A Transitivity Heuristic of Probabilistic Causal Reasoning

Momme von Sydow[1] (momme.von-sydow@bio.uni-goettingen.de)
Björn Meder[1,2] (meder@mpib-berlin.mpg.de)
York Hagmayer[1] (york.hagmayer@bio.uni-goettingen.de)
[1]Department of Psychology, University of Göttingen, Gosslerstr. 14, 37073 Göttingen, Germany
[2]Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

Abstract

In deterministic causal chains the relations '*A* causes *B*' and '*B* causes *C*' imply that '*A* causes *C*'. However, this is not necessarily the case for probabilistic causal relationships: *A* may probabilistically cause *B*, and *B* may probabilistically cause *C*, but *A* does not probabilistically cause *C*, but rather $\neg C$. The normal transitive inference is only valid when the Markov condition holds, a key feature of the Bayes net formalism. However, it has been objected that the Markov assumption does not need to hold in the real world. In our studies we examined how people reason about causal chains that do not obey the Markov condition. Three experiments involving causal reasoning within causal chains provide evidence that transitive reasoning seems to hold psychologically, even when it is objectively not valid. Whereas related research has shown that learners assume the Markov condition in causal chains in the absence of contradictory data, we here demonstrate the use of this assumption for situations in which participants were directly confronted with evidence contradicting the Markov condition. The results suggest a causal transitivity heuristic resulting from chaining individual causal links into mental causal models that obey the Markov condition.

Keywords: Transitivity; causal models; Markov condition; categorization, causal chain; syllogistic reasoning, heuristics

Deterministic Causal Chains

Deterministic causal relations imply transitivity: If *A* causes *B*, and *B* causes *C*, then *A* causes *C*. For deterministic relations this can be justified on a purely logical basis. If one treats deterministic causal relations as material implications, the syllogism Modus Barbara applies, which is known since Aristotle. Expressed in terms of modern predicate logic it states that $\forall(x)\ (A(x) \rightarrow B(x))\ \&\ (B(x) \rightarrow C(x)) \Rightarrow (A(x) \rightarrow C(x))$. Also according to Mental Model Theory, transitivity is predicted (cf. Goodwin & Johnson-Laird, 2005). Likewise, causal theories of reasoning, using causal strength estimates such as ΔP (Jenkins & Ward, 1965) or causal power (Cheng, 1997) entail transitivity for deterministic relationships. For instance, if $\Delta P_{AB} = P(B \mid A) - P(B \mid \neg A) = 1$ and $\Delta P_{BC} = P(C \mid B) - P(C \mid \neg B) = 1$, it follows that $\Delta P_{AC} = 1$. An analogous case can be made for *causal power* (Cheng, 1997) formalized by $w = \Delta P_{AC} / (1 - P(A \mid \neg C))$.

Probabilistic Causal Chains and the Markov Condition

While the validity of transitive inferences in deterministic chains is undisputed, transitivity is not necessarily entailed when the causal relations are probabilistic. Nevertheless, transitivity may intuitively appear reasonable as well. For example, smoking (*A*) increases the probability of having tar (*B*) in your lungs which, in turn, is causally related to lung cancer (*C*). Thus, the three events constitute a generative causal chain $A \rightarrow B \rightarrow C$ entailing that the probability of lung cancer is higher for smokers than for non-smokers. (i.e., $P(C \mid A) > P(C \mid \neg A)$). Thus, a transitive inference from *A* to *C* seems valid here as well.

The representation of such causal relations in *mental causal models* (Sloman, 2005; Waldmann, 1996; Waldmann, Cheng, Hagmayer, & Blaisdell, 2008) as well as in Bayes nets (Spirtes, Glymour, & Scheines, 1993; Pearl, 2000) implies transitivity in causal chains. A Bayes net consists of nodes representing the domain variables and directed edges ("causal arrows") representing the causal dependencies among the variables. At the heart of the Bayes nets formalism is the *Markov condition*, which states that a variable, conditioned on its direct causes, is independent of all other variables in the causal network except its effects (Hausman & Woodward, 1999; Pearl, 2000; Spirtes, Glymour, & Scheines, 1993). For example, applying the Markov condition to the causal chain $A \rightarrow B \rightarrow C$ entails that *A* and *C* become independent conditional on *B*. This assumption is important since it secures a modular representation of individual causal links, separate manipulability, and is crucial for inferring causal relations from probabilistic dependencies (Hausman & Woodward, 1999; Pearl, 2000).

The Markov condition is also essential for basic inferences across complex causal networks because it allows for chaining the individual links to make quantitative predictions. For example, the conditional probability $P(C \mid A)$ can be derived by a multiplicative combination of all possible paths leading from *A* to *C*:

$$P(C \mid A) = P(B \mid A)\, P(C \mid B) + P(\neg B \mid A)\, P(C \mid \neg B) \quad (1)$$

Thus, the Markov condition allows us to infer the conditional probability $P(C \mid A)$ by combining the causal links constituting the chain without observing this relation directly.

The Markov assumption has been postulated to be a necessary and universal feature of causal relations in the world (Hausman & Woodward, 1999). However, the Markov condition has also been criticized. Particularly Cartwright (2001, 2002) argued that the proof for the necessity of the Markov condition in the deterministic case is valid, but vacuous, and the proof of necessity in the probabilistic case is invalid or at least question begging.

Markov Condition on the Type Level

It is possible that some data may provide evidence for a (probabilistic) generative causal relation between events *A* and *B* on the one hand and events *B* and C on the other hand, but this does not necessarily entail that *A* and *C* are positively related. Consider Figure 1, which illustrates this point. Here $P(B\,|\,A) > P(B\,|\,\neg A)$ as well as $P(C\,|\,B) > P(C\,|\,\neg B)$. Thus, both partial relations indicate a generative causal relationship. Nonetheless, paradoxically, $P(C\,|\,A)$ is lower than $P(C\,|\,\neg A)$. Thus, although the initial event *A* (probabilistically) leads to the intermediate event *B,* and *B* (probabilistically) leads to the final effect *C*, *A* does not lead to *C* but rather to ¬*C*. This paradox would also arise if one infers $P(C\,|\,A)$ using causal strength measures like ΔP or causal power. The single links will both be positive. But if we marginalize over *B* and then derive estimates of causal strength directly for the relation *A* and *C*, one would obtain a negative (preventive) relationship. A similar logic holds for refined measures for the existence of single causal links, like causal support (Griffiths & Tenenbaum, 2005).

Hence, on the level of the categories used here, the Markov assumption is violated for this data set. While a causal chain obeying the Markov condition entails that *A* and *C* become independent when conditionalizing on *B*, in this case *C* does not become independent from *A*.

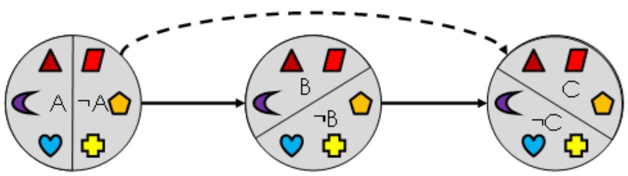

Figure 1: Example for positive causal relations between classes *A* and *B* and between classes *B* and *C*, but a negative relation between classes *A* and *C*.

Note that this paradox only arises because causal relations are usually defined on the level of types, not on the level of individual tokens. The paradox does not arise on the level of individual tokens. For example, in regard of the individual symbolised by a 'triangle' '*A* → *B* → *C'* and '*A* → *C'* hold; similarly for the individual symbolised by a 'heart' '*A* → ¬*B* → ¬*C'* and '*A* → ¬*C'* hold. If causal relations were only concerned with individual tokens the problem would not arise, but then the very concept of a causal law would be meaningless and without any predictive content.

Goals and Hypotheses

In this paper we aim to we examine whether participants' causal reasoning conforms to the Markov condition when making judgments about the relation between *A* and *C*, even when presented with data that violate this assumption.

We predict that participants' judgments will conform to a Markov based integration of single links into mental causal models (Waldmann et al., 2008, cf. in a more general context, Johnson & Krems, 2001), although the presented data violates the Markov condition thereby invalidating transitive inferences from *A* to *C*. To examine peoples' inferences we here focus on the simplest and most uncontroversial measure, conditional probabilities (cf. Oberauer, Weidenfeld, & Fischer, 2007; Evans & Over, 2004).

Related research (Ahn & Dennis, 2000; Baetu & Baker, in press), has indicated that positive causal relations between *A* and *B* and between *B* and *C* lead participants to assume that there is also a positive relation between *A* and *C*. However, in these experiments covariation information about the *A*-*C* relation was not available, thus participants could not directly assess the relation between *A* and *C*. By contrast, in our studies we provide our participants with data indicating non-transitive causal relations. The individuating information about the involved tokens would allow participants to construct chains in which the Markov assumption does not hold. In such chains transitive inferences would not be warranted. Hence, we do not only investigate the assumptions about the relation between *A* and *C* in the absence of information, but we provide information that runs counter to inferences based on combining single links by applying the Markov assumption. We tested this prediction in three experiments in which we successively increased the availability of the data indicating a violation of the Markov assumption (Table 1). Thus, we investigate whether and under which conditions participants use the Markov assumption in causal chains psychologically, even if this assumption is objectively violated.

Table 1: Overview of the experiments

	Intransitive token data	Directly observing data on *A* and *C*	Data available while judging $P(C\,|\,A)$
Experiment 1	✓		
Experiment 2	✓	✓	
Experiment 3	✓	✓	✓

Experiment 1

In Experiment 1 we investigated whether and how an inference regarding the relation between two events *A* and *C* is affected by an intermediate event *B* while keeping everything else equal. In the *control condition* (*A*→*C*) participants were only presented with data regarding the bivariate relation between *A* and *C,* which indicated that the two events were independent (i.e., $P(C\,|\,A) = P(C\,|\,\neg A) = 0.5$). In the *chaincondition* (*A*→*B*→*C*) participants were additionally presented with data regarding an intermediate event *B*.

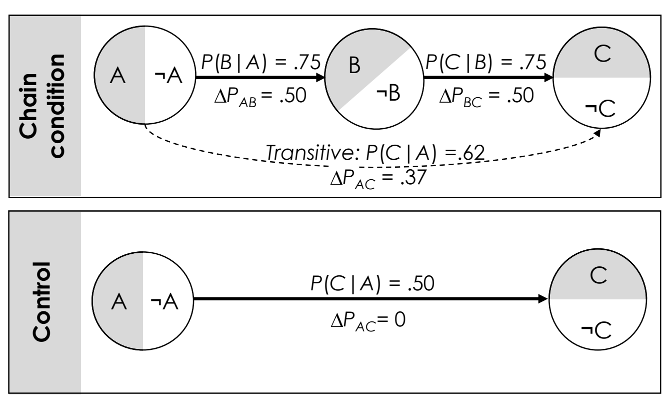

Figure 2: Design of Experiments 1 to 3.

The data (cf. Figure 2) showed positive dependencies between *A* and *B* as well as between *B* and *C*, but the relation between *A* and *C* was identical to the control condition (i.e., $P(C\,|\,A) = P(C\,|\,\neg A) = 0.5$). Using individuating item information and simple one-dimensional rules participants could find out that *A* and *C* occurred actually independent of each other. Thus, our study goes beyond previous research in which the data on the relation of *A* and *C* was not available (Ahn & Dennis, 2000; Baetu & Baker, in press).

However, if people only base their judgments on the integration of individual links, different judgments for the relation between *A* and *C* are expected depending on whether people make a transitive inference from *A* to *C* based on the *A* to *B* and *B* to *C* relations, or estimate the relation between *A* and *C* directly from the data.

Participants Sixty-two students from the Georg-August Universität Göttingen took part in exchange for candy. Additionally they could win a prize. They were randomly assigned to one of the two experimental conditions ($A \rightarrow B \rightarrow C$ vs. $A \rightarrow C$).

Materials Participants were told to imagine being a developmental biologist who investigates the metamorphoses of a species of microbes. The participants were requested to investigate three phases of the microbes' metamorphosis. In each of the three phases they should examine whether the microbes do or do not generate a certain kind of carotene (α-carotene, β-carotene and γ-carotene, respectively). Next participants received information about 40 individual microbes.

The exemplars ("microbes") were represented by circles factorially combining two dimensions, grayscale and size. Figure 3 shows the used item space. The circles varied in eight steps along the dimensions grayscale (white to black) and size (small to big). A pretest showed that subjects were clearly able to distinguish the items. The different event categories *A*, *B*, and *C* were created by rotating the category boundary through the item space, resulting in orthogonal categories *A* and *C* (Fig. 2). To allow the three categories to cut clearly through the item space, we eliminated some items (cf. Fig. 3).

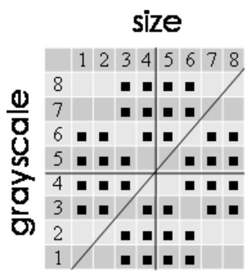

Figure 3: Stimulus material: microbes with three cuts through item space in Phases 1 to 3.

Procedure First participants were presented with data concerning the first and the second developmental stage (Fig. 4). The data for phase 1 arranged the microbes according to whether they did have or did not have produced α-carotene (event *A*). The data was presented in overviews to ease the task. The same microbes in the second phase were arranged according to whether they did or did not generate β-carotene (event *B*). Both panels – each on a large page - were presented simultaneously to participants (Fig. 4). Participants had about a minute time to examine the data sets.

Next, participants were requested to judge whether α-carotene (*A*) rather leads to β-carotene (*B*) or to no-β-carotene (¬*B*). While asking the question, the experimenter pointed to the corresponding classes on the panels. Participants then had to provide an estimate of the conditional

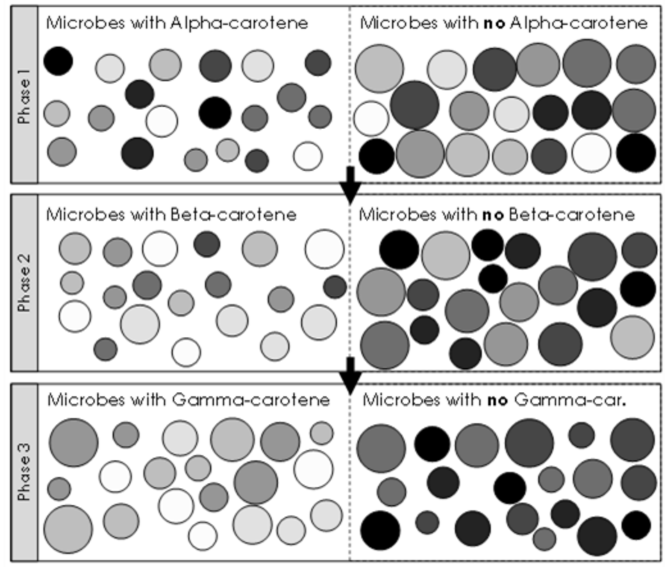

Figure 4: Data panels successively shown to the participants.

probability that given α-carotene (*A*) in the first phase, these microbes produced β-carotene (*B*) (i.e., provide an estimate of $P(B \mid A)$). Because participants may not understand that the absence of a relation (i.e, $P(B \mid A) = P(B \mid \neg A) = 0.5$) implies an equal split of cases, we used a rating scale from -100 to + 100 (cf. Figure 5).

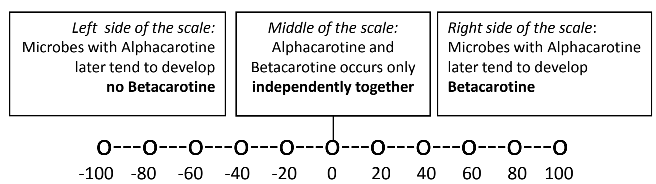

Figure 5: Example of a used rating scale.

Subsequently, the Phase 1 data was removed and the Phase 3 data was added. These data presented information which microbes had produced γ-carotene (*C*) or no γ-carotene (¬*C*) (Fig. 4, cf. Fig. 2 and Fig. 3). Participants then had to estimate the conditional probability $P(C \mid B)$ (i.e. asked to estimate whether microbes which produced β-carotene (*B*) would also produce γ-carotene (*C*)). Finally, all the two remaining data sets were removed and participants were asked to judge whether microbes that produced α-carotene (*A*) produced rather γ-carotene (*C*) or rather no γ-carotene (¬*C*). In the control condition, participants were shown data sets 1 and 3 only and asked to estimate $P(C \mid A)$.

Results and Discussion The obtained and predicted conditional probability judgments are shown in Table 2. The predicted probabilities were linearly transformed to accommodate for the used scale between -100 and + 100 (e.g., $P(C \mid A) = .62$ is +24 on this scale, cf. Fig. 2).

With regard to the evaluation of the single links, participants clearly judged that *A* leads rather to *B* than to ¬*B* and that *B* leads rather to *C* than to ¬*C*. Both estimates differed significantly from zero ($P(A \mid B)$: $t(31) = 5.43$, $p < .001$; $P(B \mid C)$: $t(31) = 8.29$, $p < .001$). The critical judgment on

$P(C \mid A)$ was significantly higher than zero ($t(31) = 3.82$, $p < .01$). A quite different judgment was obtained in the control condition, in which participants even gave slightly negative ratings. Consequently, these estimates differed significantly between conditions, $t(60) = 4.03$, $p < .001$). Both estimates also closely corresponded to the probability estimate derived from a transitive inference from *A* to *C* via *B* (linearly transformed to the scale between -100 and +100).

Table 2: Mean estimates (±SE) in Exp. 1.

		B | *A*	*C* | *B*	*C* | *A*	
				data	transitive
Condition *A*→*B*→*C*	Predicted	+50	+50	±0	+24
	Human	+38 (7.1)	+48 (5.8)	+25 (6.6)	
Condition *A*→*C*	Predicted	-	-	±0	
	Human	-	-	-12 (6.7)	

Note: All judgments were given on a scale -100 to +100.

These findings indicate that although *A* and *C* were not statistically related, participants appeared to reason transitively based on the type level. Although bottom-up information was available showing that *A* and *C* are independent, participants in the chain condition gave positive ratings that clearly differed from those obtained in the control condition. These results are consistent with the idea that participants first induced two separate causal links *A*→*B* and *B*→*C* and then integrated them into a causal chain obeying to the Markov condition. Later they used this causal model to make inferences about the relation between *A* und *C*. The findings corroborate the idea that participants reasoned transitively even if there is bottom-up information available that speaks against a transitive inference.

Experiment 2

In Experiment 1 participants were sequentially provided with data regarding the relation between *A* and *B*, and *B* and *C*, respectively. However, participants never saw the data sets regarding *A* and *C* at the same time. Although it should have been easy to use categories to detect the independence between *A* and *C*, one may object that these transitive inferences were due to the fact that participants could not observe the violation directly. To test this hypothesis, in Experiment 2 all three data sets (cf. Fig. 4) were simultaneously presented to participants. We assumed that despite the simultaneous presentation of the data participants may nevertheless derive the estimates from causal model representations, thereby making transitive inferences.

Participants Sixty-four students volunteered in exchange for candy. They were randomly assigned to either of the two conditions (*A*→*B*→*C* vs. *A*→*C*) (cf. Fig. 2).

Material and Procedure We used almost the same material and procedure as in Experiment 1. To control for potential salience differences in the stimuli we used several category schemes in both conditions (Figure 6). In the chain condition, participants were first presented with data sets *A* and *B* (cf. Fig. 2). Then they had to judge $P(B \mid A)$. We used the same scale as before, ranging from -100 to +100 (Fig. 5). Subsequently, data set *C* was added, this time without removing the data regarding event *A*. Thus, all three data sets were visible at the same time, thereby allowing participants to directly examine the relation between *A* and *C*. Participants had time to inspect the data for about a minute. Then they were asked to estimate $P(C \mid B)$. Finally, all panels were removed and participants were asked to judge $P(C \mid A)$.

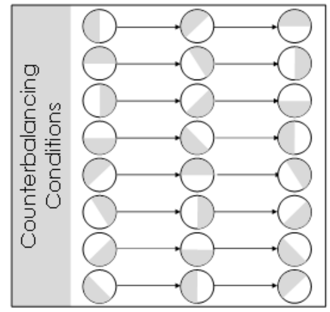

Figure 6: Counterbalancing conditions of Experiment 2 (analogously for the Control Conditions).

Table 3: Mean estimates (±SE) in Exp. 2.

		B | *A*	*C* | *B*	*C* | *A*	
				data	transitive
Condition *A*→*B*→*C*	Predicted	+50	+50	±0	+24
	Human	+50 (5.0)	+39 (6.9)	+26 (7.1)	
Condition *A*→*C*	Predicted	-	-	±0	
	Human	-	-	- 10 (7.6)	

Note: All judgments were given on a scale -100 to +100.

Results and Discussion Table 3 shows that participants again judged the relation between *A* and *B* and *B* and *C*, respectively, as being positive. Again both estimates differed significantly from zero [$P(A \mid B)$: $t(31)= 9.96$, $p < .001$ and $P(B \mid C)$: $t(31) = 5.68$, $p < .001$). The result for $P(C \mid A)$ closely resembled the value entailed by a transitive inference from *A* to *C* via *B*. Thus, even in the presence of data showing that *A* and *C* are not related, participants gave significantly higher estimates in the chain condition than in the control condition ($t(62)= 3.54$, $p < .001$).

Experiment 3

In Experiment 2 participants could inspect all data sets simultaneously, but the data was removed prior to judging $P(C \mid A)$. In Experiment 3 we aimed to investigate whether peoples' inferences adhere to transitivity even when contradictory data is directly available when judging $P(C \mid A)$. Hence, this time participants judged the relation between *A* and *C* in the presence of data indicating a zero contingency.

Participants Sixty-four students from the University of Göttingen participated voluntarily in exchange for candy. Participants were randomly assigned to either the chain condition (*A*→*B*→*C*) or the control condition (*A*→*C*).

Material and Procedure We used the same material and procedure as in Experiment 2. As before, participants were first presented with data on events *A* and *B*. After inspecting the data they were requested to judge the relation $P(B \mid A)$. Then the third data set regarding event *C* was added to the two other sets. Then, while all three data sets were visible, participants were asked to provide an estimate of $P(C \mid A)$. Thus, the data indicating that there is no relation between *A* and *C* was directly available, so that judgments had not to be based on inferences from memory at all. In the control condition participants were only shown the two data sets regarding *A* and *C* and asked to judge $P(C \mid A)$.

Results and Discussion Table 4 shows that participants again judged the relations between *A* and *B* as well as between *B* and *C* positively. Both estimates differed significantly from zero $P(A \mid B)$: $t(31) = 6.89$, $p < .001$; $P(B \mid C)$: $t(31) = 9.93$, $p < .001$). The main result was that the mean rating for $P(C \mid A)$ again closely resembled the value expected on the basis of the Markov assumption and resulting transitive reasoning (Table 3). As before, participants gave significantly higher estimates in the chain condition than in the control condition, $t(62) = 2.84$, $p < .001$.

Table 4: Mean estimates (±SE) in Exp. 3.

		$B \mid A$	$C \mid B$	$C \mid A$ data	$C \mid A$ Transitive
Condition $A \to B \to C$	Predicted	+50	+50	±0	+24
	Human	+44 (6.5)	+47 (4.7)	+17 (5.8)	
Condition $A \to C$	Predicted	-	-	±0	
	Human	-	-	- 9 (6.9)	

Note: All judgments were given on a scale -100 to +100.

Thus, although all participants had the same information available when judging $P(C \mid A)$, the two conditions strongly differed. Participants in the chain condition stated that events *A* and *C* are positively related, whereas participants in the control condition judged them to be independent. Moreover, the absolute mean value in the chain condition (+17) was very close to the value entailed by a transitive inference from *A* to *C* (+24).

General Discussion

Experiment 1 to 3 provide a crescendo of increasingly strict tests of the hypothesis that people reason transitively even if the Markov condition does not hold and, therefore, transitive inferences are not warranted (cf. Table 1). Experiment 1 showed that people made judgments about $P(C \mid A)$ as if transitivity and the Markov condition were given, even when we provided contradictory information. Experiment 2 replicated this finding in a context in which participants had the opportunity to directly assess the relation between *A* and *C*. Nevertheless, the obtained estimates provided strong evidence for transitive inferences based on a causal model constructed by chaining the individual causal links. Finally, Experiment 3 showed that the differences in the judgment regarding $P(C \mid A)$ between the chain and control condition remained stable, even though participants had data about *A* and *C* available when making their judgment. Hence, even in presence of contradictory data participants seem to reason transitively by integrating the two single causal links $A \to B$ and $B \to C$ into one overall mental causal model by using the (objectively violated) Markov assumption. As a consequence they made transitive inferences, not consistent with the actual data. Interestingly, the obtained judgments did not differ much across the three experiments. The deviations from the bottom-up data were only a bit less pronounced in Experiment 3. The average ratings were always close to the predictions derived from the Markov assumption.

These findings go clearly beyond previous studies (Dennis & Ahn, 2000; Baetu & Baker, in press) in which reasoners were never presented with data that contradict transitivity and the Markov condition. While these studies showed, that people seem to believe in the Markov assumption in general, we were able to show that they do so even when contradictory evidence was directly available to them.

Cartwright (2001, 2002) criticized that the Markov assumption has been treated as a universal and necessary aspect of causal reasoning. One may object that the Markov assumption always holds on the level of individual links (i.e., the token level) and that the violation of the Markov assumption in our experiments only occurred on the categorical level. However, categories play an indispensible role in causal prediction and are a necessary prerequisite for causal induction since causal relations can only be noticed on the basis of events that are categorized (Waldmann & Hagmayer, 2006). Standard Bayes net account assume that events (or causally effective objects) are classified into categories, and that causal laws are defined on the categorical level (cf. Kemp & Tenenbaum, 2009). We here worked with categories that were given, but did not adhere to the Markov condition. As Cartwright correctly pointed out, it is perfectly possible that the Markov condition does not hold in the world. Our results show that participants seem to employ the Markov condition subjectively, even with categories that violate Markov objectively. .

Although this is a new finding in the causal domain and in relation to the discussion of transitivity and the Markov condition, there are related findings pointing into a similar direction. For example, *Simpson's paradox*, describes the fact that statistical dependencies can vanish, or even be reversed, when moving from populations to subpopulations. In a number of studies Waldmann and Hagmayer (2001; see also Fiedler, Walther, Freytag, & Nickel, 2003) demonstrated that participants have problems to adequately control for a confounding third variable which reverses the relation between two events. While in the experiments presented here participants aggregated the individual causal links and thereby violated the overall non-existent relation, participants in these former experiments integrated subpopulations violating the relation among the variables in the overall population. In a related context, so-called *pseudo-contingencies* have been discussed (cf. Meiser & Hewstone, 2004). Pseudo-contingencies, however, normally result from highly skewed distributions, which was not the case in the present studies in which all three events *A*, *B*, and *C* had equal probabilities. In our studies the paradox arose not because of a skewed distribution of events but rather are due to the used rotation of the used categories over tokens, that is, categories with overlapping boundaries within a causal chain.

In conclusion, we have provided evidence consistent with the idea that people derive probability estimates by combining single causal links into complex causal models in a modular way (Waldmann, Cheng, Hagmayer, & Blaisdell, 2008). This finding also supports the psychological validity of a core assumption of the Bayes net account, the causal Markov assumption. However, for other causal structures this may not hold. Whether participants believe in the Markov assumption for other causal structures is still an open question, although there is some counterevidence (e.g. Rehder & Burnett, 2005, cf. Mayrhofer, Goodman, Waldmann, & Tenenbaum, 2008). Research on common cause structures (property attribution tasks) suggests that properties that are actually probabilistically related need not to be treated independently (cf. v. Sydow, 2009). Future research needs to investigate these questions in more detail for different causal structures.

In regard of causal chains, however, the current results suggest a transitivity assumption for non-transitive data – hence, a kind of transitivity heuristic.

Acknowledgements

The research was supported by a grant 'Bayeslogik' by the *Deutsche Forschungsgemeinschaft* (DFG, Sy 111/1-1 [MvS]). The authors would like to thank Christin Corinth for her assistance in data collection and several anonymous reviewers for their feedback and advice.

References

Ahn, W., & Dennis, M. (2000). Induction of Causal Chain. *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 19–24). Lawrence Erlbaum Associates, NJ: Mahwah.

Baetu, I., & Baker, A. G. (in press). Human judgments of positive and negative causal chains. *Journal of Experimental Psychology: Animal Behavior Processes.*

Cartwright, N. (2001). What is Wrong with Bayes Nets? *The Monist, 84*, 242-264.

Cartwright, N. (2002). Against Modularity, the Causal Markov Condition, and the Link between the Two: Comments on Hausman and Woodward. *British Journal for Philosophy of Science, 53*, 411-453.

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

Evans, St. B. T., & Over, D. E. (2004). *If.* Oxford University Press.

Fiedler, K., Walther, E., Freytag, P., & Nickel, S. (2003) Inductive reasoning and judgment interference: Experiments on Simpson's paradox. *Personality and Social Psychology Bulletin*, *29*, 14-27.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about Relations. *Psychological Review, 112*, 468-493.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51,* 334-384.

Hagmayer, Y., Sloman, S. A., Lagnado, D. A., & Waldmann, M. R. (2007). Causal reasoning through intervention. In A. Gopnik, & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 86-100). Oxford: Oxford University Press.

Hausman, D., & Woodward, J. (1999) Independence, Invariance, and the Causal Markov Condition, *British Journal for the Philosophy of Science, 50*, 521-583.

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied, 79*, 1-17.

Johnson, T. R., & Krems, J. F. (2001). Use of current explanations in multicausal abductive reasoning. *Cognitive Science, 25*, 903-939.

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review. 116*, 20-58.

Mayrhofer, R., Goodman, N. D., Waldmann, M. R., & Tenenbaum, J. B. (2008). Structured correlation from the causal background. In *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp. 303-308). Austin, TX: Cognitive Science Society.

Meiser, T., & Hewstone, M. (2004). Cognitive Processes in Stereotype Formation: The Role of Correct. Contingency Learning for Biased Group Judgments. *Journal of Personality and Social Psychology. 87*, 599–614.

Oberauer, K., Weidenfeld, A., & Fischer, K. (2007). What makes us believe a conditional? The roles of covariation and causality. *Thinking and Reasoning, 13, 340-369.*

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference.* Cambridge, MA: Cambridge University Press.

Rehder, B. & Burnett, R. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology, 50*, 264-314.

Sloman, S. (2005). *Causal Models. How People Think about the World and Its Alternatives*. Cambridge, MA: Oxford University Press.

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

von Sydow, M. (2009). On a General Bayesian Pattern Logic of Frequency-Based Logical Inclusion Fallacies. In *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

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, Vol. 34: Causal learning* (pp. 47-88). San Diego: Academic Press.

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. (2006). Categories and causality: The neglected direction. *Cognitive Psychology, 53*, 27-58.

Waldmann, M. R., Cheng, P. W., Hagmayer, Y., & Blaisdell, A. P. (2008). Causal learning in rats and humans: a minimal rational model. In N. Chater, & M. Oaksford (Eds.), *The probabilistic mind. Prospects for Bayesian Cognitive Science* (pp. 453-484). Oxford: University Press.