A Structure-dependent Causal Diversity Effect in Diagnostic Reasoning

Felix G. Rebitschek (felix.rebitschek@uni-greifswald.de)
University of Greifswald, Department of Psychology

Josef F. Krems (josef.krems@phil.tu-chemnitz.de)
Technische Universität Chemnitz, Department of Psychology

Georg Jahn (georg.jahn@uni-greifswald.de)
University of Greifswald, Department of Psychology

Abstract

The purpose in diagnostic reasoning is to find the cause of observed effects by applying knowledge about the effects and their potential causes. In the causal structure linking causes and effects, effects can share causes or be linked more indirectly. The causal diversity effect reflects the increased support of a cause by a more widespread distribution of effects within the underlying causal structure. We report two experiments, in which participants acquired knowledge about causal structures and then evaluated diverse and proximal effect patterns with regard to their support for inferring a cause. The diversity effect in diagnostic reasoning was stronger if participants had acquired integrated knowledge about causal structures. Moreover, teaching a reduced structure with less nodes open to alternative causation of proximal effects decreased the diversity effect. This confirmed that the causal diversity effect results from considering alternative causation and more generally that diagnostic reasoning draws on causal representations.

Keywords: causal diversity effect; diagnostic reasoning; causal reasoning; causal models

Introduction

In finding the cause of observed effects, a diagnostician consults knowledge about the effects and their relation to potential causes. However, there are varying ideas about how this causal knowledge is represented. Reasoning about causality can proceed without representations of causal concepts (e.g. Goldvarg & Johnson-Laird, 2001; Shanks, 2010). Nevertheless, causal model representations have many proponents (Krynski & Tenenbaum, 2007; Sloman, 2005; Waldmann & Hagmayer, 2013), and are assumed in causal learning (Holyoak & Cheng, 2011) and causal reasoning (Fernbach & Erb, 2013; Sloman & Lagnado, 2005; Waldmann , 2000).

Evidence consistent with causal model representations was provided, for instance, by the causal status effect (Ahn, Kim, Lassaline, & Dennis, 2000). Furthermore, the distribution of attributes in causal networks affects categorization (Rehder & Hastie, 2001), and causal diversity of symptoms was found to influence diagnostic reasoning (Kim & Keil, 2003). The present study examines the causal diversity effect.

Diversity in a causal structure means the widespread distribution of effects. Regarding Figure 1, a single root cause spreads into two intermediates, which in turn spread each into two effects. The effects 1 and 2 form a proximal pair of effects (likewise effects 3 and 4). In contrast, distally located effects, which do not share an intermediate cause form a diverse pair (e.g., effects 2 and 3).

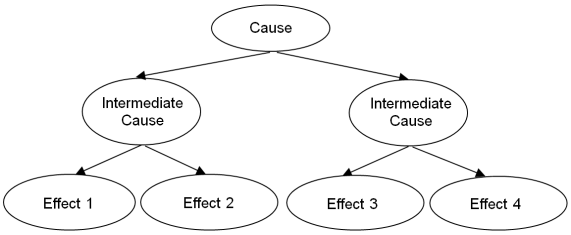

Figure 1: Exemplary causal structure depicted as a graphical causal model; equals the causal structure regarding a single chemical in Experiment 2.

If your task was to evaluate support for the root cause, would you rate the support provided by a diverse pair of effects as higher than the support provided by a proximal pair? You should. Scientists are encouraged to search widely varying support for their hypotheses (Heit, Hahn, & Feeney, 2005) and statistical theory points to the value of diverse evidence (Heit, 1998; Horwich, 1982): According to Bayes' theorem, hardly likely diverse evidence rules out many of the most plausible causes, from which probably one or more would cause proximal effects. In psychological research, effects of diversity of information on thinking and reasoning have been shown in the evaluation of categorical arguments (premise diversity phenomenon in categorical induction, Osherson, Smith, Wilkie, & López, 1990), in the testing of arguments (López, 1995), in the search for diagnostic information (Kim, Yopchick, & Kwaadsteniet, 2008), and in diagnostic reasoning (Kim & Keil, 2003).

Kim and Keil's (2003) participants were taught a causal structure similar to the one shown in Figure 2. The causal chains linking the root cause with the effects were presented separately in order to prevent visuo-spatial encoding of the distance between effects. The subsequent diagnostic reasoning task required to decide for whom of two patients the root cause was more probable. One patient presented a pair of proximal symptoms (e.g. in Figure 2, impaired speech and disability of motion), the other patient presented a pair of diverse symptoms (e.g., disability of motion and stomachache). The patient with the diverse symptoms was

chosen more often (Experiments 1 and 2) and the probability of the root cause was rated higher for patients presenting diverse symptom pairs (Experiment 3).

Only the proximal pair of symptoms can be parsimoniously explained by the shared intermediate cause (paralysis for the symptoms impaired speech and disability of motion). The intermediate cause suffices as a simple explanation (Read & Marcus-Newhall, 1993). In contrast, for a diverse pair, no intermediate cause suffices as the single cause, but the root cause is a simple sufficient explanation.

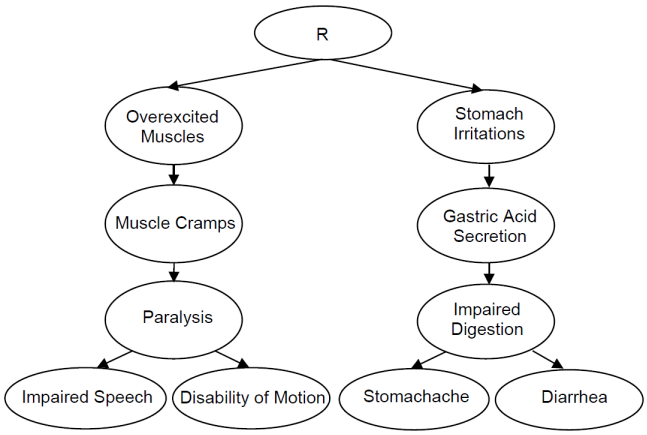

Figure 2: Causal structure equal to the structure used in the study of Kim and Keil (2003); here including an exemplary assignment as used in Experiment 1.

Kim and Keil (2003) listed three possible explanations for the diversity effect that was obtained with causal structures involving three levels of intermediate causes (Figure 2). Diversity effects could have resulted (1) because participants counted the number of supported causal chains, (2) because participants regarded correlated effects as less informative, and (3) because participants considered possibilities of alternative causation. Only the third explanation predicts a reduced diversity effect for shorter causal chains (Figure 1).

According to the first explanation – counting supported chains, the diversity effect is based on a decomposition of the causal structure. The root cause linking two causal chains (Figure 2) may be mentally decomposed into two distinct causes. Their chains are then expected to be separately regarded by the reasoner. Whereas proximal effects share the same causal chain, diverse effects are located in two distinct chains. Thus, if participants had judged the presence of the root cause by the number of causal chains that the observed effects belong to, they would have produced the diversity effect. To our knowledge there is no study confirming this explanation. Because this explanation does not relate to the length of causal chains, there is no change of the diversity effect to be expected from a causal structure with fewer levels.

The second explanation assumes that sets of effects that are expected to regularly occur together due to their shared intermediate cause are assigned a lower information value. Such correlated proximal effects cannot be considered as independently informative with respect to the root cause. But if the co-occurrence of effects is more surprising as it is for diverse effects (Kim, Yopchick, & de Kwaadsteniet, 2008) because the effects do not regularly appear together, stronger inferences are licensed (see also Heit, Hahn, & Feeney, 2005; correlation approach). This explanation focuses on the sharing of an intermediate cause and does not concern the length of causal chains either.

The third explanation concerns imaginable possibilities of alternative causation. Diverse effects are caused by distinct intermediate causes, while proximal effects share a direct parent, a common intermediate cause. The presence of a single intermediate cause can be more easily attributed to causes different from the root cause in question. Thus, an alternative causation for proximal effects is easy to imagine. However, a diverse effect pair requires the presence of two intermediates at once. Hence, there are less alternative hypotheses of causation explaining the diverse effects. The smaller set of contenders for the synchronous causation of diverse effects affords stronger inferences (see also Heit, Hahn, & Feeney, 2005; eliminative approach). This difference between proximal and diverse effects still exists with just one level of intermediate causes. Hence, a diversity effect should be observed even with shorter causal chains, but it should be smaller because with fewer intermediate causes in the chain leading to proximal effects, there are fewer nodes at which alternative causes can be imagined to trigger the chain. Thus, the possibilities of alternative causation differ less between proximal and diverse pairs than in causal structures with longer causal chains.

The subsequently reported experiments explore the effect of the length of causal chains on the diversity effect in diagnostic reasoning with a task very similar to the one used by Kim and Keil (2003) but with different materials. Experiment 1 uses a causal structure with three levels of intermediate causes as in the original experiments. Extending Kim and Keil (2003), in Experiment 2, the levels are reduced to one intermediate. Reducing the number of levels does not change the number of chains and proximal effects do still share an intermediate cause as parent. Just the possibilities of alternative causation should differ less between proximal and diverse pairs if causal chains are shortened. Thus, a reduced diversity effect in Experiment 2 would suggest that participants consider alternative causation.

Experiments

The first experiment aimed to replicate the causal diversity effect with a paradigm from research on diagnostic reasoning (Jahn & Braatz, 2014; Mehlhorn, Taatgen, Lebiere, & Krems 2011; Rebitschek, Scholz, Bocklisch, Krems, & Jahn, 2012). The causal structure from Kim and Keil (2003) was adapted using new quasi-medical material (an exemplary assignment is illustrated in Figure 2).

Participants were instructed that they would learn about a chemical that causes symptoms and they would later evaluate patients' symptoms with regard to this chemical. For acquiring the causal structure without seeing it fully depicted (Kim & Keil, 2003), the structure was presented as four individual linear causal chains (each with one cause, the intermediate causes, and one effect). While participants in Experiment 1 had to acquire a causal structure with three intermediate levels (Figure 2), Experiment 2 presented a structure with one intermediate level (Figure 1). Participants studied, memorized, and reported the causal chains to the experimenter. When all four chains could be reported, the diagnostic task started.

Each presentation slide in the diagnostic task presented two patients with symptom sets; one diverse and one proximal set. Participants had to rate the probabilities that the patients had come into contact with the respective chemical. Participants evaluated symptom sets with regard to one chemical, and subsequently learning and diagnostic judgments were repeated for a second chemical.

Method

Participants Fifty (33 female, mean age 22.9, *SD* = 3.1) undergraduate students from the University of Greifswald participated in Experiment 1, forty-nine (39 female, mean age 22.6, *SD* = 2.8) from the Technische Universität Chemnitz took part in Experiment 2. Two participants were excluded from the analysis of Experiment 1 and one participant from the analysis of Experiment 2 because they reported professional medical expertise, which was an a-priori criterion of exclusion (regarding our quasi-medical scenario).

Design The design included one within-factor reflecting the different symptom sets that were presented in the diagnostic task (diverse and proximal). Furthermore, the number of hierarchical levels in the causal structures was varied between Experiments 1 (three intermediates) and 2 (one intermediate).

Material In Experiment 1, the causal structure for each of the two chemicals that had to be acquired as diagnostic knowledge, consisted of the root cause (the chemical, e.g., *W*), two chains of intermediate causes (e.g., the two middle columns in the bottom half of Table 1: *Dry Eyes, Eye Irritations, Reddened Eyes* and *Skin Tingling, Itching, Scratch Wounds*), and two pairs of effects (the symptoms, e.g., *Epiphora, Impaired Vision* and *Dermatitis, Scarring*); another example with the two columns in the top left of Table 1 is depicted in Figure 2. The used causal structures (Table 1) reflect the outcome of several pretests, regarding plausibility, the instructions, and symptoms' assignments. Prior to Experiment 2, the plausibility of reduced strands including only one intermediate cause was pretested (formal structure in Figure 1) and the underlined intermediate causes in Table 1 were finally selected.

Table 1. Causal chains. Participants learned about two chemicals (e.g., R and K), each with two causal chains spreading into two symptoms. The two causal chains of a chemical were selected from the possibilities listed in the table. In Experiment 1, all listed intermediate causes were learned, in Experiment 2, just the underlined intermediate causes were learned.

	Chemical B / R (1st level)			
Level	Intermediate causes	Intermediate causes	Intermediate Causes	Intermediate causes
2nd	Overexcited Muscles	Stomach Irritations	Throat Irritations	Impaired Lymph No.
3rd	Muscle Cramps	Gastric Acid Secretion	Mucosa Tears	Imm. Cell Deficiency
4th	<u>Paralysis</u>	<u>Impaired Digestion</u>	<u>Sore Throat</u>	<u>Susceptib. to Infection</u>
5th				
Effects	Disability of Motion	Diarrhea	Bleeding Throat	Pneumonia
	Impaired Speech	Stomachache	Mucous Congestion	Fungal Disease
	Chemical W / K (1st level)			
Level	Intermediate Causes	Intermediate causes	Intermediate causes	Intermediate causes
2nd	Bleedings	Dry Eyes	Skin Tingling	Allergic Reaction
3rd	Blood Deficiency	<u>Eye Irritations</u>	<u>Itching</u>	Broncho-constriction
4th	<u>Low Blood Pressure</u>	Reddened Eyes	Scratch Wounds	<u>Asthma Attack</u>
5th				
Effects	Paleness	Epiphora	Dermatitis	Difficult Breathing
	Freezing	Impaired Vision	Scarring	Chest Pain

Note. Originally material was in German

In addition to the symptoms taught as part of causal chains, participants were told that general symptoms (*Tiredness, Thirst*) could occur. They were introduced as unspecific symptoms that can be caused by any chemical.

In each trial of the diagnostic reasoning task, symptom sets for two patients were presented (Table 2 lists exemplary symptom sets). One patient showed proximal symptoms (e.g. *Impaired Speech, Disability of Motion*) and the other showed diverse symptoms (e.g. *Impaired Speech, Stomachache*). Either both patients showed an additional unspecific symptom or neither showed unspecific symptoms.

Procedure Participants worked through the learning phase and diagnostic reasoning trials for one chemical and then again for a second chemical. First, the cover story was introduced, according to which the participant in the role of a doctor has to diagnose patients who are workers from an industrial plant processing chemicals.

Participants should learn how the presented chemical could cause symptoms in people. First, they were informed about unspecific symptoms that can be caused by both chemicals. Then, slides presented the knowledge about the first chemical causing symptoms via intermediate causes. Slide presentation was self-paced. The chemical's letter and two out of four possible chains of intermediate causes and symptoms (Table 1) were assigned in a pseudo-randomized way. Likewise, a letter and two out of the other four chains were assigned to the second chemical. The assignment of chains was counterbalanced.

Each presentation slide showed a single linear chain depicting the cause (e.g. *W*), the intermediates (e.g. *Skin Tingling, Itching, Scratch Wounds*), and one effect (e.g. *Dermatitis*); the terms were vertically arranged and linked by downwards pointing arrows. Four slides contained the entire structure spreading from the root cause (the chemical) down to four effects (the symptoms). A landing slide informed the participants that they could repeat studying the slides by pressing letter "w". Presentation of slides was pseudo-randomized; slides showing strands with the same intermediates were never consecutively presented.

After participants had memorized the chains, they reported the acquired knowledge to the experimenter. Reports were only required to be complete, which means that all terms that were presented on the slides should be mentioned. Any specific order, any causal or temporal links, or links at all were not required. While Kim and Keil (2003, p. 159) had prompted participants to explain "all the information about that medical condition that they had read in the preceding diagrams" in a written single paragraph, we prompted our participants to explain to the experimenter how the chemical causes symptoms. The experimenter invited the participant to study the slides again as long as the report was incomplete and noted the way the knowledge was reported. This learning procedure ensured complete knowledge and allowed to note whether participants had linked the chains into a single causal structure.

Subsequent to the learning phase, participants were informed that the patients to be diagnosed suffered from two to three symptoms. Missing symptoms were explicitly claimed to be absent. The diagnostic task was instructed as prompting to rate the probabilities that patients had come into contact with the given chemical. Pressing the space bar started the presentation of items.

On the right side of each slide presenting symptom sets to be rated, two lists of symptoms corresponding to two patients were shown in a vertical arrangement. The lists consisted of either two symptoms or three symptoms (additionally including an unspecific symptom). One list was a proximal set of symptoms, the other was a diverse set. To the left of each symptom set, it was asked: *How probable is it (on a scale from 0 to 100) that this patient had come into contact with chemical <X>?* with X replaced by the letter of the learned chemical. Additionally, a block of text at the top of the slide instructed participants to consider both patients prior to any judgment and to start rating by pressing the space bar. The key press opened the input field for the first rating about the top patient. Participants rated the probability for the first symptom set with the number keys and switched to rating the next patient by pressing the return key.

Four slides with symptom sets were presented for each chemical. The two possible proximal sets from a chemical's structure were presented two times, each time combined with a diverse set randomly drawn from four diverse sets. Table 2 shows the possible proximal and diverse symptom sets for the structure in Figure 2. On half of the slides, the symptom sets included additional unspecific symptoms (*Tiredness, Thirst*)**.**

Table 2. Exemplary proximal and diverse sets of effects matching with the effects in Figure 2. Each of the proximal sets could be combined with each of the diverse sets for trials of the diagnostic reasoning task.

Proximal Sets		Diverse Sets	
Effect 1 Impaired Speech	Effect 2 Disability of Motion	Effect 1 Impaired Speech Effect 3 Stomachache	
		Effect 1 Impaired Speech Effect 4 Diarrhea	
Effect 3 Stomachache	Effect 4 Diarrhea	Effect 2 Disability of Motion Effect 3 Stomachache	
		Effect 2 Disability of Motion Effect 4 Diarrhea	

Vertical positions of proximal and diverse symptom lists were balanced. The order of symptoms within lists was randomized.

Subsequently, participants repeated the entire procedure with a second chemical. In the end, eight diverse and eight proximal symptom lists were rated per participant. While the first experiment lasted about 40 minutes, the second lasted about 30 minutes, because participants took less time to acquire the reduced causal structure.

Results

In Experiment 1, the probability ratings for diverse sets were higher than those for proximal sets with a mean difference of $M = 6.1$, 95% CI [0.2, 12.0]. The standardized effect size of this diversity effect in ratings was $d = 0.30$. For comparison, notice that Kim and Keil (2003, Experiment 3) had obtained a stronger diversity effect with a (derived) size of Cohen's d of 0.82.

The diversity effect could depend on participants' knowledge representation. It could be stronger if the representation was similar to the actual causal structure of the causal chains linked by the root cause. Hence, we analyzed the diversity effect with regard to participants' reports on the acquired knowledge (top half of Figure 3). While the majority of participants (77%), prior to the first

rating, reported knowledge whose representation obviously equaled the causal structure underlying the task, others' reports lacked any structure or plausible links. Indeed, ratings of participants in the latter group did not indicate any diversity effect (d = -0.01), but those whose report equaled a causal structure showed a clear diversity effect with a mean rating difference of M = 8.0, 95% CI [1.8, 14.1], d = 0.43.

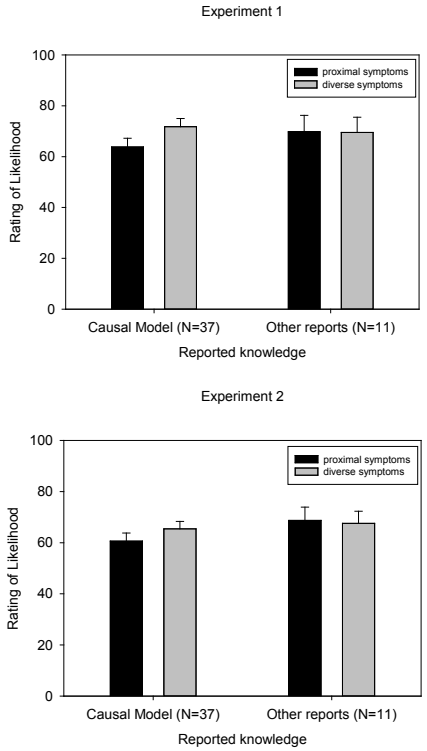

Figure 3: Experiments 1 (above) and 2 (below): Diagnostic ratings contrasting diverse and proximal symptoms, separated according to participants' reports of acquired knowledge.

Experiment 2 presenting a reduced causal structure that included only one intermediate cause did not confirm the diversity effect when ratings were analyzed irrespective of the reported knowledge. The mean difference in ratings was M = 3.4 in favor of diverse symptom sets, 95% CI [-1.7, 8.6]. As shown in the bottom diagram in Figure 3, participants reporting a causal structure after learning (again 77%) produced a small but reliable diversity effect with M = 4.8, 95% CI [-0.8, 10.4], d = 0.29. In contrast, the diversity effect was absent for the remaining participants (d = -0.06). In Experiment 1, 67% of participants showed a diversity effect (higher probability ratings for the diverse symptom set) compared to only 56% in Experiment 2.

Discussion

The described experiments presented sets of symptoms, which were manipulated according to their distribution in a causal structure with a chemical at the root causing these symptoms along one or two causal chains. Diverse symptom sets, that a chemical caused via two different causal chains, and proximal sets sharing one causal chain were contrasted regarding participants' ratings of the probabilities that the chemical caused the respective symptoms. Adapting the causal structure used in the experiments reported in Kim and Keil (2003), we replicated the diversity effect in diagnostic reasoning with new material. Participants' rated diverse symptoms as more probably caused by the chemical than proximal symptoms.

Moreover, both experiments revealed that the diversity effect clearly depends on knowledge representations of cause-effect-relationships that equal the causal structure underlying the construction of symptom sets to be rated. Hence, this precondition of the causal diversity effect leads us to conclude that participants consulted a knowledge representation that reflected some causal structure information.

Important for deciding between explanations of the diversity effect, a smaller diversity effect was shown for causal structures with shorter causal chains in Experiment 2. Although the difference in effect size between experiments is small and comparisons between experiments should be drawn with caution, the diversity effect in Experiment 2 was probably smaller than in Experiment 1.

The reduced diversity effect in Experiment 2 provides some evidence that participants considered possibilities of alternative causation in evaluating symptom sets. Shorter causal chains contain fewer nodes that could be activated by alternative causes that then would bring about the effects instead of the root cause. Thus, there are more possibilities for alternative causation of proximal symptom sets at the end of a longer causal chain. For diverse symptom sets, a synchronous alternative causation in two separate chains is hard to imagine even for longer chains.

The two other explanations of a diversity effect in the present diagnostic task that were discussed by Kim and Keil (2003) do not account for the presumed reduced effect size in Experiment 2. A decomposition of the root cause and its dependents into two separate causal chains that then are checked for confirmation by observed effects should have produced similar effects in both experiments. Likewise, if participants had rated proximal symptom pairs lower because they are linked by a single parent node and do not provide independent evidence, they should have done so in both experiments.

To confirm that the studied diversity effect in diagnostic reasoning varies with the length of causal chains, the length of causal chains should be manipulated as a between-subjects factor in a single experiment. Furthermore, the hypothesis that participants evaluate possibilities of alternative causation could be tested more directly with direct manipulations of possibilities for alternative causation. Such a manipulation could also help to explain why the effect size in Experiment 1 was smaller than in Kim and Keil (2003, Experiment 3) even for participants reporting causal structures.

Notwithstanding possible extensions and methodological improvements, the reported experiments confirm the

diversity effect and show that it requires the representation of a causal structure. In addition, they show that varying a branching hierarchical structure changes diagnostic judgments. This extends previous evidence showing that manipulating a causal model changes causal reasoning (Meder, Hagmayer, & Waldmann, 2009). More generally, our results support causal model theories of causal reasoning under uncertainty (for conditionals, see Fernbach & Rehder, 2013).

The idea that evidential information is mapped onto a qualitative model representation of the acquired causal structure for diagnostic judgments was elaborated by Krynski and Tenenbaum, 2007. Future experiments targeting the diversity effect can examine manipulations of causal strengths or causes' base rates. Such manipulations can test how well causal model theories can predict modulations of the diversity effect and will help to clarify the causal representations underlying diagnostic reasoning.

Acknowledgments

We are grateful to Ramona Groß and Agnes Scholz who helped to manage the experiments and Susann Arnold, Theresa Beuster, Miriam Müller-Bardorff, Maria Stephan, and Barbara Wulfken for their help in conducting pretests and in collecting data. This research was supported by German Research Foundation (DFG) grants JA 1761/7-1 and KR 1057/17-1.

References

Ahn, W. K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology, 41(4),* 361-416.

Fernbach, P. M. & Erb, C. D. (2013). A quantitative causal model theory of conditional reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39 (5),* 1327-1343.

Fernbach, P. M., & Rehder, B. (2013). Cognitive shortcuts in causal inference. *Argument & Computation: Vol. 4, Formal Models of Reasoning in Cognitive Psychology,* 64-88.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naïve causality: a mental model theory of causal meaning and reasoning. *Cognitive Science, 25,* 565-610.

Heit, E. (1998). A Bayesian analysis of some forms of inductive reasoning. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition (pp. 248-274)*. Oxford University Press.

Heit, E., Hahn, U. & Feeney, A. (2005). Defending diversity. In W. Ahn, Goldstone, B.C., Love, A.B. & Wolff, P. (Eds.), *Categorization inside and outside of the lab: Festschrift in Honor of Douglas L. Medin (pp. 87-100)*. Washington, DC: American Psychological Association.

Holyoak, K. H., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology, 62,* 135-163.

Horwich, P. (1982). *Probability and evidence.* Cambridge, England: Cambridge University Press.

Jahn, G., & Braatz, J. (2014). Memory indexing of sequential symptom processing in diagnostic reasoning. *Cognitive Psychology, 68*, 59-97.

Kim, N. S., & Keil, F. C. (2003). From symptoms to causes: Diversity effects in diagnostic reasoning. *Memory & Cognition, 31,* 155-165.

Kim, N. S., Yopchick, J. E., & de Kwaadsteniet, L. (2008). Causal diversity effects in information seeking. *Psychonomic Bulletin & Review, 15,* 81-88.

Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General, 136(3),* 430-450.

López, A. (1995). The diversity principle in the testing of arguments. *Memory & Cognition,* 23, 374-382.

Meder, B., Hagmayer, Y., & Waldmann, M. R. (2009). The role of learning data in causal reasoning about observations and interventions. *Memory & Cognition, 37,* 249-264.

Mehlhorn, K., Taatgen, N. A., Lebiere, C., Krems, J. F. (2011). Memory Activation and the Availability of Explanations in Sequential Diagnostic Reasoning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 37,* 1391-1411.

Osherson, D., Smith, E. E., Wilkie, O., & López, A. (1990). Category-based induction. *Psychological Review, 97(2),* 185-200.

Read, S. J., & Marcus-Newhall, A. R. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology, 65,* 429-447.

Rebitschek, F. G., Scholz, A., Bocklisch, F., Krems, J. F., & Jahn, G. (2012). Order effects in diagnostic reasoning with four candidate hypotheses. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society (pp. 905-910).* Austin, TX: Cognitive Science Society.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General, 130,* 323-360.

Shanks, D. R. (2010). Learning: From Association to Cognition. *Annual Review of Psychology, 61,* 273-301.

Sloman, S. A. (2005). *Causal models: How we think about the world and its alternatives.* New York: Oxford University Press.

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

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., & Hagmayer, Y. (2013). Causal reasoning. In D. Reisberg (Ed.), *Oxford Handbook of Cognitive Psychology.* New York: Oxford University Press, pp. 733-752.