Inference Suppression and Working Memory Capacity: Inhibition of the Disabler Search

Wim De Neys (Wim.Deneys@psy.kuleuven.ac.be)
Kristien Dieussaert (Kristien.Dieussaert@psy.kuleuven.ac.be)
Walter Schaeken (Walter.Schaeken@psy.kuleuven.ac.be)
Géry d'Ydewalle (Géry.dYdewalle@psy.kuleuven.ac.be)
Department of Psychology, K.U.Leuven, Tiensestraat 102
B-3000 Leuven, Belgium

Abstract

We examined whether individual differences in WM-capacity affected the acceptance of an extended MP problem that explicitly mentioned a possible disabler. The explicit disabler presentation was assumed to stimulate the spontaneous disabler search process. Two experiments showed that the acceptance ratings of the extended MP problems followed a U-shaped, quadratic trend with low and high spans showing the highest MP acceptance. Contrasting performance with extended and standard MP problems indicated that all span groups showed the standard suppression effect. Findings support the claim that high spans manage to inhibit the spontaneous disabler search and underline the generality and robustness of this inhibition phenomenon.

Introduction

Suppose that on a hot, summer day you hear someone claiming "If Jenny turns on the air conditioner, then she will feel cool". Next, you hear that Jenny did turn the air conditioner on. It is likely that you will conclude that Jenny will feel cool. This inference ('If P then Q. P. Therefore, Q') is known as the Modus Ponens (MP). The MP inference is considered valid in standard logic. Now, suppose that you would also have been reminded of the fact that the air conditioner might be broken or that Jenny might have a fever. In this case you would probably have been rather reluctant to accept the standard MP inference that turning on the air conditioner will make Jenny feeling cool. Thereby, the additional information would have tempted you to commit a fallacy.

Cognitive scientists have spent a great deal of research to establish how people reason with 'if, then' sentences or conditionals. One of the main findings is that additional, 'background' knowledge about the conditional relation affects the inferences people are willing to draw (Evans, Newstead, & Byrne, 1993; Manktelow, 1999). The crucial kind of background knowledge for the evaluation of the MP inference is referred to as 'disabling conditions'. A disabling condition (also 'disabler' or 'additional requirement') is a condition that prevents the antecedent specified in the conditional (e.g., turning on the air conditioner or the P part) from bringing about the consequent (e.g., feeling cool or the Q part). In the introductory example a broken air conditioner or Jenny having a fever will both function as disablers.

In a pioneering study Byrne (1989) showed that when a possible disabler was explicitly presented to participants (e.g., If she has an essay to write, then she will study late in the library. If the library is open, then she will study late in the library. She has an essay to write. Thus, she will study late in the library?) the MP inference was less frequently accepted compared to the standard MP condition without presented disabler. Further studies established that during conditional reasoning people spontaneously search their long-term memory for stored disablers. Cummins (1995) used causal conditionals for which a pilot group could retrieve many (e.g., If you put fertilizer on plants, then they grow well) or only few (e.g., If Tom grasps the glass with his bare hands, then his fingerprints are on it) disablers. Cummins reasoned that for conditionals with many (vs. few) disablers spontaneous retrieval of a disabler would be more likely. Although no specific disablers were explicitly presented, the results indeed showed that MP inferences based on conditionals with many disablers were rejected more frequently. Numerous studies confirmed these findings (e.g., Bonnefon & Hilton, 2002; Byrne, Espino, & Santamaria, 1999; Thompson, 1994; De Neys, Schaeken, & d'Ydewalle, 2002; Stevenson & Over, 1995; George, 1997; see Politzer & Bourmaud, 2002 for a review). Thus, it is well established that finding a disabler (either spontaneously or presented by the experimenter) will result in a decreased MP acceptance. This impact of disablers on the MP inference acceptance is known as the suppression effect.

In the present study we present the first experiments that look at individual differences in the suppression impact. More precisely, we will examine whether differences in working memory (WM) capacity affect the acceptance of MP problems when a possible disabler is explicitly presented. Such WM-mediation could be expected on the basis of recent findings pointing to the role of WM in the retrieval and inhibition of stored disablers (e.g., De Neys, Schaeken, & d'Ydewalle, 2003a, 2003b; Markovits, Doyon, & Simoneau,

2002; Simoneau & Markovits, 2003; Verschueren, De Neys, Schaeken, & d'Ydewalle, 2002).

De Neys et al. (2003a) and Verschueren et al. (2002) established that the efficiency of the disabler retrieval process is mediated by WM-capacity. In a task where people were asked to generate disablers for a set of conditionals in limited time, participants higher in WM-capacity retrieved more disablers. Putting a load on WM also reduced the efficiency of the retrieval process. In a further experiment, De Neys et al. (2003b) tested a group of low, medium, and high spans (participants in the bottom, middle and top quintile of first-year psychology students' WM-capacity distribution, respectively) in an everyday conditional reasoning task. Consistent with the more efficient disabler retrieval, medium spans were more likely to reject the MP inference than low spans. On the other hand, despite the intrinsic superior retrieval capacity high spans showed nevertheless higher MP acceptance ratings than the medium spans (see also Markovits et al., 2002).

Based on findings of Stanovich and West (2000), it was assumed that a basic decontextualization ability would allow high spans to put background knowledge aside when it conflicts with the logical standards. Remember that in standard logic MP is a valid inference. Since disabler retrieval will result in the rejection of MP, a basic validity notion will conflict with the disabler retrieval process. De Neys et al. reasoned that high spans would therefore use their WM-resources for an active inhibition of the disabler search.

Simoneau and Markovits (2003) showed that more efficient inhibitory processing (as measured by a negative priming task) was indeed linked with higher MP acceptance. In a related dual-task study De Neys et al. found additional support for the inhibition hypothesis. The basic assumption states that lower spans allocate WM-capacity to the disabler retrieval, while high spans allocate WM-capacity primordially to the retrieval inhibition. Consistent with the hypothesis, the dual-task study showed that a less efficient disabler retrieval under WM-load resulted in higher MP acceptance ratings under load (vs. no load) for low spans, while the less efficient inhibition resulted in lower MP ratings under load (vs. no load) for high spans.

In sum, there is evidence for the claim that high spans are inhibiting the disabler retrieval process during conditional reasoning. Inhibition of cognitive processes deemed inappropriate is indeed one of the key executive working memory functions (e.g., Baddeley, 1996; Levy & Anderson, 2002; Miyake & Shah, 1999).

The present study will allow a further test of the disabler inhibition hypothesis. Presenting extended MP problems where a possible disabler is explicitly mentioned will push the inhibition demands to the limit. The explicit disabler presentation will stimulate the search process. Note that one of the difficulties of retrieving disablers in a reasoning task is that there is no explicit retrieval cue (e.g., Markovits & Barrouillet, 2002; Markovits & Quinn, 2002). The standard MP premises do not tell you what kind of information you should look for. If a possible disabler is added, it will be incorporated in the elementary mental representation of the inference problem. It is assumed that this representation is held in working memory. As suggested by many authors, activation will automatically start to spread from the information stored in WM (or "the focus of attention" see Cowan, 1995) to related long-term memory elements (Anderson, 1993; Cowan, 1995; see also Markovits & Barrouillet, 2002 for an integrated account). Therefore, stored disablers will receive more activation by the explicit presentation of a possible disabler with the MP inference. Consequently, it will be more likely that additional stored disablers will be automatically retrieved.

In Experiment 1 participants were given a measure of WM-capacity and extended MP problems that mentioned a possible disabler. If the high spans still manage to inhibit the stimulated disabler search we expect to see a U-shaped, quadratic trend in the acceptance ratings in function of WM-capacity. Low spans were expected to show high MP acceptance ratings because of the inefficient disabler retrieval. Medium spans should show lower MP ratings because the search will be more efficient. If high spans still manage to inhibit the disabler retrieval, acceptance ratings should increase again for the high spans.

In Experiment 2 we compared the acceptance ratings of standard and extended MP problems for different WM-span groups.

Experiment 1

Method

Participants

A total of 105 first-year psychology students from the University of Leuven (Belgium) participated in the experiment in return for course credit. None of the students had had any training in formal logic.

Material

Working memory task. Participants' working memory capacity was measured using a version of the Operation span task (Ospan, La Pointe & Engle, 1990) adapted for group testing (Gospan, for details see De Neys, d'Ydewalle, Schaeken, & Vos, 2002). In the Ospan-task participants solve series of simple mathematical operations while attempting to remember a list of unrelated words The main adaptation in the Gospan is that the operation from an operation-word pair is first presented separately on screen (e.g., 'IS (4/2) – 1 = 5 ?'). Participants read the operation silently and press a key to indicate whether the answer is correct or not. Responses and

response latencies are recorded. After the participant has typed down the response, the corresponding word (e.g., 'BALL') from the operation-word string is presented for 800 ms. As in the standard Ospan three sets of each length (from two to six operation-word pairs) are tested and set size varies in the same randomly chosen order for each participant. The Gospan-score is the sum of the recalled words for all sets recalled completely and in correct order.

Participants were tested in groups of 21 to 48 at the same time. Participants who made more than 15% math errors or whose mean operation response latencies deviated by more than 2.5 standard deviations of the sample mean were discarded (participants already in the bottom quartile of the Gospan-score distribution were not discarded based on the latency criterion). De Neys, d'Ydewalle et al. (2002) reported an internal reliability coefficient alpha of .74 for the Gospan. The corrected correlation between standard Ospan and Gospan-score reached .70.

Reasoning task. Participants received three extended MP problems. These were Dutch translations of the three Byrne (1989) MP problems (see Dieussaert, Schaeken, Schroyens, & d'Ydewalle, 2000). The following item format was used:

Rule: If she has an essay to write, then she will stay late in the library.
If the library stays open, then she will stay late in the library.
Fact: She has an essay to write
Conclusion: She will stay late in the library.

All three MP problems were presented on a separate page of a booklet together with a 7-point rating scale ranging from 1 (*Very certain that I cannot draw this conclusion*) to 7 (*Very certain that I can draw this conclusion*) with 4 representing *can't tell*. Participants placed a mark on the number of the scale that best reflected their evaluation of the conclusion.

Procedure

Participants were tested in groups of 21 to 42 at the same time in a large computer room with an individual booth for every participant. All participants started with the Gospan task that was run on computer. After all participants of a group had finished the Gospan-task the extended MP evaluation task was presented. The three items were presented on separate pages of a booklet. The first page of the booklet included the task instructions. They showed an example item that explained the specific task format. Participants were told that the task was to decide whether or not they could accept the conclusions. Care was taken to make sure participants understood the precise nature of the rating scale. The task instructions did not mention to accept the premises as true or to endorse conclusions that follow necessarily. Instead participants were told they could evaluate the conclusions by the criteria they personally judged relevant (see Cummins, 1995).

Results and discussion

Rejection probability for all reported statistical analyses was .05. For completeness, we always report the individual estimated p-values.

Three participants were discarded because they did not meet the operation correctness or latency requirements of the WM-task (see De Neys, d'Ydewalle et al., 2002). The remaining 102 participants were split in three span groups of equal n based on the boundaries of the Gospan-score distribution. Mean Gospan-score for the three successive span groups was 23.27 (SD = 4.34, low span), 35.15 (SD = 2.79, medium span), and 45.89 (SD = 4.97, high span).

For every participant we calculated the mean acceptance rating for the three extended MP problems. The means were subjected to an ANOVA with span group as between-subject variable. There was a significant effect of span group, $F(2, 99) = 5.55$, $MSE = 1.23$, $p < .01$. The acceptance rating showed the expected pattern: Medium spans ($M = 3.84$, $SD = 4.67$) showed lower MP acceptance ratings than the low ($M = 4.56$, $SD = 1.11$) and high spans ($M = 4.67$, $SD = .99$). A trend analysis confirmed that there was a significant U-shaped, quadratic trend, $F(1, 99) = 10.85$, $MSE = 1.23$, $p < .005$ without mediation of a linear trend, $F(1, 99) < 1$.

Thus, even when the disabler search was specifically stimulated high spans showed the highest levels of MP acceptance. This is consistent with the claim that high spans are inhibiting the disabler search and underlines the generality and robustness of the inhibition phenomenon.

Experiment 2

The first experiment showed that the acceptance ratings for the extended MP problems differed for participants of different WM-capacity. In Experiment 2 we compare the acceptance ratings of standard vs. extended MP problems in function of WM-span. This allows us to establish the impact of the explicit disabler presentation per se. For the validity of our framework it is crucial that the acceptance ratings decrease when a disabler is explicitly presented.

First, for low spans it is assumed that the disabler search with standard MP problems will not be very successful. Although low spans' limited resources will restrict the impact of the extra search stimulation, the extended disabler manipulation does present low spans a disabler they will probably not retrieve in the standard condition. Therefore, low spans' inference acceptance should decrease for the

extended MP problems. Second, because of the more efficient retrieval, medium spans in the standard condition will probably retrieve the disabler presented in the extended MP condition themselves. Hence, the mere presentation of the disabler should not affect medium spans. Nevertheless, if we are right that the search process is stimulated by the disabler presentation one should expect that additional disablers will be retrieved in the extended condition and this should further decrease the MP acceptance (see De Neys, Schaeken, & d'Ydewalle, 2003c, for a study on the effect of the number of retrieved disablers on MP acceptance). High spans are expected to inhibit the search both for the standard and extended problems. However, it is explicitly assumed that the inhibition is not automatic, but draws on WM-resources. Therefore, the inhibition should be less successful when the process is more demanding. De Neys et al. (2003a) already observed that the increasing inhibition demands caused by an increasing number of available disablers resulted in a less efficient disabler inhibition. Hence, although the high spans should overall show a high MP acceptance, their acceptance level should nevertheless be affected by the stimulated disabler search.

In sum, we expected a standard suppression effect for all span groups: Acceptance ratings should be lower for the extended (vs. standard) MP problems. In addition, overall MP acceptance ratings should be affected by WM-span: Extended and standard MP acceptance in the successive span groups should follow the U-shaped trend observed in Experiment 1.

Method

Design

As standard condition or baseline we used the MP evaluations of the 282 participants in the study of De Neys et al. (2003b). In this study participants were presented a standard conditional inference task with causal conditionals and a measure of WM-capacity. We calculated the mean MP acceptance for different span groups and used this as a baseline to compare the MP acceptance of matched span groups with similar extended causal MP problems.

Participants

All 105 participants of Experiment 1 evaluated the extended MP inferences in the present experiment. The data for the standard MP condition were taken from the study of De Neys et al. (2003b) where 282 first-year psychology students evaluated standard conditional inferences.

Material

Working memory task. All participants' working memory capacity was measured with the Gospan-task (see De Neys, d'Ydewalle et al., 2002).

Reasoning tasks. All conditionals were selected from the generation studies of De Neys et al. (2002) and Verschueren et al. (2002). Eight causal conditionals were used for the standard condition and six causal conditionals for the extended condition. Half of the conditionals in each condition were previously classified as having many possible disablers, while the other half had only few possible disablers. The number of possible alternative causes (see Cummins, 1995) of the selected conditionals with few and many disablers was kept constant. The item format for the extended and standard task was similar to the format used in Experiment 1, except that for the extended items a possible disabler was mentioned. We always presented the disabler that was most frequently generated for that conditional in the generation task (e.g., see De Neys et al., 2002). As in Byrne (1989) the disablers (e.g., engine broken) were always presented as an additional requirement, embedded in a conditional (e.g., If the engine works, then the car starts). This resulted in the following format:

Rule: If the ignition key is turned, then the car starts.
If the engine works, then the car starts.
Fact: The ignition key is turned.
Conclusion: The car starts.

It should be noted that the set of conditionals in the standard and extended condition was not completely similar. Although both conditions used causal conditionals with a comparable number of possible disablers, the standard condition should therefore not be conceived as a control condition per se. Rather, the standard condition serves as a baseline against which the performance of the different WM-span groups can be compared.

Procedure

Participants were tested in groups of 21 to 48 at the same time in a large computer room with an individual booth for every participant. All participants started with the Gospan task that was run on computer. After all participants of the group had finished the Gospan-task the extended MP evaluation task or the standard conditional inference task was presented. The standard task was run on computer. Participants evaluated eight standard MP inferences mixed with other conditional inferences. The six items of the extended MP task were presented on separate pages of a booklet. This booklet was presented before the booklet with the items of Experiment 1. Task instructions for the standard and extended MP task were similar to the instructions given in Experiment 1.

Results and discussion

In order to match the span groups in the extended and standard conditions as closely as possible we decided to split both samples up in five span groups

each, based on the quintile boundaries of the Gospan-score distribution of the 282 participants in the standard condition. A 5 (span group, between-subjects) x 2 (MP task, between-subjects) ANOVA on the Gospan-scores established that there were no WM-capacity differences for participants in both task conditions [effect of MP task, $F(1, 374) < 1$; interaction MP task x Span-group, $F(4, 374) = 1.84$, $MSE = 10.91$].

Each participant evaluated eight or six MP evaluations. The mean of these ratings was calculated and subjected to a 5 (WM-span, between-subjects) x 2 (MP task, between-subjects) ANOVA.

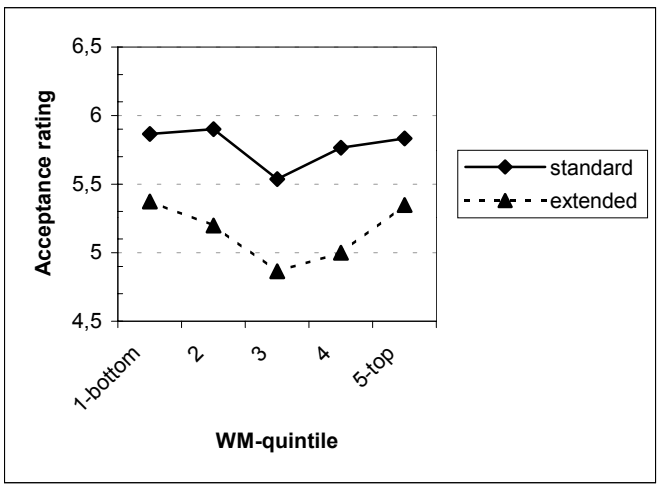

Figure 1. Mean MP acceptance rating in function of WM-capacity with (extended) and without (standard) explicitly presented disabler. The rating scale ranged from 1 (very sure cannot draw this conclusion) to 7 (very sure can draw this conclusion).

Explicitly presenting a disabler clearly decreased the MP acceptance, $F(4, 374) = 37.56$, $MSE = .72$, $p < .0001$. Figure 1 shows that, as expected, this effect was present for all WM-span groups, span group x MP task interaction, $F(4, 374) < 1$. There was also a marginal main effect of WM-span, $F(4, 374) = 2.28$, $MSE = .72$, $p < .06$. As Figure 1 indicates, a trend analysis clearly established that the MP ratings followed a U-shaped, quadratic trend in function of WM-span, $F(1, 374) = 6.77$, $MSE = .72$, $p < .01$. There was no sign of a linear trend, $F(1, 374) < 1$, and the quadratic trend did not differ for the standard and extended MP problems, $F(1, 374) = 1.07$, $MSE = .71$, $p > .35$. Thus, as expected, all span groups showed an impact of the explicit disabler presentation, but both on the standard and extended problems the MP acceptance ratings were affected by WM-capacity.

Finally, one might note that the number of disablers of the adopted conditionals in the present experiment varied systematically (e.g., half of the conditionals had few vs. many possible disablers). We had no specific hypotheses concerning the impact of this factor on the manipulations. For completeness, the variable was entered as a within-subjects factor in the ANOVA. We replicated the traditional (e.g., Cummins, 1995) main effect of the number factor: MP acceptance was always lower for conditionals with many disablers than for conditionals with few disablers, $F(1, 374) = 139.40$, $MSE = .51$, $p < .0001$. However, none of the interactions with the other factors reached significance. Thus, the crucial effects of span group and the explicit disabler presentation were not affected by the number factor.

General Discussion

The two experiments clearly established that even when a disabler is explicitly presented MP acceptance ratings of the successive working memory (WM)-span groups follow a U-shaped trend. Previous studies already suggested that the higher MP acceptance ratings of high vs. medium spans, despite high spans' superior retrieval capacities, result from an active inhibition of the disabler search. The fact that in the present study the same pattern is found under conditions that can be assumed to stimulate the search process points to the robustness and generality of the inhibition phenomenon.

As De Neys et al. (2003a, 2003b) we hypothesized that the disabler inhibition is not occurring in a cognitive vacuum but draws on working memory resources. Therefore, higher inhibition requirements were expected to result in a less efficient inhibition process. The goal of the search stimulation by the explicit disabler presentation was precisely to increase the inhibition demands. Disablers that would be inhibited under less demanding inhibition conditions could 'slip through' the filter and decrease the MP acceptance. Consistent with these hypotheses Experiment 2 clearly showed that even high spans' MP acceptance decreased for the extended MP problems.

The present findings have implications for traditional suppression studies. The results indicate that Byrne's (1989) findings can be generalized over the whole WM-capacity distribution: For all WM-span groups MP acceptance decreased when a possible disabler was explicitly presented. Thus, all WM-span groups show the basic suppression effect. However, it is important to note that the final acceptance level is systematically affected by WM-capacity: Reasoners with an inefficient disabler retrieval and reasoners that inhibit the retrieval show the highest levels of MP acceptance. These people will be typically situated in the bottom and top levels of the WM-capacity distribution, respectively. Reasoner that can allocate sufficient resources to the retrieval and do not inhibit the search process, typically people with medium sized WM-span, will be most likely to reject MP. These findings further

emphasize the role of WM-capacity in the retrieval and inhibition of disablers during conditional reasoning.

Acknowledgements

Preparation of this manuscript was supported by grants from the Fund for Scientific Research-Flanders (FWO).

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Bonnefon, J.-F., & Hilton D. J. (2002). The suppression of modus ponens as a case of pragmatic preconditional reasoning. *Thinking & Reasoning, 8*, 21-40.

Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition, 31*, 61-83.

Byrne, R. M. J., Espino, O., & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language, 40*, 347-373.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Cummins, D. D. (1995). Naive theories and causal deduction. *Memory and Cognition, 23*, 646-658.

De Neys, W., d'Ydewalle, G., Schaeken, W., & Vos, G. (2002). A Dutch, computerized, and group administrable adaptation of the operation span test. *Psychologica Belgica, 42,* 177-190.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Causal conditional reasoning and semantic memory retrieval: A test of the semantic memory framework. *Memory & Cognition, 30,* 908-920.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003a). *Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples*. (Tech. Rep. No. 295). Leuven, Belgium: University of Leuven, Laboratory of Experimental Psychology.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003b). *Working memory span and everyday conditional reasoning: A trend analysis*. Paper presented at the Twenty-Fifth Annual Meeting of the Cognitive Science Society, Boston, MA.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003c). Inference suppression and semantic memory retrieval: Every counterexample counts. *Memory & Cognition, 31,* 581-595.

Dieussaert, K., Schaeken, W., Schroyens, W., & d'Ydewalle, G. (2000). Strategies during complex conditional inferences. *Thinking & Reasoning, 6*, 125-160.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hove, UK: Erlbaum.

George, C. (1997). Reasoning from uncertain premises. *Thinking & Reasoning, 3*, 161-189.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1118-1133.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6*, 299-305.

Manktelow, K. I. (1999). *Reasoning and thinking*. Hove, UK: Psychology Press.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22*, 5-36.

Markovits, H., Doyon, C., & Simoneau, M. (2002). Individual differences in working memory and conditional reasoning with concrete and abstract content. *Thinking & Reasoning, 8*, 97-107.

Markovits, H., & Quinn, S. (2002). Efficiency of retrieval correlates with logical reasoning from causal conditional premises. *Memory & Cognition, 30*, 696-706.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge, MA: Cambridge University Press.

Politzer, G., & Bourmaud, G. (2002). Deductive reasoning from uncertain conditionals. *British Journal of Psychology, 93*, 345-381.

Simoneau, M., & Markovits, H. (2003). Reasoning with premises that are not empirically true : Evidence for the role of inhibition and retrieval. *Developmental Psychology, 39*, 964-975.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences, 23*, 645-726.

Stevenson, R. J., & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology, 48A*, 613-643.

Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory and Cognition, 22*, 742-758.

Verschueren, N., De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). *Working memory capacity and the nature of generated counterexamples*. Paper presented at the Twenty-Fourth Annual Meeting of the Cognitive Science Society, Fairfax, VA.