Title
The Activation of Hypotheses during Abductive Reasoning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN
1069-7977

Authors
Baumann, Martin R.k.
Mehlhorn, Katja
Bocklisch, Franziska

Publication Date
2007

Peer reviewed

The Activation of Hypotheses during Abductive Reasoning

Martin R. K. Baumann (martin.baumann@phil.tu-chemnitz.de)
Katja Mehlhorn (katja.mehlhorn@phil.tu-chemnitz.de)
Franziska Bocklisch (franziska.bocklisch@phil.tu-chemnitz.de)
Chemnitz University of Technology, Department of Psychology,
Wilhelm-Raabe-Str. 43, 09107 Chemnitz, Germany

Abstract

Abductive reasoning, that is, finding an explanation for a set of observations, can be understood as a process of sequentially understanding and integrating new observations into a mental model about the current situation (Johnson & Krems, 2001; Josephson & Josephson, 1994). Whereas Johnson and Krems' model focuses on conscious deliberate processes, it has been argued that automatic implicit processes also play an important role in abductive reasoning (e.g. Johnson, Zhang, & Wang, 1997). Adopting Kintsch`s (1998) construction-integration theory, we assume that automatic activation processes regulate the availability of possible explanations during the reasoning process. In our experiment, participants solved an artificial diagnosis task while the activation of explanatory hypotheses was measured. We found that explanatory hypotheses relevant in the current context for explaining a set of observations are kept in a more active state in memory than irrelevant or rejected hypotheses.

Keywords: abductive reasoning; causal reasoning; automatic processes; explanations; activation.

Introduction

Generating a hypothesis to explain one or more observations is an essential part of many real world tasks. This kind of reasoning is called abductive reasoning (Josephson & Josephson, 1994). It is a vital subprocess, for example, in scientific discovery, medical diagnosis, software debugging, social attribution processes, and discourse comprehension. While explaining a given set of observations, the reasoner has often to decide between different alternative explanations to find the best explanation for the observations. We assume that both deliberate reasoning processes and automatic comprehension processes contribute to the generation of hypotheses (Johnson, Zhang, & Wang, 1997; Sloman, 1996). The goal of this paper is to examine how automatic comprehension processes constrain the consideration of hypotheses to the most plausible ones in the given context by making these hypotheses highly available to the reasoner and reducing the availability of implausible ones.

Constructing an explanatory hypothesis can be a quite complex task. First, in many cases there is more than one possible explanation for a given observation. For example, headache is a common symptom of many diseases and is associated with many different causes. Second, the task is often not to explain one observation but a set of observations where each observation can be explained with more than one explanation. In such a case, a combination of elementary hypotheses has to be found that best explains all observations. Following the above example, in most cases a patient complains not only about one symptom, such as a headache, but about a set of observations that could be a headache, sickness, and fever. Each of these symptoms can be caused by many different diseases. The physician's task is to find the best explanation for the whole symptoms set. And, despite the complexity of the problem, the physician solves the problem in most cases quickly and accurately.

How is this accomplished? Johnson and Krems (2001) suggested on the basis of their results on abductive reasoning that people use initial observations to construct a preliminary explanation for these observations. Succeeding observations are sequentially comprehended and integrated to generate a single current explanation for all observations seen so far. If an observation can be comprehended in different ways, that is, if there exist alternative elementary explanations for this new observation, the current explanation is used to decide between these alternatives. Only those elementary explanations for the new observations are considered as relevant that are compatible with the current explanation. Thus, the current explanation acts as an explanatory context for the comprehension and explanation of new observations. It reduces the complexity of the abductive reasoning problem as not all possible elementary explanations for a new observation are considered as relevant but only those that are compatible with the current explanation.

Whereas Johnson and Krems' model focuses on deliberate reasoning processes to describe the use of the current explanatory context, we assume that automatic comprehension processes based on spreading activation and constraint satisfaction also play a key role. It has been argued recently that both deliberate and automatic processes are involved in many reasoning tasks (Sloman, 1996) such as impression formation (Thagard & Kunda, 1998), hypothesis evaluation (Johnson, Zhang, & Wang, 1997), and medical diagnosis (Arocha & Patel, 1995). Thagard and Kunda explain how spreading activation processes can explain the effect of social stereotypes on the interpretation of behavior. Johnson, Zhang, and Wang show how automatic processes can provide information for the evaluation of hypotheses that is used subsequently in more deliberate processes to revise existing or generate new hypotheses in an abductive reasoning task.

In our view these automatic processes also serve the function of making those elementary explanations of new observations highly available to the reasoner that have a high probability of being the relevant explanations in the current context. Effortful and deliberate reasoning processes are thereby focused on those hypotheses that are the most promising candidates. A precondition for this is that the reasoner's knowledge is highly adapted to the task at hand, which normally requires many years of domain experience. If, for example, a physician possesses this experience, observing a given set of observations immediately reminds him of at most a small set of possible explanations for these observations. This set of hypotheses is then used for further examinations (Groen & Patel, 1988).

We assume that these automatic comprehension processes can be described using Kintsch's (1998) construction-integration theory. According to Kintsch, comprehension involves a two-stage process. In the first phase, the construction phase, knowledge stored in long-term memory (LTM) is activated by new information via associative links. The result of this activation process is a knowledge network that is unstructured and in most cases inconsistent. In this network, compatible nodes are connected with excitatory links, incompatible nodes with inhibitory links. The network becomes then integrated by a constraint satisfaction process that keeps compatible knowledge activated whereas incompatible knowledge is suppressed and thereby removed from the activated network. The result of this two-stage process is a coherent representation of the current situation - the situation model that in our case represents the current explanation for the given observations.

Kintsch's theory has been used to explain a wide variety of behavioral phenomena, such as discourse comprehension (Kintsch, 1988), completing the Tower of Hanoi task (Schmalhofer & Tschaitschian, 1993), human computer interaction skills (e.g., Doane, McNamara, Kintsch, Polson, & Clawson, 1992), and action planning in piloting (Doane & Sohn, 2000). Arocha and Patel (1995) adopted the construction-integration theory to explain the di*f*ferences in medical diagnostic reasoning performance between experts and novices. Whereas their focus was to apply Kintsch's construction-integration theory to predict the performance differences between experts and novices in diagnostic reasoning given the different knowledge structures of experts and novices we use Kintsch's theory to make predictions about the activation and inhibition of hypotheses during the abductive reasoning process.

We assume that observations, such as the symptoms of a patient, are linked to explanatory hypotheses in the reasoners LTM (Arocha & Patel, 1995). According to Kintsch (1998), observing a symptom should activate at least the subset of the most common of these explanatory hypotheses. This activated network represents the current explanation after, for example, the first symptom. When observing a new symptom, those hypotheses that are compatible both with the previous and the new observation should receive additional activation. In contrast, hypotheses that were compatible with the previous but incompatible with the new observation should decrease in activation. As these hypotheses cannot explain the whole set of observations they should be rejected. Whereas others assume that previously activated knowledge that has become irrelevant is actively inhibited (Conway & Engle, 1994; Gernsbacher, 1993; Lustig, Hasher, & Tonev (2001) Kintsch (1998) is somewhat unclear about this. The construction-integration theory postulates that incompatible knowledge structures should be associated via inhibitory links leading to the inhibition of incompatible structures. But he also assumes that activated nodes lose activation once they are rejected, but "retain considerable activation as a sort of memory of past trouble" (Kintsch, 1998, p. 169). An additional objective of our research was to address this issue and test the "active inhibition" assumption against the "decay of activation" assumption.

Figure 1 illustrates the postulated construction-integration process. The first symptom (headache) activates associated explanatory hypotheses stored in long-term memory resulting in a preliminary explanation with two alternative explanations for headache - concussion and influenza. The second symptom (fever) is compatible with the influenza explanation for headache but not with the concussion explanation. After the integration phase the influenza explanation as a still relevant explanation should be strengthened as it receives activation both from the symptom fever and the symptom headache, whereas the to-be-rejected concussion explanation should decrease in activation as it is inhibited by the symptom fever. Thus, after integrating the network, influenza remains the only possible explanation for the observed symptoms.

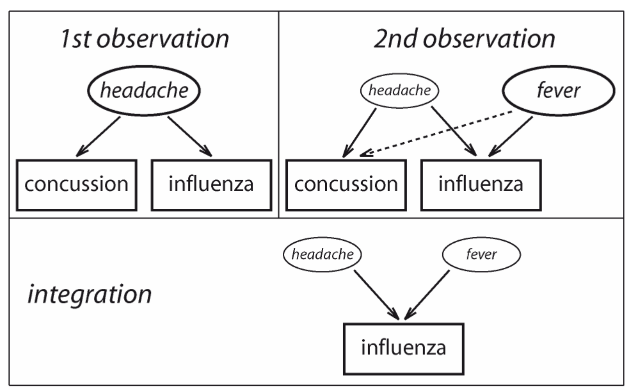

Figure 1: Illustration of the generation of an explanation during an abductive reasoning task according to the construction-integration theory. Solid lines represent excitatory links, dashed lines inhibitory links.

Our experiment tested these predictions. We measured the activation of explanatory hypotheses during an abductive reasoning task. These were either hypotheses that should be considered as *relevant* as they were compatible with all observations presented so far or hypotheses that had to be *rejected* as they were compatible with previous symptoms but incompatible with a new observation or hypotheses that were *irrelevant* as explanation for the presented symptoms because they had never been part of the current explanation. For relevant hypotheses we assumed the activation to be higher than for irrelevant hypotheses because relevant hypotheses should be activated by the associated symptoms, irrelevant hypotheses not as they were not part of the activated network of explanations. For rejected hypotheses, two competing assumptions were tested against each other: If rejected hypotheses are actively inhibited, we would expect the following pattern of activation: relevant > irrelevant > rejected. If activation simply decays, the activation pattern should be: relevant > irrelevant = rejected.

Experiment

The Abductive Reasoning Task

To examine the activation of explanatory hypotheses during an abductive reasoning task an artificial diagnosis task was developed. It was introduced to the participants with the cover story: “You are a doctor in a chemical plant. After a chemical accident an employee complaining about several symptoms comes to see you. Your job is to find out which chemical caused the symptoms.” Table 1 shows the structure of the knowledge necessary to solve the task. The right column displays the symptoms that are caused by the chemicals shown in the middle column. The chemicals were grouped into three categories, “Landin”, “Amid”, and “Fenton”. This hierarchical structure should ease the learning of the material and reflects in a simplified form the hierarchical knowledge organization found in medical diagnosis (Arocha & Patel, 1995). Each chemical caused three or four symptoms[1]. As each symptom could be caused by several chemicals only the combination of symptoms allowed to unambiguously identify the chemical causing the symptoms. In total there were nine different symptoms that were caused either only by chemicals of one group, such as cough in the group Landin, (specific symptoms), or by chemicals of different groups, such as headache (non-specific symptoms).

Table 1: Summary of the material participants had to learn (original material in German).

Group	Chemical	Symptoms
Landin	B	short breath, cough, headache, eye inflammation
	T	short breath, cough, headache, itching
	W	cough, eye inflammation, itching
Amid	Q	redness, skin irritation, headache, eye inflammation
	M	redness, skin irritation, headache, itching
	G	skin irritation, eye inflammation, itching
Fenton	K	vomiting, diarrhea, headache, eye inflammation
	H	vomiting, diarrhea, headache, itching
	P	diarrhea, eye inflammation, itching

[1] The reason for this difference in the number of symptoms was to keep the material of the experiment as close as possible to the material of another experiment (Baumann, Bocklisch, Mehlhorn & Krems, in press). Baumann et al. used some of the trials with three symptom chemicals to present an additional symptom that contradicted the other symptoms.

Figure 2 shows an example of a trial demonstrating the principal idea of this task. In each trial the symptoms of one hypothetical patient were presented sequentially. The participants’ task was to find the correct diagnosis for the set of observations shown in this trial. Therefore they had to sequentially understand the symptoms and integrate them into their current situation model. More precisely, the first symptom should initiate the explanatory context for the following symptoms by activating all explanations compatible with this symptom. The second symptom should either confirm the current set of explanations (if they were compatible with both symptoms) or allow excluding one or more hypotheses from the set of possible explanations (if they were incompatible with the new symptom). By this mechanism the set of relevant explanations should be reduced during the trial until only one possible explanation, the correct diagnosis remained.

To measure the activation of the different types of explanations during the task, a probe reaction task was used. After one of the symptoms in each trial a probe was presented and participants had to decide as fast as possible whether or not the probe was the name of one of the nine chemicals learnt before. Half of the probes, the targets, were names of chemicals, the other half, the distractors, were not. Only the reactions to targets were of interest with regard to our hypotheses. These targets varied in terms of their relation to the current explanation of the symptoms presented so far. The first type of targets, the relevant targets, were the names of relevant explanations, that is, chemicals that were compatible with all symptoms presented so far. The second type, the rejected targets, were the names of rejected explanations, that is chemicals that were compatible at least with the first symptom but incompatible with one of the succeeding symptoms. The third type of targets, the irrelevant ones, were the names of irrelevant explanations, that is chemicals that had never been part of the explanatory context because they were incompatible with the first symptom of the trial.

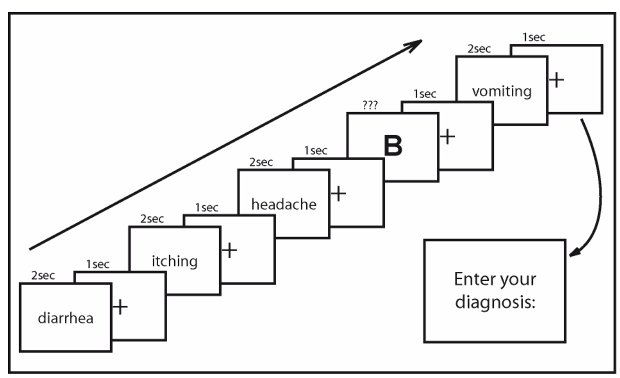

Figure 2: Example trial with irrelevant target after the third symptom and H as final diagnosis.

The reaction times and response accuracy to these targets were used as a measure for the activation of the respective chemicals. The higher the activation of an explanation the shorter should be the reaction time and the higher response accuracy to the corresponding target (Gernsbacher, 1990). According to our predictions about the activation of hypotheses during an abductive reasoning task, we expected the reaction times to the relevant targets to be shorter (and response accuracy to be higher) than the reaction times (and response accuracy) to irrelevant targets. Reaction times to rejected targets should be either slower (given active inhibition) or similar (given decay of activation) to that of irrelevant ones. Respective, response accuracy for rejected probes should be either lower or similar to that of irrelevant ones.

Experimental design

We manipulated in this experiment the type of probe (distractor vs. relevant, irrelevant, or rejected target) and the position at which the probe was presented (after first, second, third, or fourth symptom). Rejected targets were additionally varied according to the number of symptoms between the first incompatible symptom and the target presentation resulting in three different types of rejected targets: “just rejected”, “rejected one (symptom) ago” and “rejected two (symptoms) ago”.

The most relevant dependent variables in terms of our predictions were reaction times and response accuracy to the targets. We also measured diagnosis performance by assessing accuracy and time for generating the diagnosis at the end of each trial.

Participants and procedure

Twenty six (16 female and 10 male; mean age 22.8, SD = 3.6) undergraduate students from the Chemnitz University of Technology took part in this experiment that consisted of 3 sessions within one week. The first session was a pure practice session, to ensure a high familiarity with the material and the task. After learning the material shown in Table 1, participants had to perform a series of practice blocks until they achieved a level of at least 80% correct trials. The actual data collection took place in the following two sessions. It was spilt into two sessions to prevent fatigue among participants. In each of these sessions participants had to solve 170 abductive reasoning trials.

In these 340 trials the sequence of symptoms was chosen such that the different levels of the manipulated factors could be realized, therefore controlling for a) the number of symptoms before presentation of the probe, b) the number of symptoms between the rejection of an explanation and presentation of the respective rejected target, and c) the frequency of the different probe types. Additionally, to control for confounding variables that could possibly affect the performance in the probe reaction task, such as the number of currently possible alternative hypotheses, but could not be included as factor in the experimental design due to efficiency reasons, the frequency of specific and non-specific symptoms on the different positions of a trial sequence and the frequency with which the different chemicals caused the symptoms were balanced across trials.

Results

First, trials were analyzed with respect to accuracy and time for generating the diagnosis at the end of each trial. In 96.1% of all trials participants correctly solved the task. Correct diagnoses were generated on average 731.16 ms (SD = 468.42), wrong diagnoses 1954.30 ms (SD = 1158.18) after onset of the screen asking for the final diagnosis. The high accuracy and short duration for generating correct diagnoses indicates that participants were able to solve the task quite easily. Neither accuracy nor time for generating the final diagnosis differed between the trials with different probe types. This indicates that these trials were comparable in terms of task difficulty.

Second, to test the predictions about the activation of explanations, response accuracy and reaction times for correct responses to the targets were analyzed. Therefore, only trials with correct final diagnosis were used. Because the results for accuracy and reaction time showed a similar pattern, here only the results for the reaction times are presented.

Relevant vs. irrelevant. To test the prediction that relevant explanations are more activated than irrelevant explanations, reaction times to relevant and irrelevant targets were compared. Consistent with the prediction that relevant explanations should be more activated than irrelevant explanations, reactions to relevant targets were faster at all positions in the trial sequence than reactions to irrelevant targets (see Figure 3), $F(1, 25) = 14.537$, $p = .001$, partial eta-square = .368. Furthermore, reaction times to both target types decreased the later in the trial sequence the target was presented, $F(3, 75) = 13.224$, $p < .001$, partial eta-square = .346. This was somewhat more pronounced for the relevant targets than for the irrelevant ones, but the respective interaction term in the ANOVA did not reach significance.

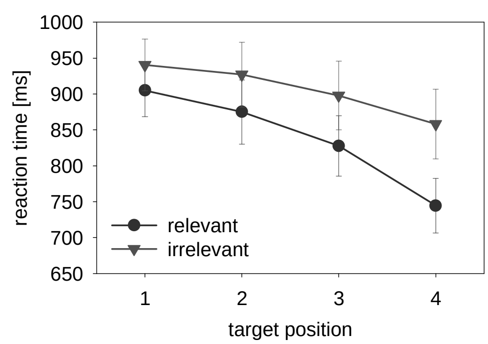

Figure 3: Reaction time to relevant and irrelevant targets at the different target positions.

Relevant vs. irrelevant vs. just rejected. To test the predictions regarding rejected explanations, the reaction times to relevant, irrelevant, and just rejected targets were compared. In accordance with our prediction, reactions to rejected targets were slower than to relevant ones. However, they were not slower than reactions to irrelevant targets (see

Figure 4). A significant main effect for target type was confirmed by a 3 (target type) x 3 (target position) within-subjects ANOVA, $F(2, 50) = 9.720$, $p < .001$, partial eta-square = .280. Pairwise Bonferroni-adjusted comparisons showed that reaction times were significantly faster for relevant targets than for both irrelevant ($p = .002$) and rejected targets ($p = .030$). Irrelevant and rejected targets did not differ significantly. Again, a significant main effect for target position was found, ($F(2, 50) = 23.916$, $p < .001$, partial eta-square = .489), whereas the interaction between target type and position was not significant.

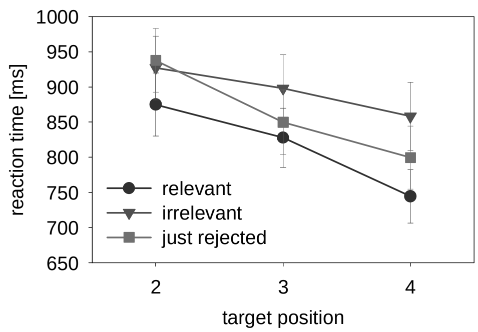

Figure 4: Reaction time to relevant, irrelevant and just rejected targets at the different target positions.

Effect of time since explanation rejection. In addition to the above analysis we examined the effect of the time elapsed since a hypothesis had to be rejected. Therefore, reaction times to targets presented after the 4th symptom were compared (see Figure 5).

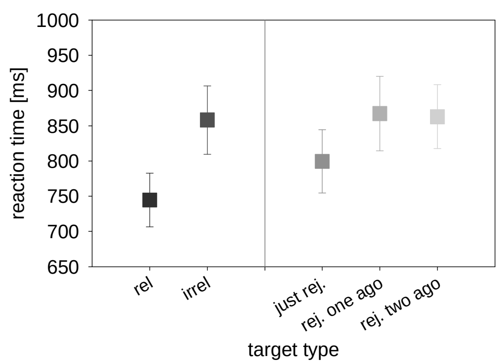

Figure 5: Reaction time to all types of targets (x-axis) presented after the 4th symptom.

As expected, reactions to relevant targets were the fastest. Reactions to rejected targets were slower and the response time increased the more time elapsed between the presentation of the contradicting symptom and the target presentation. A one-factorial ANOVA confirmed, that the reaction times after the 4th symptom differed significantly between the different target types, $F(4, 100) = 4.643$, $p = .002$, partial eta-square = .157. For further analysis, pairwise comparisons with Bonferroni adjustment were carried out. Relevant targets were significantly faster than targets rejected one ($p = .030$) or two ($p = .023$) symptoms ago. Just rejected targets were marginally faster than targets rejected one symptom ago ($p = .077$). All other comparisons were not significant.

Discussion

The results of the probe reaction task show that reactions to targets that are related to relevant explanations are faster than reactions to targets that are related to irrelevant explanations or explanations that had to be rejected during the abductive reasoning task. This indicates that explanatory hypotheses that are relevant in the current context for explaining a set of observations are kept in a more active state in memory than irrelevant or rejected hypotheses.

Considering the effect of target position on reaction times three aspects seem noteworthy. First, the difference between relevant and irrelevant hypotheses tended to increase with the number of symptoms shown in a trial (see Figure 2). On the one hand this might be due to the fact that with a growing number of symptoms presented in a trial the number of symptoms confirming the relevant hypotheses grew, leading to an increase of activation for relevant hypotheses. On the other hand, with an increasing number of symptoms more and more hypotheses had to be rejected as they were not compatible with all symptoms presented so far. Hence, the available activation became focused to a decreasing number of relevant hypotheses (cf. Anderson, 1983) leading to an increase of the activation of the still relevant hypotheses. This also means that the reaction times in the probe tasks are not simply an effect of associative priming due to links between the symptoms and causal hypotheses. Focusing activation on a decreasing number of hypotheses is only possible when only those explanations for a new symptom are considered as relevant that are also compatible with the previously presented symptoms. If the reaction times in the probe task were only the result of an associative priming process due to the links between the just presented symptom and the associated causes, the course of activation should be the same for relevant and irrelevant hypotheses.

Second, the fact that the reaction times to irrelevant explanations also decreased, even though to a smaller extent than the reaction times to relevant explanations, might be due to a growing activation of the probe task itself. The more symptoms were shown in a trial the more likely the presentation of a probe became, thereby increasing the activation of the probe task set after each symptom.

Third, regarding the question whether rejected explanatory hypotheses are actively inhibited or show just a decay of activation to baseline level, the results support Kintsch's (1998) decay assumption. Reaction times to rejected targets increased with increasing time interval since rejecting the respective hypothesis until they reached the level of irrelevant targets. This indicates that the activation of rejected explanatory hypotheses decreased over time to the level of irrelevant hypotheses but not beyond this level.

Hence, our results provide no support for an inhibition process of rejected hypotheses.

Conclusions

The goal of this paper was to examine how automatic comprehension processes are involved in the generation of explanations for observations. We assumed that these comprehension processes support the abductive reasoning process by making the most plausible explanatory hypotheses for the given context highly available so that deliberate reasoning processes are constrained to these hypotheses. This should be the case especially in routine situations where the reasoner has available a rich body of domain knowledge that is fine-tuned to the task structure. We used Kintsch's (1998) construction-integration theory as a framework for these automatic comprehension processes to predict that relevant hypotheses should become activated during the reasoning process. Hypotheses that were first considered as relevant but had to be rejected later because of new incompatible observations should decrease in activation or should be actively inhibited.

The results of the experiment indicate that automatic comprehension processes are indeed involved in abductive reasoning. They lead to the activation of relevant hypotheses in comparison to irrelevant ones. And this activation of relevant hypotheses increases in the course of the abductive reasoning task when more confirming observations are processed and the activation of hypotheses becomes focused to a decreasing number of still relevant hypotheses. The results did not provide support for the notion of an active inhibition process of rejected hypotheses. It rather seems that after being rejected hypotheses are not kept activated anymore and therefore loose activation until they reach the baseline level activation.

Our results demonstrate how automatic comprehension processes in the context of abductive reasoning tasks serve to reduce the complexity of these tasks by constraining the number of explanations that are considered first as possible explanations to those that remain highly relevant in the developing context.

Acknowlegements

We thank Frank Ritter, Josef Krems, Georg Jahn, and three anonymous reviewers for their helpful comments on earlier drafts of this paper.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behaviour, 22*, 261-295.

Arocha, J. F. & Patel, V. L. (1995). Construction-integration theory and clinical reasoning. In Ch. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension. Essays in honor of Walter Kintsch.* Hillsdale, NJ : Lawrence Erlbaum Associates.

Baumann, M. R. K., Bocklisch, F., Mehlhorn, K., & Krems, J.F. (in press). Changing explanations in the face of anomalous data in abductive reasoning. *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. August 01-04, 2007, Nashville, USA.

Conway, A. R. A. & Engle, R. W. (1994). Working Memory and Retrieval: A Resource-dependent Inhibition Model. *Journal of Experimental Psychology: General, 123*, 354-373.

Doane, S. M., McNamara, D. S., Kintsch, W., Polson, P. G., & Clawson, D. (1992). Prompt comprehension in UNIX command production. *Memory and Cognition, 20*, 327–343.

Doane, S. M. & Sohn, Y. W. (2000). ADAPT: A predictive cognitive model of user visual attention and action planning. *User Modeling and User-Adapted Interaction 10*, 1-45.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science, 4*, 294-298.

Gernsbacher, M. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Groen, G. J., & Patel, V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. Chi & R. Glaser (Eds.), *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Johnson, T. R., Zhang, J., & Wang, H. (1997). A hybrid learning model of abductive reasoning. In R. Sun & F. Alexandre (Eds.), *Connectionist symbolic integration*. Hillsdale, NJ: Lawrence Erlbaum.

Josephson, J., & Josephson, S. G. (1994). *Abductive inference: Computation, philosophy, technology*. New York, NY: Cambridge University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York, NY: Cambridge University Press.

Lustig, C., Hasher, L., & Tonev, S. T. (2001). Inhibitory control over the present and the past. *European Journal of Cognitive Psychology, 13*, 107-122.

Schmalhofer, F. & Tschaitschian, B. (1993). The acquisition of a procedure schema from text and experiences. *Proceedings of the 15th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3-22.

Thagard, P. & Kunda, Z. (1998). Making sense of people: Coherence mechanisms. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior*. Hillsdale, NJ: Erlbaum.