knowledge guides the assessment strategies. The goal of Experiment 1 is to investigate whether human participants are sensitive to this crucial distinction between causally relevant and causally irrelevant co-factors.

In this experiment all participants received identical learning input and identical instructions to rate the strength of a causal relation between a target cause and a target effect. The participants' task was to assess the causal relation between irradiation of fruit and the quality of fruit. The crucial manipulation involved a grouping variable according to which the learning cases could be partitioned into two exhaustive sets of cases. This grouping variable was either characterized as causally relevant or causally irrelevant. In the condition with the causally relevant grouping variable, information about two types of fruit was given which may affect the efficacy of irradiation. In the condition with the causally irrelevant grouping information, it was pointed out that the fruit samples were randomly assigned to two

investigators. A third "control" condition presented no information about possible groupings. The crucial question was whether participants would conditionalize the causal relation on the grouping variable when it was characterized as causally relevant but ignore this variable when it was causally irrelevant.

A variant of Simpson's paradox was imposed on the learning data to empirically test this question. Whereas the contingency computed across all cases pointed in one direction (positive or negative), the contingencies within each subgroup had signs in the opposite direction. This reversal allows for a clear test of whether participants focus on the total distribution or on subgroups. If participants proved sensitive to the causal relevance of the grouping variable they should focus on the total distribution in the condition with the causally irrelevant grouping variable (as in the control condition with no grouping information), whereas participants who assumed causal relevance of this variable should analyze the subgroups separately. Thus, a clear reversal of the sign of the causal ratings is to be expected.

4.1. Method

4.1.1. Participants and design

The participants were 36 students from the University of Tübingen, Germany, who were randomly assigned to one of three conditions. Two of the three conditions varied the causal relevance of the co-factor (causally relevant vs. irrelevant co-factor). The third condition was a control condition in which no co-factor was presented.

4.1.2. Procedure and materials

The participants' task in this experiment was to assess the strength of the causal relation between irradiation of tropical fruit and the quality of the fruit. All instructions and materials (in this and all other experiments) were in German. Participants received written instructions in which they were told that importers of tropical fruit are trying to improve the quality of the fruit by irradiating them. However, so far it is unknown whether the irradiation has a positive, a negative, or no effect on the quality of the fruit. To assess the efficacy of irradiation, participants received information about the quality of samples of fruit that either had or had not been irradiated. The participants were handed a two-page list which contained information about 80 samples of fruit. Each sample was represented on a separate line, and for each sample participants could see whether or not the sample had been irradiated ("yes" or "no"), and whether the quality of this sample was "good" or "bad". Participants were instructed to study the list carefully in order to be able to assess whether irradiation has an effect or not. They were requested to express their impression on a rating scale that ranged from -10 ("irradiation leads to a strong deterioration of the quality") to $+10$ ("irradiation leads to a strong improvement of the quality"). Participants were allowed to take notes on a separate sheet of paper.

Participants were assigned to one of three conditions. Participants in all conditions saw the same list with the 80 cases, and received the same rating instructions. Thus, all participants were requested to rate the overall strength of the causal relation

Table 2
Contingencies and relative frequencies of fruit with good quality within and across the subcategories A and B of the grouping variables (Experiment 1)

	A	B	Total
Irradiation	16/36 (0.44)	0/4 (0.00)	16/40 (0.40)
No irradiation	3/4 (0.75)	5/36 (0.14)	8/40 (0.20)
Contingency	− 0.31	− 0.14	+ 0.20

between irradiation and quality of fruit. The only difference was that in two of the conditions an additional grouping variable was mentioned which either was causally relevant or irrelevant. The third condition represents a control condition in which no grouping variable was introduced.

Participants in the condition with the *causally relevant* variable were told that there are two types of fruit, Taringes and Mamones. Additionally it was pointed out that it was expected that irradiation affects these two types of fruit differently. Furthermore, information was added to the list which indicated that one of the two pages showed Taringes, and the other page Mamones. In the condition with the *causally irrelevant* grouping variable, participants were told that due to the large number of tests, the samples of fruit described on the two pages were randomly assigned to two different investigators. The participants in this condition saw the same list as in the condition with the causally relevant variable except that "Mamones" and "Taringes" were respectively replaced by "A" and "B" as a shorthand for the two investigators.

To test whether participants' contingency judgments reflect their prior assumptions about the additional grouping variables, the organization of the list corresponded to a variant of Simpson's paradox. Table 2 shows how the cases were distributed. The table displays the proportion of fruit that were of good quality after they were irradiated, and the proportion of fruit that were good without being irradiated. For example, within subgroup A 36 fruit samples were presented that were irradiated. Forty-four percent of these samples (i.e. 16 out of 36) had good quality after irradiation. As can be seen in Table 2, the arrangement of the cases resulted in a reversal of the sign of the contingencies within as opposed to across the grouping variable. Disregarding the grouping variable yields a positive contingency between irradiation and quality of fruit (+0.20). By contrast, the contingencies within each of the subgroups are negative (−0.31 and −0.14). For half of the participants, the mapping between irradiation and quality of fruit was switched so that these participants saw a symmetric situation with a negative overall contingency, and positive contingencies within the subgroups.

4.2. Results and discussion

Table 3 shows the results of this experiment. The signs of the ratings of the group who saw the negative overall contingency were reversed so that the two subgroups were comparable. The results show that participants behaved according to the Cart-

Table 3
Mean ratings of the causal relation between irradiation and quality of fruit (Experiment 1)

Relevant	Irrelevant	Control
−4.33	5.17	4.75

wright (1983) advice. The mean ratings in the control condition and in the condition with the irrelevant grouping variable were positive, and statistically indistinguishable from each other. Thus, participants in these two conditions believed that irradiation *raises* the quality of fruit. This finding indicates that participants based their assessments on the total distribution of cases while disregarding subgroups. By contrast, participants in the condition with the causally relevant grouping variable got the impression that the cause prevents the effect. These participants concluded that irradiation *lowers* the quality of fruit. The negative mean rating indicates that many participants computed contingencies for each subgroup separately before these contingency estimates were integrated. The mean rating of this group was very different from the mean ratings of the two other groups ($F(1, 33) = 71.1$, $P < 0.001$, MSE $= 9.71$).

5. Experiment 2

In Experiment 1 two different grouping variables were compared, type of fruit (Mamones and Taringes) and investigators (A and B). Experiment 2 attempted to replicate the results of Experiment 1 with a grouping variable that was kept constant across the two conditions. Thus, all participants saw identical cases, received identical rating instructions, and were informed about identical subcategories. The only manipulation involved a hint about the potential causal relevance of the co-factor describing the subcategories. Only half of the participants received this hint.

The participants' task in this experiment was to assess whether a newly developed medicine reduces the incidence of a symptom caused by a tropical disease. Again the learning data embodied a variant of Simpson's paradox. Overall the contingency between the medicine and the reduction of the incidence of the symptom was positive, whereas within the subgroups of males and females no influence could be seen (see Table 4). Two instruction conditions were compared which only

Table 4
Contingencies and relative frequencies of patients with the symptom (Experiment 2)

	Males (%)	Females (%)	Total (%)
Medicine	24/30 (80)	2/10 (20)	26/40 (65)
No medicine	8/10 (80)	6/30 (20)	14/40 (35)
Contingency	0	0	+ 0.30

differed in one sentence. In this sentence the hint was given that the medicine may affect male and female patients differently.

A second aim of Experiment 2 was to explore whether the findings of Experiment 1 generalize to more naturalistic learning situations in which information is presented sequentially in a trial-by-trial fashion.

A third goal of this experiment was to provide a test of the assumption that processing effort also influences learning strategies. The claim to be tested was that participants should tend to neglect subgroups when the separate analysis of the subgroups is cognitively demanding. Processing load was manipulated by presenting the male and female subgroups either in a blocked fashion or by presenting them in a random order. We expected that the participants would resort to the simpler strategy of estimating unconditional contingencies in the conditions that are more taxing for the cognitive system.

5.1. Method

5.1.1. Participants and design

The participants were 40 students from the University of Tübingen, Germany, who were randomly assigned to one of four conditions generated by the two factors causal relevance (hint vs. no hint) and presentation mode (blocked vs. random).

5.1.2. Procedure and materials

In this experiment all participants received initial written instructions that were identical except for one additional sentence in the condition with the hint. The instructions mentioned an increased incidence of tropical diseases due to international travel. One (fictitious) disease, "Ariana Tropicans", was singled out which causes a number of symptoms including large amounts of water in the extremities. Furthermore, it was pointed out that no medicines exist that alleviate the symptoms but that a new experimental medicine, "Caldivon", is currently being tested on 80 patients from a tropical clinic. According to the instruction, some of these patients receive the medicine, and others are in the control condition. Half of the participants of Experiment 2 who were in the hint condition were then told that Caldivon may have different effects on males and females as it is based on a hormone.

Then the instructions showed pictures of the front and back side of an index card which served as an example for the cards from the later learning phase. The front side gave information about the initials of the patient, the gender of the patient, the treatment ("Caldivon"), and the initial symptom ("large amounts of water in the extremities"). On the backside, the state of the symptom after the treatment was shown ("no water in the extremities"). The instructions pointed out that the task was to learn to predict the symptom described on the backside of the index card on the basis of the information on the front side. Furthermore, at the end of the instructions a rating scale was shown along with the instruction that the task would require the participants to assess the causal efficacy of the medicine by using a scale that ranged from 0 ("no effect") to +10 ("strong effect").

After the instructions participants received a stack of 80 index cards. The index

cards looked like the one presented on the instruction sheet. On the front sides information was given about the initials of the patient, gender, the initial symptom (all patients initially had "large amounts of water in the extremities") and whether Caldivon was given to the particular patient or not. On the backsides, the state of the symptom after treatment was shown ("water in the extremities" or "no water in the extremities"). Participants went through the stack by themselves by turning around one index card after the other. Taking notes or reordering the cards was not permitted. In the blocked condition the index cards were grouped according to the gender of the patients. In the random condition participants received the cases in a random sequence. After the learning task all participants received identical rating instructions and a rating scale. The task for all participants was to assess on the ten-point scale how Caldivon affected water in the extremities.

Table 4 shows the statistical structure of the learning domain. Overall the percentage of patients without water is higher in the treatment condition in which Caldivon was given than in the control condition in which no medicine was taken. The contingency between the medicine and the symptom amounts to +0.30. However, within the groups of male and female patients the contingencies are zero. Thus, the medicine appears to help the whole group but it does not have an effect within the subgroups of males and females.

5.2. Results and discussion

Table 5 shows the mean ratings of the participants in the four conditions. The experimental manipulations yielded a highly reliable interaction between the two factors hint and presentation mode ($F(1, 36) = 12.2$, $P < 0.01$, MSE = 3.62). This interaction is due to the fact that a clear lowering of the ratings, consistent with the use of the co-factor, was only observed in the condition with hint and blocked presentation of cases. This impression was confirmed by a post-hoc comparison of the four conditions (least significant difference test) which revealed that the condition with hint and grouped presentation proved significantly lower than the other three conditions which, in turn, did not statistically differ among each other. This pattern is consistent with the prediction that the hint led a significant number of participants to base their judgments on conditional contingencies (i.e. 0), whereas the participants in the groups without hint and with random ordering of cases seemed to have based their estimations on unconditional contingencies (i.e. +0.30).

The results obtained in this experiment show that prior knowledge and processing effort interacted. Experiment 2 showed a general tendency to ignore subgroups unless the causal relevance of the co-factor was explicitly pointed out *and* the

Table 5
Mean ratings of the causal relation between the medicine and the symptom (Experiment 2)

	With hint	Without hint
Grouped	1.6	4.9
Ungrouped	4.5	3.6

task of separating into subgroups was cognitively manageable. This pattern of results is consistent with the claim that both prior knowledge and cognitive effort constrain the learning strategies.

6. Experiment 3

The previous experiments have explored the conditions in which participants calculate contingencies between a potential cause and an effect conditional upon a third correlated event. This third event was either described as causally relevant or irrelevant for the cause–effect relation in question. In these experiments, the co-factor was typically introduced as a variable that potentially interacted with the target cause in producing the effect. This is a situation in which it is appropriate to hold the co-factor constant to assess the causal relation between the target cause and the effect (Eells, 1991). Common-effect situations in which a co-factor is linked to a target effect is only one type of structure that links three events. Other structures such as causal chains or common-cause structures also are possible.

Experiment 3 systematically explores different causal structures. In this experiment all events are described as causally relevant. However, the causal role in the underlying causal network is varied. Four conditions are compared in which the relationship of a target cause (C) to a target effect (E) is to be assessed. Cause C and effect E are directly linked in all causal structures. The causal role of the co-factor K is varied, however. Fig. 2 displays the four causal structures that can be generated given the constraint that C and E are directly linked, and K is linked to either C or E.

As elaborated in Section 2.3.4, if the goal is to assess the causal strength between C and E, K should be held fixed in the common-effect condition in which K is independently causally relevant for E (Fig. 2A), whereas it should not be held fixed in the chain in which K represents the final effect (Fig. 2B) and the common-cause situation in which E and K are collateral effects of C (Fig. 2D). In the causal chain condition in which K plays the role of the initial cause (Fig. 2C) it is permissible to ignore K if the chain embodies the Markov condition (i.e. K and E are independent conditional upon the intermediate cause C). If the Markov condition is violated, however, it is necessary to hold K constant.

Experiment 3 explores whether participants learning about these causal structures are sensitive to these different conditionalization requirements involving K. Four conditions were created in which identical learning data were presented to the participants. All participants were asked to assess the causal relation between C and E. The only difference was that the causal role of event K was manipulated through initial instructions yielding the four mental causal representations shown in Fig. 2.

As an additional factor we manipulated the abstractness of the learning materials. From an experimental point of view abstract materials, for which no prior knowledge about the learning events is available, are ideal because the four causal models can be mapped to the same set of three events, thus holding everything constant except for the initial instructions. According to causal model theory the use of

abstract causal models should suffice to induce different learning strategies. However, this prediction requires that participants actually use the initially instructed causal models when processing the learning input. It seems plausible that this may be difficult with extremely abstract materials in which the event labels do not even allow learners to distinguish whether the respective event represents a potential cause or a potential effect.

To investigate the potential role of familiarity with causal relations we compared the four causal models using two types of materials. In the abstract scenario we described the learning events as fictitious chemical substances (Anin, Billon, Cyran) which could be assigned arbitrary causal roles within the causal models. In the contrast condition, however, we used cover stories that presented plausible examples for the four models which were consistent with prior knowledge about the potential causal status of the mentioned events. We kept the causal relation to be assessed constant; all participants had to assess whether genetically altered fruit damages the gastrointestinal flora. However, in order to increase plausibility we varied the semantic contents of the co-factor *K*. For example, in the common-effect condition (Fig. 2A) *K* was described as bacteria (a possible alternative cause of gastrointestinal disease), whereas in the common-cause condition (Fig. 2D) *K* was described as an alternative symptom of the food, an allergic reddening of the skin. Although prior knowledge does not determine whether genetically altered food indeed leads to gastrointestinal disease or a reddening of the skin, the potential causal roles of the events are unambiguous. Reddening of the skin, for example, cannot possibly be the cause of genetically altered food.

Comparing the abstract with the plausible condition allows us to directly test the generality of the predicted effects and assess the role that prior knowledge about specific plausible mechanisms plays.

6.1. Method

6.1.1. Participants and design

Ninety-six students from the University of Göttingen, Germany, participated in this experiment. The participants were randomly assigned to one of the eight conditions generated by crossing the factor "plausibility of cover stories" (plausible vs. abstract) with the factor causal model (common effect, common cause, chain-1, chain-2) (see Fig. 2). Only students who had not taken any course in statistics or experimental methodology were admitted to participate in this experiment. Thus, students from a variety of areas participated in the experiment. Only five participants were (first semester) psychology students.

6.1.2. Procedure and materials

Participants in the conditions with the *abstract* cover stories were told that researchers in medicine are increasingly interested in metabolic processes. In particular, the researchers are investigating whether the substance Anin (event *C* in Fig. 2), by means of polymerization, stimulates the production of the substance Billon

(event *E*). It was pointed out that thus far it is unclear whether Anin influences the production of Billon or has no effect.

The four conditions in which the causal models were manipulated differed with respect to the assignment of the causal role of the third event (event *K*), substance Cyran. In all conditions Cyran was linked to Anin or Billon by means of a fictitious oxydation process. (The processes oxydation and polymerization were mentioned to add credibility to the cover stories. Other than that they did not provide any help as to distinguishing between causes and effects.) In the *common-effect condition*, participants were told that it is presumed that Cyran may also lead to Billon. Thus, event *C* as well as event *K* may independently cause event *E* (see Fig. 2A). Participants in the *common-cause condition* read instead that it was assumed that event *C* leads to event *E* and to event *K* (see Fig. 2D). (In all four conditions, the full description of the events was used.) In the *chain-1 condition* event *C* causes event *E* which in turn causes event *K* (see Fig. 2B), whereas in the *chain-2 condition* event *K* leads to event *C* which in turn leads to event *E* (see Fig. 2C). Before participants inspected the learning data they were informed that later they would be asked to assess the causal relation between event *C*, Anin, and event *E*, Billon.

After studying the instructions, all participants were requested to summarize what they had read. If anything was unclear, clarifications were given. As an additional help a diagram was placed in front of the participants which displayed the causal structure by linking the three events with arrows (similar to Fig. 2). The target relation (between Anin and Billon) was signified by a dotted arrow with a question mark placed above the arrow. Then participants in all four conditions received an identical two pages on which 100 observations of a fictitious medical study were described. The two pages showed a four-column list of the descriptions of the 100 cases. The first column numbered the observations. The second, third and fourth columns displayed whether Cyran, Anin, and Billon, respectively, were present or not ("yes" or "no"). The ordering of cases was randomized. Participants were permitted to take notes. After studying the cases all participants were given the same task. They had to rate the causal relation between Anin (*C*) and Billon (*E*) using a rating scale that ranged from 0 ("no influence") to +10 ("very strong stimulation").

Again the statistical structure of the learning data followed a variant of Simpson's paradox. Table 6 shows the statistical distribution of the conditions. The most important information involves the unconditional and conditional contingencies between events *C* and *E*. Whereas the unconditional contingency is positive

Table 6
Contingencies and relative frequencies of the presence of event *E* in the presence and absence of events *C* and *K* (Experiment 3) (see also Fig. 2)

	K (%)	~ *K* (%)	Total (%)
C	32/40 (80)	1/10 (10)	33/50 (66)
~*C*	8/10 (80)	4/40 (10)	12/50 (24)
Contingency	0	0	+ 0.42

(+0.42), the contingencies are zero when the potential co-factor *K* is held constant. All participants studied the same list with identical arrangements and ordering of the observations, and were requested to respond to the identical rating question. The only difference between the conditions was the assignment of different causal roles to event *K*, substance Cyran, in the initial instructions.

In the *plausible condition* cover stories were used that allowed participants to use prior knowledge about causal directionality during learning. Participants in the four conditions in which the causal model was manipulated were generally told that medical researchers were interested in diseases of the stomach. In particular, the researchers suspected that genetically altered fruit (event *C*) may damage the gastrointestinal flora (event *E*). As in the abstract condition it was pointed out that, thus far, it is unclear whether the new fruit has an effect or not. All participants had the same task; they were requested to assess the causal relation between fruit and damaging of the gastrointestinal flora. The four causal models were manipulated by mentioning a third event (event *K*) which, unlike in the abstract condition, differed across causal models. In the *common-effect condition* (Fig. 2A) bacteria were mentioned as an alternative potential cause of gastrointestinal problems. In the *common-cause condition* (Fig. 2D) a possible reddening of the skin was introduced as a potential second effect of genetically altered fruit. In the *chain-1 condition* (Fig. 2B) flatulence was mentioned as a potential effect of gastrointestinal disease, and in the *chain-2 condition* (Fig. 2C) it was pointed out that irradiation of the tropical fruit may be the cause of genetic alterations.

Apart from these instructions the learning procedure was modeled closely after the one used in the abstract conditions. Again participants were requested to summarize the instruction and diagrams were placed in front of participants which displayed the causal structure of the labeled events. Again the diagrams showed the instructed causal structure with arrows representing causal relations. The learning events were represented by pictorial symbols and verbal labels.

The same learning list as in the abstract condition was used, only the labels on top of the two sheets were exchanged in a way consistent with the respective condition. After studying the lists participants were requested to rate the strength of the causal relation between genetically altered food and the gastrointestinal flora by using a scale that ranged from 0 (“no influence”) to +10 (“very strong damage”).

6.2. Results and discussion

Table 7 displays the mean ratings in the eight conditions. The pattern of results corresponds to the philosophical analyses outlined in Section 2.3.4 (Eells, 1991; Pearl, 2000). The causal relation between the target cause (event *C*) and the target effect (event *E*) significantly differed across the four conditions ($F(3, 88) = 8.14$, $P < 0.01$, MSE = 6.12). This effect did not interact with the second factor which manipulated the plausibility of the cover stories ($F < 1$). Further analyses (least significant difference tests) revealed that this effect is solely based on the fact that the ratings for the target relation within the common-effect model proved significantly lower than the ratings for the other three conditions. Thus, no significant difference

Table 7
Mean ratings of the causal relation between event *C* and event *E* in the four conditions of Experiment 3 (see also Fig. 2)

	Common cause	Chain-1	Chain-2	Common effect
Abstract	4.92	4.67	6.58	3.00
Plausible	6.42	6.50	6.83	3.58
Total	5.67	5.58	6.71	3.29

was observed between the common-cause, chain-1, and chain-2 conditions. This pattern of results is consistent with the recommendation that one should conditionalize on a co-factor when this factor represents an alternative cause within a common-effect model. Although the ratings for this condition deviated from the conditional contingency (0) displayed in the data, it still was clearly lower than the ratings for the other three conditions. It is a well known fact that people have difficulties with contingencies of 0 especially when the effect also occurs in the absence of the cause (see Shanks, 1993). This may explain the slight increase in this condition.

The ratings for the other three causal models are consistent with the assumption that in these conditions the assessments were based on the unconditional contingency (0.42) which suggests that in these conditions event *K* was ignored. This is recommended by normative analyses for the common-cause model and the chain-1 model. For the chain-2 model in which the co-factor represents the initial cause of the chain this strategy is appropriate if it is assumed that chains honor the Markov condition, a standard assumption made by most researchers who use Bayesian models to represent causal structures (e.g. Glymour & Cooper, 1999). Whenever the Markov condition is valid, computations are greatly simplified by enabling the learner to focus on the direct causal links (e.g. between *C* and *E*). All indirect relations (e.g. between *K* and *E*) can be generated from the direct relations. The results of the present experiment suggest that learners untutored in statistics and methodology share the Markov assumption with experts, which allows it to simplify the task of learning about causal chains.

The second factor (plausibility of cover stories) also proved significant ($F(1, 88) = 4.25$, $P < 0.05$, MSE $= 6.12$). There was a general tendency of higher ratings in the plausible as compared to the abstract condition, especially in the common-cause and chain conditions. The increase in these conditions may be viewed as supporting Experiment 5 of Koslowski (1996) which showed higher ratings for identically described relations when the underlying causal relation seemed plausible as compared to an implausible relation.

Unlike in Experiment 2, participants were able to control for co-factors even though the presentation of cases was not blocked according to the states of the co-factor (see also Waldmann & Hagmayer, 1995, Experiment 2). However, this is consistent with the explanation that processing load is the crucial variable. In Experiment 2, the cases were presented in a trial-by-trial fashion which made it extremely hard to separate out the groups in the condition with intermixed presenta-

tion of cases. By contrast, Experiment 3 presented the cases in list format which allowed participants to go back and forth between cases. Furthermore, only in Experiment 3 were participants permitted to take notes.

7. General discussion

The goal of causal induction is to arrive at representations of objective causal relations. Typically causal relations cannot be observed directly but must be inferred on the basis of statistical relations between observable events (see Cartwright, 1989; Cheng, 1997). The presented studies focus on the problem that there are multiple ways of extracting statistical information from a set of events. Which method is appropriate is partly determined by the hypothesized causal structure underlying the observed patterns of events. As pointed out by Cartwright (1989), covariation is one of the most potent ways to *measure* causal capacities. However, like other measuring instruments it needs to be read properly.

Previous experiments have shown that assumptions about causal directionality influence how the participants of the experiments compute statistical contingencies (Waldmann, 1996, 2000, in press; Waldmann & Holyoak, 1992; Waldmann et al., 1995). The present set of studies focuses on how assumptions about the causal role of potential co-factors or other correlated events influence statistical estimation. The three presented experiments show that identical observations may yield radically different causal strength estimates depending on assumptions about the causal role of the learning events. All three experiments studied causal structures with three events in which the task consists of estimating the strength of the relation between a target cause and a target effect. The manipulations involved the role of a third event. The general goal was to explore the conditions in which participants would tend to take this third event into account and hold it constant or tend to neglect it.

In the past few years, philosophers and statisticians have analyzed the conditions under which it is necessary to hold a co-factor constant and when it is prohibited to hold it constant (Cartwright, 1989; Eells, 1991; Pearl, 1996, 2000). The behavior of the participants conformed surprisingly well to these philosophical analyses (see also Section 2.3.4). Experiments 1 and 2 show that, at least under favorable circumstances, participants tended to hold a co-factor constant when they assumed that it was independently causally relevant for the target effect, whereas they tended to ignore it when it was causally irrelevant. This strategy is in line with the Cartwright (1983) prescription only to conditionalize on causally relevant co-factors. Experiment 3 extends these findings to more complex structures. The results show that participants tended to conditionalize on co-factors when they represented potential alternative causes of the target effect (common-effect structure), whereas they ignored these events when they represented a collateral effect of the target cause (common-cause structure) or when they were part of a causal chain. This strategy implicitly follows the prescription to conditionalize on factors that are independent alternative causes or interact with the target cause and not to conditionalize on factors that are direct or indirect effects of the target cause (Eells, 1991). The results

about causal chains furthermore reveal that, at least in the type of task used in Experiment 3, untutored participants share the intuition with researchers in the area of Bayesian causal modeling (Glymour & Cooper, 1999) that it is reasonable to assume that chains embody the Markov condition. The Markov condition for chains states that indirectly related causal events are independent conditional upon the intermediate events.

The role of prior knowledge in the selection of processing strategies is an interesting example of how specific causal knowledge and abstract strategies of causal strength assessment are interconnected. It is true that knowledge about the causal relevance of the grouping variables is domain-specific (e.g. the fact that gender is a causally relevant factor). However, unlike in previous research on the transfer of specific causal knowledge which showed that learners tend to distort observed covariations in the direction of prior assumptions (e.g. Chapman & Chapman, 1967, 1969), this type of knowledge does not *directly* bias the sign or size of estimates of the strength of a causal relation between the target cause and the target effect. Whether or not a cause generates or prevents the effect is completely dependent on the structure of the learning input. To obtain the correct results, abstract knowledge has to be activated about what type of contingency assessment (e.g. unconditional or conditional) should be chosen. These abstract rules of causal strength assessment are then instantiated based on specific assumptions about the relevant causal factors at hand (e.g. gender). Thus, the dramatic reversals obtained in situations exemplifying Simpson's paradox are not due to selective processing of individual cases or knowledge-driven distortions of the contingency estimates. They rather are a natural consequence of unbiased processing of differentially grouped cases. Therefore, these results demonstrate that besides covariation and specific knowledge about mechanisms abstract knowledge about causal structure may also exert a powerful influence on reasoning and learning.

Prior assumptions about causal roles is only one factor influencing contingency estimation. Conditionalizing on events and computing conditional contingencies is sometimes an effortful process that may surpass the cognitive capacity of the learners. Experiment 2 provides evidence for the fact that processing effort also determines how contingencies are assessed. Despite prior knowledge about the causal relevance of a co-factors participants inappropriately tended to ignore this co-factor when the task of separating the learning events into subgroups was too demanding.

In summary, one of the general problems of causal induction is to select among the many possible ways of processing statistical information. Without prior assumptions about causal roles the amount of data and the number of necessary computations may easily surpass our capacity even in fairly simple domains. Assumptions about causal structure may greatly reduce the complexity of this task by implicitly encoding information about interdependencies and conditional independencies among events (see Pearl, 1988; Peng & Reggia, 1990; Spirtes et al., 1993). The presented experiments demonstrate that humans are attempting to make the task of learning cognitively manageable. The results of the experiments show that both assumptions about causal structure and processing demands prove to be potent pragmatic constraints on learning.

7.1. The role of the learning data

In all experiments, variants of Simpson's paradox were used in the design of the learning materials. The main reason for this choice was that Simpson's paradox offers a method to empirically decide whether participants conditionalized on a co-factor or computed unconditional contingencies. In situations modeled after Simpson's paradox these two types of contingencies reverse. However, the question may be raised whether the statistical structure of the learning data was always plausible. Experiment 1, for example, was modeled after an example discussed by Cartwright (1983) (the Berkeley admission problem) that was used to demonstrate that one should conditionalize on causally relevant (departments) but not causally irrelevant events (roller skating skills). Eells and Sober (1983) argue against this recommendation that in the latter case no paradox should arise in the real world since a causally irrelevant factor, such as roller skating skills, should turn out to be uncorrelated with the causal factors, for example the tendency to apply to the different departments. Similarly in Experiment 1 the assignment to investigators should turn out to be random unless there is a hidden causal relation involved. This observation led the philosopher Spohn (1990) to recommend to conditionalize on all temporally prior factors independent of their causal relevance. This should take care of causally relevant co-factors, and should not make a difference when irrelevant factors are held constant. However, from a psychological point of view it should be pointed out that these philosophical analyses assume the ideal case of statistical information that is representative for the population. In realistic learning situations the learning data may be sparse and therefore distorted so that accidental correlations may occur. Thus, it may be reasonable to let prior knowledge guide processing instead of relying too much on the representativeness of the learning data. Also, even if conditionalizing on all prior factors led to correct results, the present experiments show that participants, if anything, prefer to not conditionalize unless prior knowledge dictates it to take co-factors into account. Conditionalizing in many situations is a demanding process that participants rather avoid unless pressed. Thus, Cartwright's distinction between causally relevant and irrelevant co-factors seems to be adopted by human learners even when the statistical structure contradicts the irrelevance assumption.

The relatively strong influence of top-down knowledge despite statistical patterns in the learning input that contradict these assumptions generally raises the question of what role bottom-up information plays. The example of the statistical implications of causal irrelevance may be extended to other causal structures. Causal structures generally imply empirical patterns that may or may not hold in the learning data. For example, causal chains and common-cause structures imply intercorrelations among the involved events, whereas two independent causes of a common effect would be expected to be independent of each other (see Pearl, 1988). Thus, the statistical structure of the learning input presented in Experiment 3, which was identical in all four conditions, proved more or less compatible with the different causal structures suggested in the initial instructions.

The results of the experiments show little sensitivity to such potential

mismatches. Participants generally were not aware of the full statistical pattern presented to them. No participant became aware of the fact that the data contained a paradox (see also Fiedler et al., in press). The participants rather decided on the basis of their prior assumptions to follow a specific analytical strategy (e.g. to compute conditional or unconditional contingencies) without attempting to explore alternative strategies. Therefore, only a subset of the potential statistical information was generated.

There is little evidence that causal structures are induced solely by means of a complete analysis of the statistical information in the learning input. In this regard, our theory ("causal model theory") notably deviates from normative accounts that have been developed in philosophy and artificial intelligence research (e.g. Pearl, 1988; Spirtes et al., 1993). One of the most important goals of these theories is to develop methods for inducing causal structures from covariation information alone with minimal entering assumptions. However, the present experiments show that this strategy is simply too complex to be computationally feasible for humans. It is more plausible to assume that participants generally enter the task with concrete or abstract prior assumptions about the causal structure to be studied. The analysis of the learning input then allows the learner to fill in the missing information into the hypothesized causal models (e.g. statistical contingency information). Only if this information or other learning feedback is blatantly incompatible with the hypothesized causal models may the learner attempt to revise parts of the model. But revision of initial models certainly is a demanding task that learners tend to avoid unless pressed. Waldmann et al. (1995) have shown that given enough time some participants are eventually able to learn about domains that are inconsistent with prior assumptions but many participants were unable to revise the initial models within the given time frame. Unlike in the present experiments, these studies used a corrective learning procedure in which active predictions were required. Thus, the present studies may show that participants do not *explicitly* test structural implications embodied in causal models, whereas participants may *implicitly* become aware of incompatibilities when corrective feedback proves the inadequacy of the initial causal model (see also Ahn & Mooney, 1995; Schaller, Asp, Rosell, & Heim, 1996; Waldmann & Martignon, 1998).

This hypothesis raises the question of whether part of the reason for the strong impact of top-down assumptions may have been that the presentation of lists may have made it difficult to implicitly become aware of statistical relations that were not actively sought out. Whereas other learning procedures, such as trial-by-trial learning, may force learners to encode all learning events, list learning allows them to neglect parts of the list that are deemed irrelevant. However, Experiment 2 used a trial-by-trial learning format and the results closely matched the results of the experiments using the list format. Nevertheless, there are other studies using trial-by-trial learning with a random order of trials in which conditionalization could be reliably found (e.g. Spellman, 1996a,b). One reason for the divergence could be that Spellman used more salient cover stories that emphasized the causal relevance of all the causal factors. In her experiments participants observed pictorially represented trials in which different colored liquids were simultaneously poured onto a plant

which made the causal impact of both liquids visually salient. Furthermore, in the initial instruction it was made clear that these liquids were either fertilizers or weed killers. Thus, it was obvious to participants that both liquids were causally relevant, and that the two causes may even cancel each other out. By contrast, the cover stories in our Experiment 2 only mentioned the possibility that the drug may have different effects on males and females. Furthermore, there was no reason to expect reversals of effects (as with fertilizers and weed killers).

Similarly with causal chains there may be learning conditions which make it easier to become aware of the fact that the Markov condition is violated. In a recent study, Young, Johnson, and Wasserman (2000) have used an occasion setting paradigm with trials in which an effect was temporally preceded by two causal factors. The results showed that participants had a strong tendency to conditionalize the causes on each other when a temporal gap separated them (as in a causal chain). Thus, trial-by-trial learning of chain structures may increase the tendency of learners to notice interactions between the elements of the chain.

7.2. The role of processing effort

If correct, prior knowledge about causal roles reduces complexity *and* leads to correct causal representations. In contrast, the beneficial role of saving processing effort by resorting to unconditional contingencies is less clear. Causal structures following Simpson's paradox are a clear example of how this strategy may lead to erroneous judgments. However, interactions among causal factors that lead to reversals of conditional and unconditional contingencies may not be as frequent as the present examples suggest. In the majority of real-world situations, conditional and unconditional contingencies may at least point in the same direction. We hardly would be able to arrive at causal representations if non-linear interactions and reversals of the Simpson's paradox variety were the rule (Dawes, 1988). Sometimes it may even be the goal of the learner to obtain information about a causal event across a number of varying contexts, and in such situations it may be adequate to base the judgment on the overall probabilistic relation observed in a sample of representative contexts (see Shafer, 1996). A politician, for example, who is interested in the general impact of a new health-related regulation may just be interested in the general average effect of the new policy, and not in the underlying causal relations. Of course, when the goal is to assess specific causal relations independent of other causal factors, contingency estimates that take into account potential causal co-factors are less biased. And indeed the experiments show that the learners attempt to hold such causal factors constant in such tasks whenever it is possible. When the task is too complex, however, the learners will have to trade off effort against the likelihood of deviating substantially from the less distorted estimates.

An interesting question for future research is whether learners are able to adaptively regulate the effort they invest in analyzing statistical relations. Research on decision making has shown that the complexity of information processing is often adaptively varied depending on assessments of the costs and benefits of making correct judgments (Payne, Bettman, & Johnson, 1993). It seems plausible to expect

that processing effort and the plausibility of the causal model interact. Learners may have a tendency to simplify the task when there are no convincing reasons to invest more effort, whereas they may use more effortful strategies with cover stories that suggest the relevance of these strategies.

In summary, the results of the present experiments along with related research suggest that causal induction is determined by a complex interaction between top-down assumptions, goals of the learner, structure of the learning input, type of learning task, and assessments of trade-offs between strategies and learning effort. The present studies only present a first step in elucidating these complex interactions.

Acknowledgements

We would like to thank Merideth Gattis for helpful comments. Experiment 1 was presented at the 1995 meeting of the Cognitive Science Society, Pittsburgh, PA. The first two experiments were planned and conducted while the authors were affiliated with the Max Planck Institute for Psychological Research, Munich, and the University of Tübingen.

References

Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*, 299–352.

Ahn, W., & Mooney, R. J. (1995). Biases in refinement of existing knowledge. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 437–442). Hillsdale, NJ: Erlbaum.

Appleton, D. R., French, J. M., & Vanderpump, M. P. (1996). Ignoring a covariate: an example of Simpson's paradox. *American Statistician*, *50*, 340–341.

Baker, A. G., Mercier, P., Vallée-Tourangeau, F., Frank, R., & Pan, M. (1993). Selective associations and causality judgments: presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 414–432.

Bickel, B. J., Hammel, E. A., & O'Connell, J. W. (1977). Sex bias in graduate admissions: data from Berkeley. In W. B. Fairley & F. Mosteller (Eds.), *Statistics and public policy* (pp. 113–130). Reading, MA: Addison-Wesley.

Blyth, C. R. (1972). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association*, *67*, 364–366.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.

Cartwright, N. (1983). *How the laws of physics lie. Essay 1*. Oxford: Clarendon Press.

Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.

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

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, *72*, 193–204.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid diagnostic signs. *Journal of Abnormal Psychology*, *74*, 271–280.

Cheng, P. W. (1993). Separating causal laws from casual facts: pressing the limits of statistical relevance.

In D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 30, pp. 215–264). San Diego, CA: Academic Press.

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

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

Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.

Dawes, R. M. (1988). *Rational choice in an uncertain world*. San Diego, CA: Harcourt Brace Jovanovich.

Dupré, J. (1993). *The disorder of things. Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.

Eells, E. (1991). *Probabilistic causality*. Cambridge: Cambridge University Press.

Eells, E., & Sober, E. (1983). Probabilistic causality and the question of transitivity. *Philosophy of Science*, *50*, 35–57.

Fiedler, K., Walther, E., Freytag, P., & Stryczek, E. (in press). Playing mating games in foreign cultures: a conceptual framework and an experimental paradigm for trivariate statistical inference. *Journal of Experimental Social Psychology*.

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

Glymour, C. N. & Cooper, G. F. (1999). *Computation, causation, and discovery*. Cambridge, MA: MIT Press.

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs*, *79* (whole volume X).

Koslowski, B. (1996). *Theory and evidence. The development of scientific reasoning*. Cambridge, MA: MIT Press.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Mateo, CA: Morgan Kaufmann.

Pearl, J. (1996). Structural and probabilistic causality. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation. Causal learning* (Vol. 34, pp. 393–435). San Diego, CA: Academic Press.

Pearl, J. (2000). *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press.

Peng, Y., & Reggia, J. A. (1990). *Abductive inference models for diagnostic problem-solving*. New York: Springer-Verlag.

Price, P. C., & Yates, J. F. (1993). Judgmental overshadowing: further evidence of cue interaction in contingency judgment. *Memory & Cognition*, *21*, 561–572.

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

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

Salmon, W. C. (1971). *Statistical explanation and statistical relevance*. Pittsburgh, PA: University of Pittsburgh Press.

Salmon, W. C. (1980). Probabilistic causality. *Pacific Philosophical Quarterly*, *61*, 50–74.

Schaller, M., Asp, C. H., Rosell, M. C., & Heim, S. J. (1996). Training in statistical reasoning inhibits the formation of erroneous group stereotypes. *Personality and Social Psychology Bulletin*, *22*, 829–844.

Schaller, M., & O'Brian, M. (1992). "Intuitive analysis of covariance" and group stereotype formation. *Personality and Social Psychology Bulletin*, *18*, 776–785.

Shafer, G. (1996). *The art of causal conjecture*. Cambridge, MA: MIT Press.

Shafir, E. (1993). Intuitions about rationality and cognition. In K. I. Manktelow & D. E. Over (Eds.), *Rationality: psychological and philosophical perspectives* (pp. 260–283). London: Routledge.

Shanks, D. R. (1993). Human instrumental learning: a critical review of data and theory. *British Journal of Psychology*, *84*, 319–354.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.),

The psychology of learning and motivation. Advances in research and theory (Vol. 21, pp. 229–261). New York: Academic Press.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, *47*.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.

Simon, H. A. (1982). *Models of bounded rationality*. Cambridge, MA: MIT Press.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B (Methodological)*, *13*, 238–241.

Spellman, B. A. (1996a). Acting as intuitive scientists: contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, *7*, 337–342.

Spellman, B. A. (1996b). Conditionalizing causality. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation. Causal learning* (Vol. 34, pp. 167–206). San Diego, CA: Academic Press.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, *57*, 31–95.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. New York: Springer-Verlag.

Spohn, W. (1990). Direct and indirect causes. *Topoi*, *9*, 125–145.

Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North Holland.

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

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 53–76.

Waldmann, M. R. (in press). Predictive versus diagnostic causal learning: evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review*.

Waldmann, M. R., & Hagmayer, Y. (1995). Causal paradox: when a cause simultaneously produces and prevents an effect. In J. D. Moore, & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 425–430). Mahwah, NJ: Erlbaum.

Waldmann, M. R., & Holyoak, K. J. (1990). Can causal induction be reduced to associative learning? In M. Piattelli-Palmarini (Ed.), *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society* (pp. 190–197). Hillsdale, NJ: Erlbaum.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: asymmetries in cue competition. *Journal of Experimental Psychology: General*, *121*, 222–236.

Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General*, *124*, 181–206.

Waldmann, M. R., & Martignon, L. (1998). A Bayesian network model of causal learning. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1102–1107). Mahwah, NJ: Lawrence Erlbaum Associates.

Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perception of causal relations in humans: factors affecting judgments of response-outcome contingencies under free-operant procedures. *Learning and Motivation*, *14*, 406–432.

Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 174–188.

Young, M. E., Johnson, J. L., & Wasserman, E. A. (2000). Serial causation: occasion setting in a causal induction task. *Memory & Cognition*, *28*, 1213–1230.