Understanding the Causal Logic of Confounds

Björn Meder (bmeder@uni-goettingen.de)
York Hagmayer (york.hagmayer@bio.uni-goettingen.de)
Michael R. Waldmann (michael.waldmann@bio.uni-goettingen.de)

Department of Psychology, University of Göttingen, Gosslerstr. 14,
37073 Göttingen, Germany

Abstract

The detection of causal relations is often complicated by confounding variables. Handbooks on methodology therefore suggest experimental manipulations of the independent variable combined with randomization as the principal method of dealing with this problem. Recently, progress has been made within the literature on causal Bayes nets on the proper analysis of confounds with non-experimental data (Pearl, 2000). The present paper summarizes the causal analysis of two basic types of confounding: common-cause and causal-chain confounding. Two experiments are reported showing that participants understand the causal logic of these two types of confounding.

Introduction

Scientific studies and everyday causal learning aim to reveal the structure and strength of causal relations among events: Does event *C* cause event *E*? Will a manipulation of *C* generate *E*? In order to answer these questions, data have to be gathered. But even with data it is often hard to answer these questions because the statistical relation observed between *C* and *E* not only may reflect a direct causal relation but a spurious relation due to other, confounding variables.

For example, in the 1950's, a series of studies with non-experimental data was published showing that lung cancer was found to be more frequent in smokers than in non smokers (e.g., Doll & Hill, 1956). This data was interpreted as evidence that smoking is a cause of lung cancer. However, some prominent statisticians (e.g., Fisher, 1958) argued that such a conclusion was not justified on the basis of the available data. Fisher (1958) offered an alternative causal model in which the observed covariation was not interpreted as a direct causal relation but as a spurious correlation generated by a common cause, a specific genotype causing both a craving for nicotine and the development of lung cancer.

Confounding variables are statistically related to both the potential cause *C* (independent variable) and the presumed effect *E* (dependent variable). It is the relation between the confounding variable and the cause that creates serious problems. In the most extreme case the cause and the other variable are perfectly confounded, that is, they are either both present or both absent all the time. In this case it is impossible to tell whether the effect is generated by the cause or by the confounding variable. Note that the problem of confounding does not originate in the relation between the extraneous variable and the effect. Even if the extraneous variable has a very strong influence, the impact of the cause variable can be detected as long as the extraneous variable is not permanently present and the cause variable and the extraneous variable are independent of each other. Under these circumstances the impact of the cause variable can be seen as an increase (generative influence) or decrease (inhibitory influence) of the probability of the effect given the presence of the cause.

Methodology textbooks (e.g., Keppel & Wickens, 2000) strongly recommend controlled experiments to eliminate the relations between the cause and potentially confounding variables. Experiments involve the random assignment of participants to experimental and control groups (i.e., randomization) and a manipulation of the putative cause variable by an outside intervention. This procedure ensures independence of the cause variable from all other potentially confounding variables. However, in some sciences (e.g., astronomy) and in many everyday contexts controlled experimentation is impossible. Thus, people have to deal with the problem of confounding variables quite often. This paper intends to show (i) under which conditions valid causal inferences are possible on the basis of observations even in the presence of confounding variables, and (ii) that people are capable of reasoning correctly with causal models that contain confounds.

Throughout this paper we focus on the most basic type of causal induction, the detection and evaluation of a single causal relation. In addition, we assume that there is a known confounding variable which is related to both the cause and the effect. First, we will provide a causal analysis of confounding. Then we will show how causal Bayes net theory models confounding and causal inferences. Finally, we will report two experiments investigating whether participants are able to take confoundings into account.

The Causal Basis of Confounding

Two basic causal structures may underlie confounding. One possibility is that the confounding variable *X* is a cause of both the candidate cause *C* and the effect *E* (*common-cause confound*, Fig. 1a). Another possibility is a causal-chain model in which the cause variable *C* not only directly influences the effect *E* but also generates the confounding variable *X*, which, in turn, influences *E* (*causal-chain confound*, Fig. 1b). The crucial point is that both models imply a correlation between cause *C* and effect *E*, even when there is no direct causal relation between them (i.e., without the causal arrow $C \rightarrow E$). If the confounding variable is present, both the cause and the effect should tend to be present; if *X* is absent both *C* and *E* should tend to be absent. In addition to the causal relations connecting the confounding variable to *C* and *E* there is a direct causal relation between *C* and *E* whose existence and strength has to be identified.

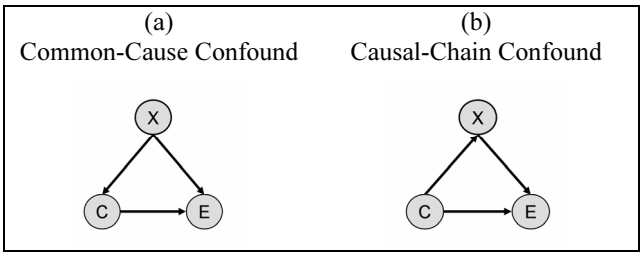


Figure 1: Two Types of Confounding

The two models shown in Fig. 1 represent two different kinds of confounding. The common-cause confound model represents the situation that some extraneous variable is causally affecting both the cause and the effect. The hypothesis that smoking and lung cancer are both caused by a specific genotype exemplifies this type of confounding. There are several possibilities to eliminate the causal relation between the common cause *X* and the candidate cause *C*. For example, *X* might be eliminated or held constant (e.g., only people without the carcinogenic genotype are studied). In addition, *C* might be manipulated independently of *X* (which would not be possible in the case of smoking for ethical reasons). Such an independent manipulation is equivalent to a randomized experiment (Fisher, 1951).

However, simple randomization combined with manipulations of the candidate cause *C* cannot eliminate causal-chain confounding. This type of confounding calls for other controls because a manipulation of *C* would directly affect *X*. Thus, other ways have to be found to block the causal relation connecting the cause *C* to the confound *X*. For example, aspirin (*C*) might not only have a direct influence on headache but also make your blood thinner (*X*), which, in turn, might also have an impact on your headache (*E*). One way to get rid of confounding in this case is to administer aspirin to people who all have thin blood or who are resistant against the side effect, which is equivalent to holding the confound constant. Another possibility is to manipulate the confounding variable in addition to the cause variable and thereby eliminate their causal relation.

In summary, there are two fundamental types of confounding which call for different measures of control. While external manipulations of the candidate cause eliminate common-cause confounding, this is not true for causal-chain confounding. However, controlled experiments are not the only way to avoid confounding. Observational studies in combination with other control techniques (e.g., holding constant hypothesized confounds) may also allow us to draw valid causal inferences. Causal Bayes net theories provide a general framework to model confounds.

Causal Bayes Net Theories

Causal Bayes net theories (Pearl, 2000; Spirtes, Glymour, & Scheines, 1993) allow us to model causal structures with confounding variables. Provided the confounding variable is observable, causal Bayes nets also enable valid inferences about the existence and strength of confounded causal relations. Moreover, they allow us to derive predictions for the consequences of interventions in causal models.

Causal Bayes nets theory integrates graph theory and probability calculus. A causal Bayes net consists of a structural causal model representing events and their directed causal relations (see Fig. 1). Associated with the model are parameters (e.g., conditional probabilities) encoding the strength of the causal relations and the events' base rates. At the heart of the causal Bayes nets framework lies the *causal Markov condition* (Spirtes et al., 1993; Pearl, 2000) which states that the value of any variable *X* in a causal model is independent of all other variables (except for its causal descendants) conditional on the set of its direct causes. By applying the Markov condition the joint probability distribution of a causal model can be factorized using components representing only direct causal relations. For example, the joint probability distribution of the common-cause confound model can be factorized into

$$(1) \quad P(X.C.E) = P(X) \cdot P(C|X) \cdot P(E|C.X)$$

Similarly, the causal-chain confound model can be formalized by

$$(2) \quad P(X.C.E) = P(C) \cdot P(X|C) \cdot P(E|C.X)$$

The conditional probabilities of the decomposed models can be directly estimated from the conditional frequencies in the available data, provided the confounding variable is observed along with the cause and effect variables.

However, the causal consequences of possible interventions cannot always be read off from conditional frequency information alone. Consider the common-cause confound model depicted in Figure 1a. The conditional probability of the effect in the presence of the cause (i.e., $P(e|c)$) reflects both the direct causal influence of *C* on *E* and the spurious relation arising from the confounding common cause *X*. However, *intervening* in *C* renders the cause independent of the confounding variable because the intervention fixes the variable's state. Therefore the probability of *E* given an intervention in *C* reflects the causal influence of *C* on *E* and the causal influence of *X* on *E* but is not distorted by a spurious relation.

To formalize the notion of an external intervention, Pearl (2000) introduced the so-called '*Do-Operator'*, written as Do (•). Formally, the Do-operator renders a variable independent of all its causes, which is graphically represented by deleting all causal links pointing towards the variable fixed by the intervention ('graph surgery'). Based on the modified causal model and the factorized joint probability distribution, the probabilities of the other events can be computed. For example, in the common-cause confound model the probability of *E* given an intervention in *C* is formalized by

$$(3) \quad P(e|\text{Do } c) = P(x) \cdot P(e|x.c) + P(\neg x) \cdot P(e|\neg x.c).$$

Contrary to the original factorization (cf. equation (1)), in this formula *C* is no longer conditionalized on *X*. This formalizes the idea that intervening in *C* eliminates the dependence on its cause *X*. Based on the assumption that the confounding variable *X* and the cause variable *C* independently influence the effect variable (*modularity assumption*) the direct causal impact of *C* on *E* can be estimated. We can only provide a rough description of the inference's logic (for more details and formal derivations, see Pearl, 2000). If *C* is generated by an intervention, the probability of the effect is a summary effect of the causal impact of the cause variable

C and the base rate and causal influence of the confounding variable X. If C is prevented by an intervention the probability of the effect depends solely on the base rate and the causal impact of the confounding variable. According to the modularity assumption the causal influence of the cause and the confounding variable are independent of each other. Therefore, the difference of the two interventional probabilities corresponds to the direct causal influence of cause C.

The predictions for the causal-chain confound model differ. Again the conditional probabilities of the effect given an observation of the cause variable can be directly estimated. What about an intervention on C? As C is the cause of both E and X the dependence of X and C is not eliminated by an intervention on C. Therefore, the interventional and observational probabilities are equal:

$P(e|\text{Do } c) = P(e|c)$ and $P(e|\text{Do } \neg c) = P(e|\neg c)$

Unfortunately, this implies that the direct causal influence of the cause variable on the effect cannot be estimated on the basis of the two conditional interventional probabilities. Given a causal-chain confounding model the difference of the interventional probabilities represents the sum of the cause variable's direct causal impact on E and its indirect causal influence on E via X. In order to assess only the cause's direct causal influence the causal relation between the cause and the confounding variable has to be eliminated by a second intervention. This aim could be achieved by eliminating the confounding variable or by blocking the causal pathway connecting the cause and confound. If this is done, the difference of the resulting interventional probabilities again reflects the cause's direct causal impact upon the effect.

In sum, causal Bayes nets allow us to model various inferences within causal models that contain confounding variables. Necessary prerequisites for these inferences are assumptions about the underlying causal structure and observational data from which the model's parameters can be estimated. No experiments are required.

Causal Learning with Confounds

A number of studies have investigated whether people distinguish between observations and interventions. For example, Gopnik and colleagues (e.g., Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004) showed that even preschoolers grasp the difference between observing and intervening, and learn better when they are allowed to intervene in a causal system. Sloman and Lagnado (2005) provided evidence that participants understand that events targeted by interventions are rendered independent of their actual causes ('undoing'). Thus, people seem to understand the causal logic of intervention (i.e., graph surgery). In our own work (Meder, Hagmayer, & Waldmann, 2005; Waldmann & Hagmayer, 2005) we were able to show that participants base their inferences about hypothetical interventions not only on the structure of the causal model but also take into account the model's parameters which could be estimated from passive observations. Some of the studies also provided first evidence that participants are able to take into account alternative causal pathways in complex models. However, other findings cast doubt on the possibility to generalize this finding. For example, Luhmann (2005) found that participants tend to overestimate how informative confounded data are.

The two experiments reported in this paper focus more closely on different types of confoundings to further explore lay people's understanding. Whereas previous research has focused on whether learners consider the influence of alternative causes (e.g., Spellman, 1996; Waldmann & Hagmayer, 2001) or known alternative causal pathways (Meder et al., 2005), the present experiments go one step further by combining the task of model identification with different kinds of causal inferences. Learners generally are provided with competing causal model hypotheses and then passively observed trial-by-trial learning data. In the test phase participants are asked about hypothetical observations and interventions in causal situations that contain different kinds of confoundings. Thus, our main interest is to investigate whether learners are capable of making correct inferences about hypothetical situations in scenarios that contain confounds.

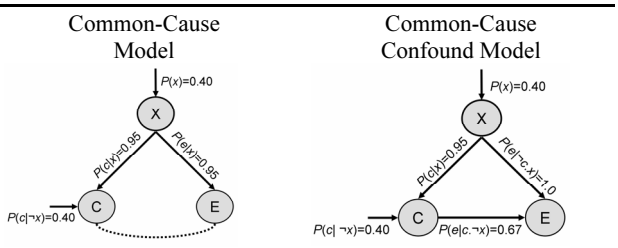
While Experiment 1 confronted participants with a common-cause confounding, Experiment 2 focused on a causal-chain confounding. In both experiments we manipulated the strength of the target causal relation, which was distorted by a superimposed spurious relation. The target causal relation was in one condition present and in the other absent. In order to tap onto participants' understanding we gave them several tasks. First, we asked them explicitly to indicate whether there is a direct causal relation between the cause and effect variable. In order to answer this question correctly, participants have to separate the causal from the spurious relation (in the case of common-cause confounding) or to disentangle the direct from the indirect causal influence (in the case of causal-chain confounding). Second, we asked participants about the probability of the effect given an observation of the absence or presence of the candidate cause. These questions refer to the summary effect of the direct and indirect causal pathways. Third, we asked participants about the probability of the effect given the cause was generated or prevented by an intervention in the cause. In case of a common-cause confounding, the estimated interventional probabilities should differ from the observational probabilities, whereas in the case of the causal-chain confounding simple interventions are not able to eliminate confounding. Thus, in this situation the interventional probabilities should include the confounded causal relation and therefore equal the observational probabilities. To test whether participants are able to extract the direct causal relation in this case, we added questions about combinations of interventions. We asked participants what would happen if the cause was set by an intervention when simultaneously the causal relation to the confounding variable was blocked by a second intervention. If participants understand the causal logic of confounding, the estimated probabilities should reflect the direct impact of the cause upon its effect.

Experiment 1

Participants and Design. Participants were 36 psychology undergraduates at the University of Göttingen. The factor 'learning data' was varied between conditions.

Task. Experiment 1 investigated participants' understanding of common-cause confounding. Participants were told that ornithologists recently had discovered a new species of bird. The biologists hypothesized that in this species singing (*C*) is causally related to reproduction (*E*). It was pointed out that it was difficult to assess this direct causal relation because of a gene (*X*), which is known to influence both the birds' capacity to sing and their fertility. Participants were then suggested two candidate causal models representing either the hypothesis that there is an additional direct causal relation between singing and breeding (common-cause confound model) or the hypothesis that there is none (common-cause model). Learners were also shown a graphical representation of the two causal models and were requested to find out which of the two models was correct. This phase was identical for all participants.

Learning Phase. To assess whether there is a direct causal relation between birdsong and breeding, learners received 50 index cards depicting observational data from individual birds. Two data sets either implemented a common-cause model without a direct causal relation between *C* and *E* or a common-cause confound model. The two parameterized models and the resulting patterns of data are shown in Figure 2. Participants received one of the two data sets and were free to explore the data at will and take notes.



Data pattern			Frequencies	
X	*C*	*E*	Common cause	Confound
yes	yes	yes	18	18
yes	yes	no	1	1
yes	no	yes	1	1
yes	no	no	0	0
no	yes	yes	0	8
no	yes	no	12	4
no	no	yes	0	0
no	no	no	18	18

Figure 2: Parameterized causal models and data sets of 50 cases, each generated from these graphs. Arrows indicate causal relations between variables; conditional probabilities encode the strength of these relations.

Test Phase. After the learning phase participants were given two blocks of questions referring to hypothetical observations and hypothetical interventions. The order of blocks was counterbalanced. Participants were allowed to refer back to the index cards and instructions while answering the questions. The *observational questions* stated that the researchers had captured a new bird and observed that this bird sings [does not sing]. Based on this observation, learners were asked to estimate the probability that this bird would breed (i.e., $P(e|c)$ and $P(e|\neg c)$). The *generative interventional question* stated that the ornithologists had attached a miniature speaker to a bird which imitates birdsong (i.e., Do *c*). The *inhibitory interventional question* stated that the biologists had surgically modified the bird's vocal cords, thereby preventing the bird from singing (i.e., Do ¬*c*). Again, participants had to estimate the probability that these birds would breed (i.e., estimate the interventional probabilities $P(e|\text{Do } c)$ and $P(e|\text{Do } \neg c)$). The correct answers to these questions derived from a causal Bayes net analysis are shown in Table 1 in parentheses. Finally, participants were given a graphical representation of the two alternative causal models and requested to select the correct one.

Table 1: Mean probability judgments for hypothetical observations and hypothetical interventions.

	Observations		Interventions	
	$P(e\|c)$	$P(e\|\neg c)$	$P(e\|\text{Do } c)$	$P(e\|\text{Do } \neg c)$
	63.89	22.78	50.00	41.39
	[58]	*[05]*	*[38]*	*[38]*
	58.33	14.44	63.06	20.56
	[84]	*[05]*	*[78]*	*[40]*

Note. Probabilities derived from Bayes nets are presented in parentheses.

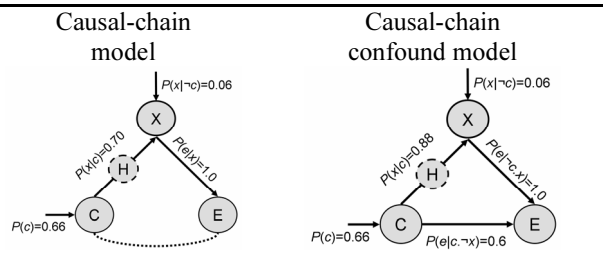
Results. 27 out of 36 participants (75%) chose the correct model. Thus, a majority of participants was able to disentangle a causal relation from a spurious relation. Table 1 shows learners' probability judgments for hypothetical observations and hypothetical interventions. Participants' judgments for observations reflected the statistical relation between the cause and the effect. In both conditions they rated $P(e|c)$ higher than $P(e|\neg c)$. An analysis of variance with 'data sets' and 'presence versus absence of *C*' as variables only yielded a significant main effect for the presence of *C*, $F(1,34)=63.70$, $p<.001$, $MSE=510.38$. In contrast, learners' estimates for the outcomes of interventions differed. As expected, an analysis of variance resulted in a significant interaction, $F(1, 34)=27.95$, $p<.01$, $MSE=420.63$, which we further analyzed by planned comparisons. In the common-cause condition, only a small, non-significant difference was obtained between the interventional probabilities, $F(1,17)=1.09$, $p=.31$, $MSE=614.42$. This result indicates that learners understood that the observed correlation is spurious and that intervening in *C* will not make *E* more or less likely to occur. In the common-cause confound condition the probability of *E* being present was judged higher when *C* was generated, $P(e|\text{Do } c)$, than when it was prevented, $P(e|\text{Do } \neg c)$, $F(1,17)=71.67$, $p<.001$, $MSE=226.84$. Thus, participants had detected the direct causal link connecting the cause to its effect.

Experiment 2

Participants and Design. Participants were 36 psychology undergraduates at the University of Göttingen. The factor 'learning data' was varied between conditions.

Task. While Experiment 1 investigated participants' understanding of a common-cause confounding, Experiment 2 focused on a causal-chain confounding (see Fig. 1). We used the same scenario as in Experiment 1. However, now

participants were told that ornithologists were investigating whether a specific gene (*C*) has a direct causal impact upon the birds' reproduction (*E*). As in Experiment 1, participants were informed about the presence of a confounding variable. They were told that the gene was known to affect the birds' ability to sing (*X*) by a (non-observable) hormone mechanism (*H*) which affects the birds' ability to sing. Moreover, singing (*X*) has, according to the instructions, a causal influence upon reproduction (*E*). Participants were then presented with two competing causal hypotheses, a causal-chain model and a causal-chain confound model (Fig. 3). The causal-chain model represents the hypothesis that the gene affects reproduction only via singing, whereas the causal-chain confound model represents the assumption that the gene has both an immediate and an indirect causal impact upon reproduction.



Data pattern			Frequencies	
X	*C*	*E*	Causal-Chain	Confound
yes	yes	yes	29	23
yes	yes	no	0	0
yes	no	yes	1	1
yes	no	no	0	0
no	yes	yes	0	6
no	yes	no	4	4
no	no	yes	0	0
no	no	no	16	16

Figure 3: Parameterized causal models and data sets of 50 cases generated from these graphs.

Learning Phase. As in the first experiment, learners received 50 index cards displaying observational data from individual birds. The models used to generate the two sets of data and the resulting distributions of event patterns are shown in Figure 3. Note that participants were never informed about the state of *H*, the mechanism connecting *C* to *X*. The causal-chain data indicated that the observable relation between *C* and *E* was merely indirect, while the data corresponding to the causal-chain confound model pointed to a fairly strong direct relation between the gene and reproduction. The unconditional relation between *C* and *E* was identical in both data sets. As before, participants were free to explore the data at will.

Test Phase. In this phase participants were given three blocks of questions with the order of blocks being counterbalanced. The *observational questions* asked participants to estimate the probability that a new bird possessing the gene [not possessing the gene] would breed (i.e., $P(e|c)$ and $P(e|\neg c)$). The *generative interventional question* stated that the researchers had activated the gene of a new bird by means of an intervention (i.e., Do *c*). The *inhibitory interventional question* mentioned that the gene was deactivated by an outside intervention (i.e., Do ¬*c*). Again, participants were requested to estimate the probability that these new birds would breed (i.e., $P(e|\text{Do } c)$ and $P(e|\text{Do } \neg c)$). Then, a question referring to a *combination of interventions* was given which requested participants to assume that researchers had activated the gene of a newly caught bird while simultaneously blocking the generation of the hormone affecting singing (i.e., Do *c*. Do ¬*h*). The second combination question stated that both the gene and the hormone production had been deactivated by inhibitory interventions (i.e., Do ¬*c*. Do ¬*h*). For both questions participants were asked about the probability of procreation (i.e., $P(e|\text{Do } c.\ \text{Do } \neg h)$ and $P(e|\text{Do } \neg c.\ \text{Do } \neg h)$). In both cases, participants received no information about whether the individual birds had the capacity to sing or not. Finally, participants had to select the correct model from a graphical representation of the two alternative causal models.

Results. 31 of the 36 participants (86%) picked the correct causal model. Thus, as in Experiment 1 a majority of participants was able to separate the causal relation between *C* and *E* from the spurious relation. Table 2 shows the mean probability estimates for the six questions along with the values derived from causal Bayes nets. Again participants gave on average the same ratings to the observational questions in both conditions and judged the effect to be more likely in the presence than in the absence of the observed cause, $F(1,34)=317.25$, $p<.01$, $MSE=231.19$. Contrary to Experiment 1 and consistent with the Bayesian causal analysis participants' estimates for the simple interventional questions did not differ between conditions. An analysis of variance resulted in a significant main effect of 'presence versus absence of *C*', $F(1,34)=250.20$, $p<.01$, $MSE=286.42$, but no interaction with the given data set ($F<1$). Participants apparently understood that intervening in *C* would generate *E* no matter whether the underlying causal model was a causal-chain or a causal-chain confound model. Participants' answers to the combination questions showed that they differentiated between the two models. An analysis of variance yielded the expected interaction, $F(1,34)=7.80$, $p<.01$, $MSE=382.55$, but also main effects of the variables 'presence versus absence of *C*', $F(1,34)=54.62$, $p<.01$, and 'data sets', $F(1,34)=9.39$, $p<.01$. A closer look at individual ratings revealed that 10 out of the 18 participants in the causal-chain condition judged *E* to be equally likely when *C* was generated or prevented by an intervention while the causal mechanism linking *C* to *X* was blocked. In contrast,

Table 2: Mean probability judgments for observations, simple interventions, and combinations of interventions

	Observations		Interventions		Combination of Interventions	
	$P(e\|c)$	$P(e\|\neg c)$	$P(e\|\text{Do } c)$	$P(e\|\text{Do } \neg c)$	$P(e\|\text{Do } c.\ \text{Do } \neg h)$	$P(e\|\text{Do } \neg c.\ \text{Do } \neg h)$
	76.89	16.72	77.08	17.28	29.03	7.83
	[88]	*[06]*	*[88]*	*[06]*	*[06]*	*[06]*
	76.11	8.61	81.11	14.72	62.22	15.28
	[88]	*[06]*	*[88]*	*[06]*	*[62]*	*[06]*

Note. Probabilities derived from Bayes nets are presented in parentheses.

all participants in the causal-chain confound condition assumed that an intervention in *C* would increase the probability of *E* despite the blocked link. Thus, a majority of participants seemed to have grasped the causal logic of causal-chain confounding.

Discussion

Causal Bayes nets allow us to analyze non-experimental data that reflect the impact of confounding variables. As long as the confounding variable is observed and not perfectly confounded with the target cause, inferences about the causal impact and the consequences of hypothetical interventions can be derived from observational data. Two basic causal structures containing confounds can be distinguished: A common-cause confound model, in which a cause and an effect are directly and spuriously related, and a causal-chain confound model, in which a cause is directly and indirectly affecting its effect. Manipulating the cause by an external intervention eliminates common-cause confounding but not causal-chain confounding, which requires that the second causal pathway is blocked by other means (Pearl, 2000).

The results of the two experiments show that participants understand the causal logic of these two types of confounding. A majority of participants in both experiments was able to disentangle the direct causal relation from the additional spurious relation. How did people achieve this? Previous research on causal judgments has shown that participants tend to control for extraneous causal variables when estimating the causal impact of a target variable (e.g., Spellman, 1996; Waldmann & Hagmayer, 2001). Confounding variables are such extraneous variables. Controlling for these variables, for example by only considering cases in which they are absent, enables us to derive correct inferences. Observations of participants' behavior during the experiment and their written comments indicate that participants used the strategy to focus on the events in which the confound was absent when assessing whether a direct causal relation was present or not.

How does our research relate to findings showing that people occasionally fail to understand confounding? One critical factor might be the data that is shown to participants. In some studies (e.g., Luhmann, 2005) participants did not receive data that allowed them to focus on the absence of the confounding variable (i.e., holding it constant), which may explain participants' failure. But even when participants receive this information, they may not succeed when the critical cases are rare and highly separated from each other (see Waldmann & Hagmayer, 2001). Common-cause confounding may serve as an example: If the common-cause confound has a high base rate and strongly affects the target cause, there will occur only very few cases in which the target cause occurs in the absence of the confounding variable. In the two experiments reported here we have provided data to participants that contained a relatively large number (>10) of such critical cases. In a pilot study (not reported here) we had presented fewer of these critical observations and participants consequently had failed to arrive at correct conclusions.

In summary, people are able to understand the causal basics of confounding. Whereas previous studies have shown that people sometimes can separate direct from spurious relations (e.g., by controlling for co-factors), our results additionally show that learners have the capacity to reason with causal models containing confounds. Based on trial-by-trial learning data they were surprisingly good at deriving correct predictions for hypothetical observations and hypothetical interventions, and were capable of separating causal from spurious relations in these predictions. This remarkable capacity may fail, however, with more complex models or less salient data.

Acknowledgments

We thank Momme von Sydow for helpful comments and Julia Iwen and Irene Warnecke for collecting the data.

References

Doll, R., & Hill, A. B. (1956). Lung cancer and other causes of death in relation to smoking. A second report on the mortality of British doctors. *British Medical Journal*, *233*, 1071-1076.

Fisher, R. A. (1951). *The design of experiments*. Edinburgh: Oliver and Boyd.

Fisher, R. A. (1958). Smoking, cancer, and statistics. *Centennial Review*, *2*, 151-166.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3-32.

Keppel, G., & Wickens, T. D. (2000). *Design and analysis*. Upper Saddle River, NJ: Pearson.

Luhmann, C. C. (2005). Confounded: Causal inference and the requirement of independence. In B.G. Bara, L. Barsalou & M. Bucciarelli (Eds.) *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 1355-1360). Mahwah, NJ: Erlbaum

Meder, B., Hagmayer, Y., & Waldmann, M. R. (2005). Doing after Seeing. In B.G. Bara, L. Barsalou & M. Bucciarelli (Eds.) *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 1461-1466). Mahwah, NJ: Erlbaum

Pearl, J. (2000). *Causality*. Cambridge: Cambridge University Press.

Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science*, *29*, 5-39.

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

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

Waldmann, M. R., & Hagmayer, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. *Cognition*, *82*, 27-58.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 216-227.