2007

The role of working memory in transfer of implicit learning

Daniel Yurovsky
Carnegie Mellon University

The role of working memory in transfer of implicit learning

Daniel Yurovsky

Carnegie Mellon University

Advisor: Ken Kotovsky, Ph. D.

H&SS Honors Thesis

Abstract:

Implicit learning, behavioral change accompanied by an inability to consciously describe the means by which it has occurred, has been demonstrated in a number of domains. One question concerns the role of working memory in the learning process – if participants do not have conscious access to the learned information, what is the role of conscious attention and working memory in their learning? This paper further explores the question by studying the role of working memory on transferability of implicitly learned knowledge in the Balls and Boxes problem. Participants were given a puzzle to solve and then either the same puzzle or a horizontally inverted isomorph under single-task or working-memory interference conditions. As hypothesized, participants have little difficulty transferring their learned knowledge to the new problem unless their working memory is loaded.

The role of working memory in transfer of implicit learning

In 1967, Arthur Reber demonstrated something extraordinary. He showed participants a series of nonsense strings (e.g 'VXVS') that were generated according to an artificial grammar; for each letter there was a small set of legal letters which could follow it. After training, he showed participants a number of novel strings, some of which were legal according to the grammar and some of which were not. Amazingly, despite demonstrating no explicit knowledge of the grammar's rules, participants were able to discriminate the legal strings from the illegal ones. Reber called this peculiar process of acquiring the unconscious ability to make judgments about complex stimuli 'implicit learning.'

This finding lead to a renewed interested in the unconscious processes that underlie learning and spawned quite a number of studies in varied domains. Research proceeded in artificial grammar (Reber & Allen, 1978, Mathews, Buss, Stanley, Blanchard-Fields, Cho, & Druhan, 1989, Poletiek, 2002), covariation learning (Lewicki, 1986, Miller, 1987, Musen & Squire, 1993) and dynamic systems (Berry & Broadbent, 1984, Squire & Frambach, 1990). Effects were found, for the most part, in relatively simple and passive paradigms, which leaves them open to criticism that participants are acquiring minor changes to uncomplicated representations. The trivial nature of these changes may be the reason for the unlikeliness of their rising to awareness. In light of this, it is useful to study the nature of implicit learning in a more complex domain: problem solving.

A crucial point is to find a good problem to study. Kotovsky and Simon (1990), in a study of the relation between difficulty and problem isomorphs (problems with identical problem states but different surface representations) created one such problem. They began with the Chinese Ring Puzzle (Figure 1A) – an incredibly difficult steel tavern puzzle in which moves are made by manipulating a set of rings on a bar. Finding that the majority of their participants were unable to solve it in the allotted 180 minute period, they created several digital isomorphs (Kotovsky & Simon, 1990). Both isomorphs are problems in which the participant is presented with five boxes containing balls and is required to move the puzzle into a state in which all balls are out of their respective boxes. In the lo-info version of the puzzle (Figure 1B) , the set of currently available moves is clear at any point in the problem state, while in the no-info version (Figure 1C) there is no visual information indicating legal moves. Kotovsky and Simon administered two trials of the Balls and Boxes puzzle, varying which puzzle (lo or no info) participants solved on each trial. They found a significant decrease in the number of moves, as well as the time required to solve the puzzle, on the second trial despite the fact that participants had no explicit knowledge of the rule they used to solve it. They also found the lo-info version of the puzzle to be a better recipient of transfer. It was characterized by the greatest proportional decrement in moves and time on the second trial relative to the first, regardless of which version was administered initially. Kotovsky and Simon concluded that this was a result of problem difficulty: since the lo-info problem was easier to solve it would also be easier to transfer to. Thus, transfer was modulated not by similarity but by ease of the target problem. Participants required fewer resources to transform their learned knowledge from the no-info problem for use on the lo-info problem than to apply it directly to the no-info problem. One likely candidate resource is working memory.

Working memory, proposed by Baddeley and Hitch (1974), refers to the resource the brain uses to temporarily store and manipulate incoming information before it is made permanent. The role of working memory in the Balls and Boxes puzzle was examined in 1997 by Reber and Kotovsky who asked participants to solve the lo-info version of the puzzle on two successive trials. Each time the participant was also given a concurrent verbal working-memory interference task. This task, the n-back task, required participants to listen to a stream of random numbers played at a rate of one every three seconds, and report back a recently heard number when prompted by a beep. The number they were to report varied by condition from the most recent number to the fourth-to-last one. They found that when participants solved the puzzle while simultaneously remembering at least three numbers at a time their performance was impaired on the first trial. However, their performance on successive trials was not impaired regardless of level of memory load. Reber and Kotovsky therefore concluded that working-memory was involved in the implicit learning of the problem's rule. But, since participants were not impaired on the following trials, they argued that working memory was not involved in the production of moves according to the already learned rule. A similar pattern was also seen in artificial grammar learning tasks (Dienes, Broadbent, & Berry, 1991). So it seems clear that working memory is involved in learning of the rule and not in production, but what about transfer?

Transfer of implicitly learned information requires participants to transform their acquired knowledge into a form that is useful in the new task. Both Kotovsky and Simon (1990) and Reber and Kotovsky (1997) showed that there was an effect of transfer from trial one to trial two of the same puzzle. They also showed that transfer is facilitated more by an easy than a hard target problem. In thus study, we wanted to explore the nature of transfer to an equally hard, but different puzzle. This way, we could be sure that trouble with transfer was not the result of variability in difficulty, but rather in the amount of working memory required to transform learned rules. What would happen if on the second trial the puzzle was horizontally inverted? In order to transfer their implicitly learned knowledge the participants would need to recruit working memory to recode information about problem states into a form that would be useful for the inverted puzzle. It seemed quite likely that participants would not show great difficulty in transferring their knowledge under normal conditions – after all the manipulation is quite a simple one. The more interesting question concerns the role of working memory in such a transfer. Given that working memory was involved in learning the rule, would it also be involved in translating the rule? We hypothesized that there would be an interaction between working-memory load and transfer puzzle. That is, participants would be relatively unimpaired in transferring knowledge to a second trial of the same puzzle regardless of their level of working memory load. On the other hand, we expected them to be significantly impaired at transferring from the original puzzle to the inverted puzzle, a task that presumably imposes more of a memory load, but only when their working memory was loaded.

Methods

In this experiment we wanted to examine the role of working memory in translating implicitly learned information to a novel problem. In order to load working memory we used a verbal memory task. Participants were given a brief training session to familiarize them with the 2-back working memory load task and then asked to solve the low-info version of the balls and boxes problem. Subsequently, they were divided into four conditions. Half of the subjects solved the same puzzle a second time while the other half solved a horizontal inversion. Additionally, half of the subjects solved the puzzle under a control condition, while the other half were given a 2-back working memory interference task. Number of moves and time to solution for each participant were recorded.

Participants

Forty undergraduate Carnegie Mellon students participated in partial fulfillment of course requirements for first or second-level psychology courses.

Materials

The problem used was the 21-move lo-info version of the Balls and Boxes puzzle introduced by Kotovsky and Simon (1990). For the transfer puzzle, a horizontal inversion was created. In this version, the leftmost ball acted as if it were the rightmost ball, and the second from the left as if it were second from the right. A starting state comparison can be seen in Figure 2.

Sound clips used in the n-back task were all recorded by the experimenter. Each was approximately one second long and all were recorded at the same volume. Each number zero through nine was uniquely represented by one clip. There was also one clip of a beep that was acquired from a freeware sound website.

All stimuli were presented on a 17 inch CRT monitor using a PC computer. Sounds were played at a volume selected by participants to ensure they did not have to strain to hear. The experiment itself was programmed in Java by the experimenter and run using the Eclipse Development Environment.

Design and Procedure

Each participant began the experiment by completing one minute of training on a 2-back working memory interference task. Numbers between zero and nine were played by the computer in random order at the rate of one every three seconds. At randomly chosen intervals of four to 10 numbers, a beep was played and participants were asked to verbally report the third to last number played. After a minute, the participant was allowed to repeat the training procedure at the discretion of either the participant or the experimenter if it was determined that they were still uncomfortable with it.

After this, each participant was instructed that they would be presented with a problem consisting of a series of boxes, each containing a ball. Their goal was to reach a state in which each ball was out of its box. They were to interact with the puzzle by clicking on the balls. Each participant then solved the puzzle and a move record, as well as time for each move, was logged.

After completion, each participant had to carry out one more task according to their randomly assigned condition in a 2 x 2 design. Half of the participants were asked to solve the same problem again (Figure 2A) while the other half were asked to solve a horizontally inverted version of the same puzzle (Figure 2B). Also, half of the participants simultaneously performed the 2-back task from the training portion while the other half were given a control task. In the control condition no numbers were played, but the beep still occurred at the same random time interval. When they heard a beep, participants in the control condition were instructed to respond with any number between zero and ten, with the provision that they were not allowed to repeat the same number twice in a row. Again, a record of moves and time for each move were recorded.

Results

In this experiment we wanted to measure the effect of our manipulations on participants' problem solving ability. To this end we measured performance in two ways – time required to reach solution and number of moves made to reach solution. Unfortunately, the variability between participants was quite large on the first trial. The fastest participant made 39 moves to solve the puzzle in 36 seconds, while the slowest participant made 767 moves to solve the puzzle in just under 18 minutes. This made it necessary to be very careful to look for outliers before analyzing the data. Box-plots of the number of moves and time to solution for the first puzzle can be seen in Figures 3 and 4 respectively. Because of the extremely atypical solution times and moves of participants 4 and 23, they were excluded from the rest of the data analysis.

We also asked each participant to describe the strategy they used to solve the puzzle during their debriefing. Participants varied somewhat, some giving deeper and more effective strategies than others. For instance, many participants mentioned that there was a pattern to which balls needed to be out in order to release a new ball. A few participants mentioned that the first ball was moved more frequently than others or that the last ball was the hardest to move. However, none of the participants were able to give an explicit account of the puzzle's rule or describe the relationship of the problem states. That is, they were learning mostly implicitly.

Mean solution times and moves for each trial by condition can be found in Table 1. It was first crucial to see whether participants were showing transfer from the first to second trial, as we expected to see participants solve the puzzle significantly faster the second time through. A 2x2 repeated measures ANOVA showed a main effect of trial number on both moves ($F(1,34) = 13.48$, $p < .01$) and time ($F(1, 34) = 10.53$, $p < .01$). This indicates that participants were learning something about the puzzle during the course of their first solution.

Next we wanted to assess the effects of puzzle and memory load condition on transfer to the second trial. We found a significant effect of memory load condition on moves ($F(1, 34) = 8.66$, $p < .01$) and time ($F(1, 34) = 15.81$, $p < .001$). This indicates that participants were being hampered by memory load regardless of condition. We also found an interaction between puzzle and memory load for both moves ($F(1,34) = 4.28$, $p < .05$) and time ($F(1, 34) = 8.98$, $p < .02$). Thus, as hypothesized, memory load had a stronger interference effect in the inverted puzzle condition. Profile plots of these effects on moves and time can be found in Figures 5 and 6 respectively.

Finally, we were somewhat uncomfortable totally excluding the 2 outliers. We thus converted each of the performance measures to a log-scale and reran the 2x2 repeated measures ANOVA with the full dataset. As before we found a main effect of trial on both moves ($F(1,36) = 22.66$, $p < .001$) and time ($F(1,36) = 25.31$, $p < .001$). We also again found a significant effect of memory load for both moves ($F(1,36) = 6.56$, $p < .02$) and time ($F(1,36) = 10.53$, $p < .01$). Interestingly, we also found an effect of puzzle condition for time ($F(1,36) = 4.73$, $p < .05$), but not moves. Thus, by this possibly more sensitive measure, participants were slightly impaired by the inverted puzzle regardless of memory load condition. Most importantly, we again found the predicted interaction between puzzle and memory load for both moves ($F(1,36) = 6.54$, $p < .02$) and time ($F(1,36) = 4.12$, $p = .05$). This suggests that we were justified in removing the two outlying participants from the original analysis and testifies to the robustness of our key finding.

Overall, the main finding was the interaction between puzzle condition and memory load. As hypothesized, participants were relatively unimpaired by the novel, inverted problem unless their working memory was loaded. This expands our understanding of the role of working memory in implicit learning. Not only is working memory involved in the encoding of implicitly learned information, it is also involved on-line in transformation of that knowledge.

Discussion

The first clear finding in our study is a replication of Kotovsky and Simon's (1990) transfer results. Our participants demonstrated significant increases in problem solving performance on trial two relative to trial one. Thus, despite having no conscious awareness of the strategy they were using, participants learned something that made them more able to solve the Balls and Boxes puzzle on trial two. This confirms that we are studying an implicit learning effect, and sets the stage for further analysis.

Second, we found a significant effect of memory load on puzzle solution. Across conditions participants had a harder time transferring their knowledge when their working memory was interfered with by a 2-back task. This is contrary to the findings of Reber and Kotovsky (1997) who found a significant effect of working memory load only on trial one. We believe this to be a result of slight differences in methodology. Our participants were given only a minute of training on the 2-back task and performed both the training and trial one under single-task conditions. We suspect that all participants were mildly impaired on trial two by the novelty of the dual-task condition and their low levels of practice on the distracter task. Also, due to the large variability in our participants, it is possible we were studying a different population than Reber and Kotovsky (1997) were.

Most importantly, as hypothesized, we found a significant interaction between working memory load and puzzle condition (see Figures 7 and 8). Participants were much more impaired by the working memory interference task when they had to transfer their learning to the horizontally inverted puzzle than when they solved the same puzzle a second time. This evidence has implications for two important questions: what is the role of working memory in implicit learning, and what is the representation of implicitly learned knowledge?

Working memory has been found to play a part in the implicit learning process during the initial learning phase, but not during subsequent production phases (Reber & Kotovsky, 1997). This suggests that working memory is involved in the encoding of implicitly learned material, but not in its retrieval or translation to motor responses. In our experiment we have demonstrated that working memory does have an intermediate role in the production process. That is, while working memory may not be involved directly in retrieval, it appears to be used on-line in the translation of the encoded information into a form useful in a novel, but similar, problem. This is consistent with the Baddely (1992) conception of working memory as a scratchpad for real-time mental manipulations. Many of our participants mentioned in their debriefings that they explicitly noticed the horizontal inversion condition. It is unclear how much of the working memory involvement was unconscious and how much was a result of purposeful processing.

Nonetheless, we have shown that working memory has a role to play in implicit learning not only in encoding information, but also in translating it on-line.

The nature of the representations used for implicitly learned information is an open question. Seger (1994) presents a two level hierarchy of possibilities. At the first level, representation is either abstract or instantiated. Abstract representation means that general rules are being learned while instantiated representation means that specific stimuli are being encoded. These abstract rules may be tied to both the deep structure and surface structure of exemplars, or only to their deep structure. Instantiated representations may be verbatim (exact stimuli are encoded) or aggregate (partial stimuli are stored in combinations). We have shown that transfer to a novel problem with similar deep but different surface structure is possible, but that it is significantly impaired by working-memory interference. Since working-memory is used on-line, participants must be recoding information between retrieval and application. This suggests that participants are not representing verbatim instances of the Balls and Boxes problem as they would be significantly impaired in transfer regardless of memory load condition. This implication is confirmed by more explicit measurements of knowledge taken by Reber and Kotovsky (1997) in their original work. When shown a problem state and asked for the correct next move their participants performed at chance. We have also shown that if the representation is abstract, it must be tied to the surface structure of the problem. If this were not the case, participants would be equally impaired by working memory load in either problem condition. Based on our findings, and the result from Kotovsky and Simon (1990), we believe that participants are encoding abstract rules that are also tied to problem representations. This is why the rule is easier to apply to the original problem than a horizontally inverted isomorph. Only this theory can explain both the relative ease of transfer to the lo-info puzzle over the no-info puzzle, which have the same surface and deep structure, and the difficulty of transfer to the horizontally inverted puzzle.

Overall we have both replicated previous findings of transfer of implicitly learned information and shown something we believe to be quite novel: working memory is involved on-line in translating implicitly learned information to a form useful for a novel problem. However, we do have to address one concern: we have a great deal of variability in our data. This is suggestive of the use of different strategies by different participants and may mean we are averaging over distinct cases. It would be useful to repeat the same experiment with more participants and also to analyze participants with similar first solution times separately. Future directions for this work would seek to further elucidate the nature of the underlying mental representations. First, the problem we chose as our transfer condition problem is very similar to the training problem. It would be instructive to look at multiple transfer problems in order to further explore both the conditions under which transfer occurs and the conditions under which working-memory is involved. We would expect to see increased working memory need as the transfer problem differed more from the source problem. Second, all participants were given a verbal interference task. We would be interested to see how outcomes might change if working memory was loaded with a spatial task. We would be interested in an interaction between transfer problem type and interference type. That is, we might expect to see greater interference if the transfer problem differed spatially from the source problem and the working memory load task involved spatial manipulation than if it involved verbal manipulation. Finally, participants performed only two trials of the task.

It would be useful to have them perform several more trials to see if the working-memory interference effect persists and if we could recreate it in transfer back to the original training problem. Would participants who solved the horizontal isomorph on trial two be more impaired in transfer to the isomorph or to the original source problem? We would thus be able to get a better handle on the malleability of the implicitly learned representation.

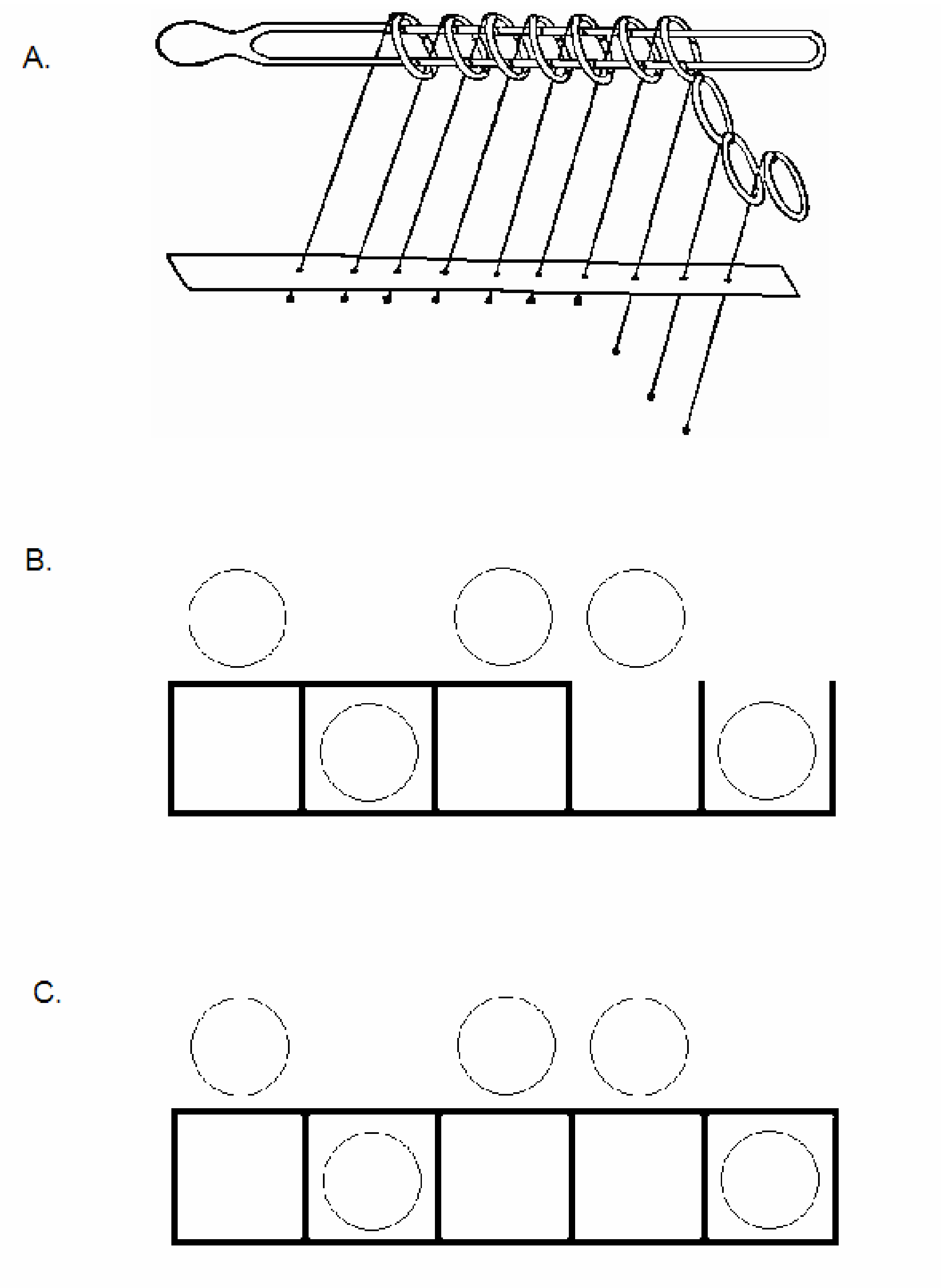

Figure 1: The various isomorphs from Kotovsky & Simon (1990)

A. The Chinese Ring Puzzle

B. The lo-info Balls and Boxes Puzzle

C. The no-info Balls and Boxes Puzzle

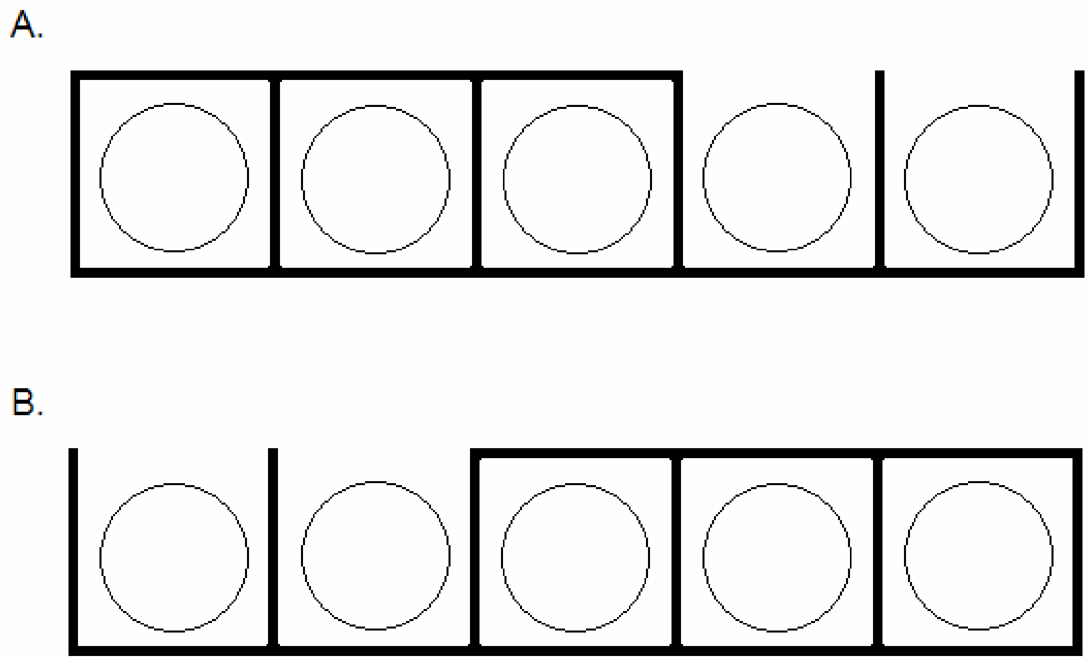

Figure 2: The starting state of the Balls and Boxes Puzzle
A. The first trial puzzle
B. The horizontally inverted puzzle

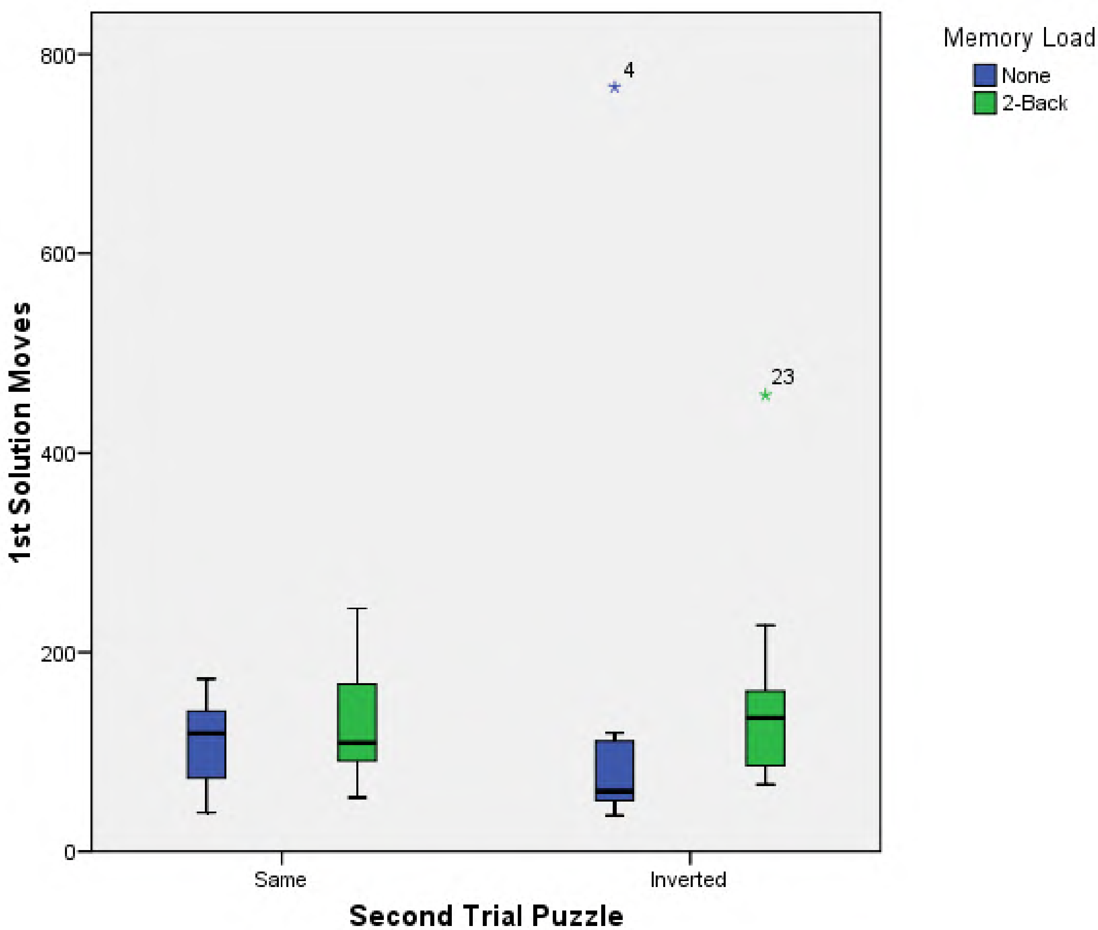

Figure 3: Number of moves for first solution of the Balls and Boxes Puzzle

* *Extreme Outlier*

o *Outlier*

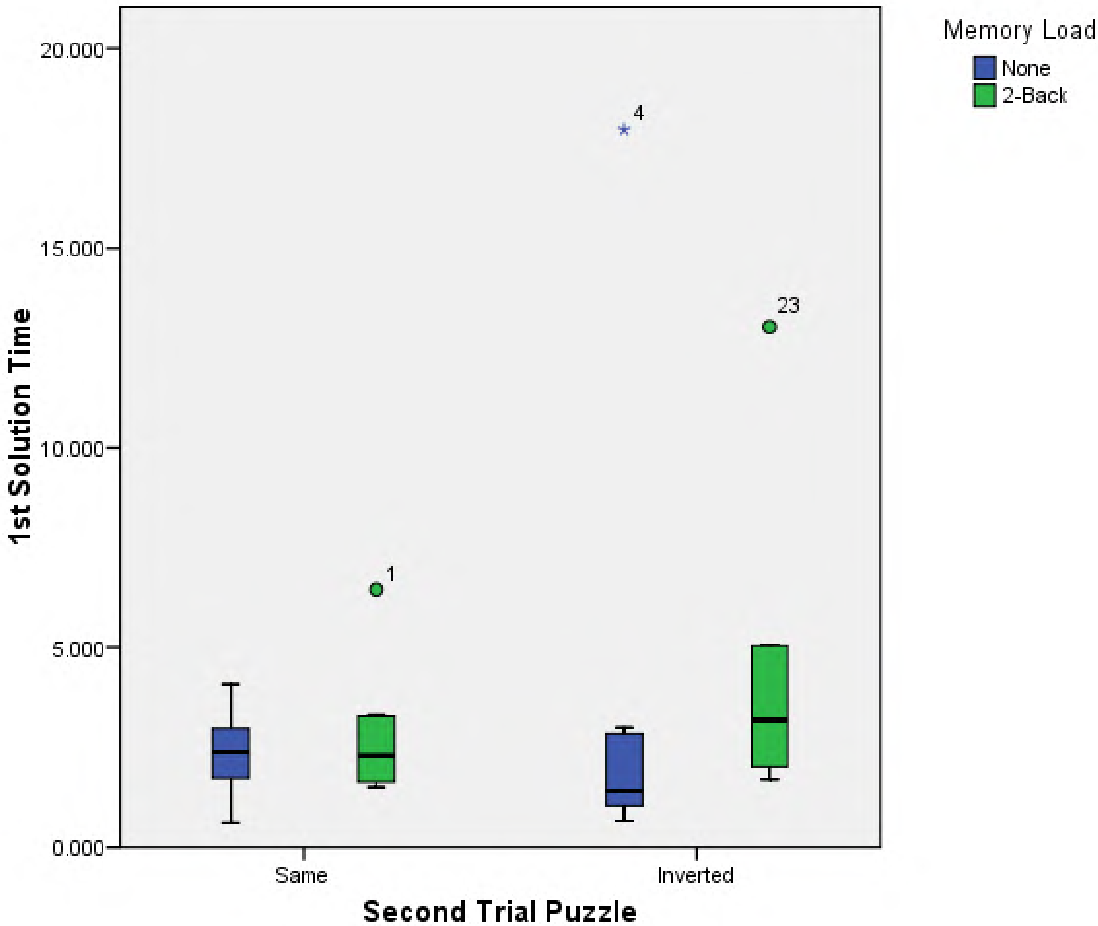

Figure 4: Time for First solution of the Balls and Boxes Puzzle

* *Extreme Outlier*
o *Outlier*

Table 1

Average moves and time to solution by condition for each trial

Puzzle	Memory Load	Trial 1		Trial 2	
Same	No load (n = 10)	109 moves (41.04)	2.39 min (1.05)	54.6 moves (55.32)	.95 min (.97)
	2-back (n = 10)	126 moves (60.27)	2.70 min (1.46)	55.80 moves (42.67)	1.29 min (1.22)
Inverted	No load (n = 9)	71.11 moves (32.08)	1.65 min (.91)	52.11 moves (29.11)	1.02 min (.72)
	2-back (n = 9)	128.33 moves (54.48)	3.20 min (1.34)	99.11 moves (62.6)	2.88 min (1.60)

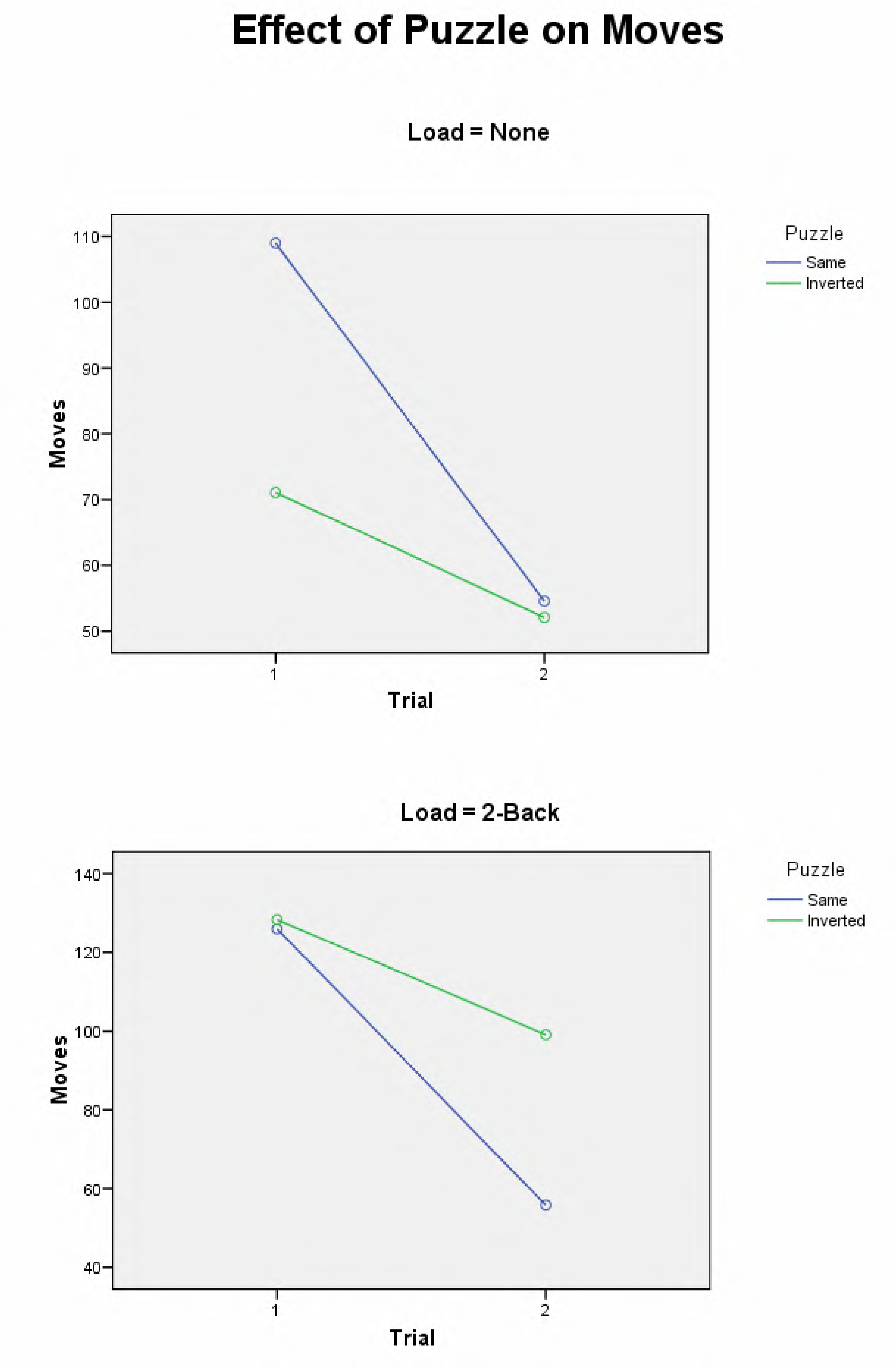

Figure 5: Interaction between puzzle and memory load – moves

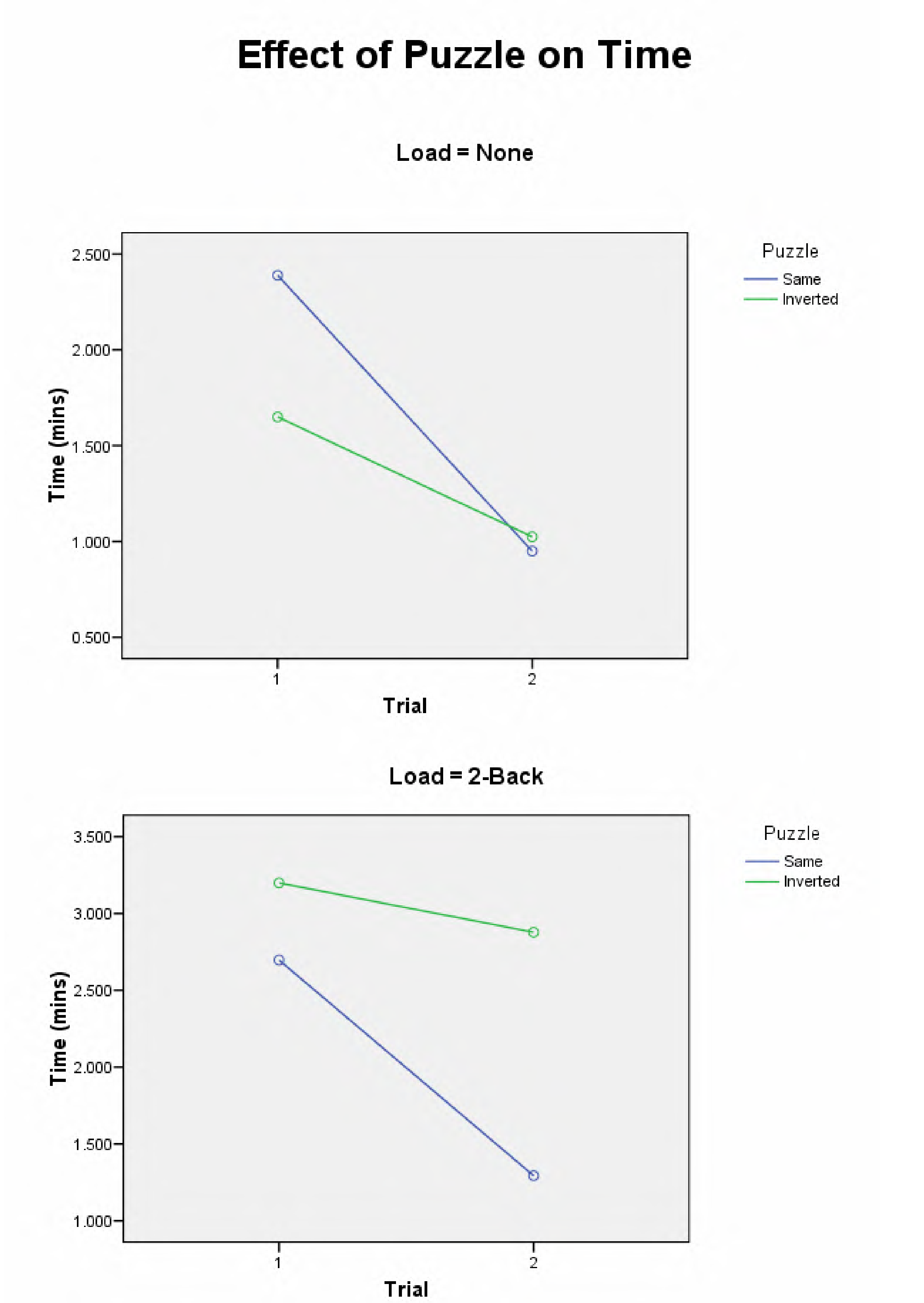

Figure 6: Interaction between puzzle and memory load – time

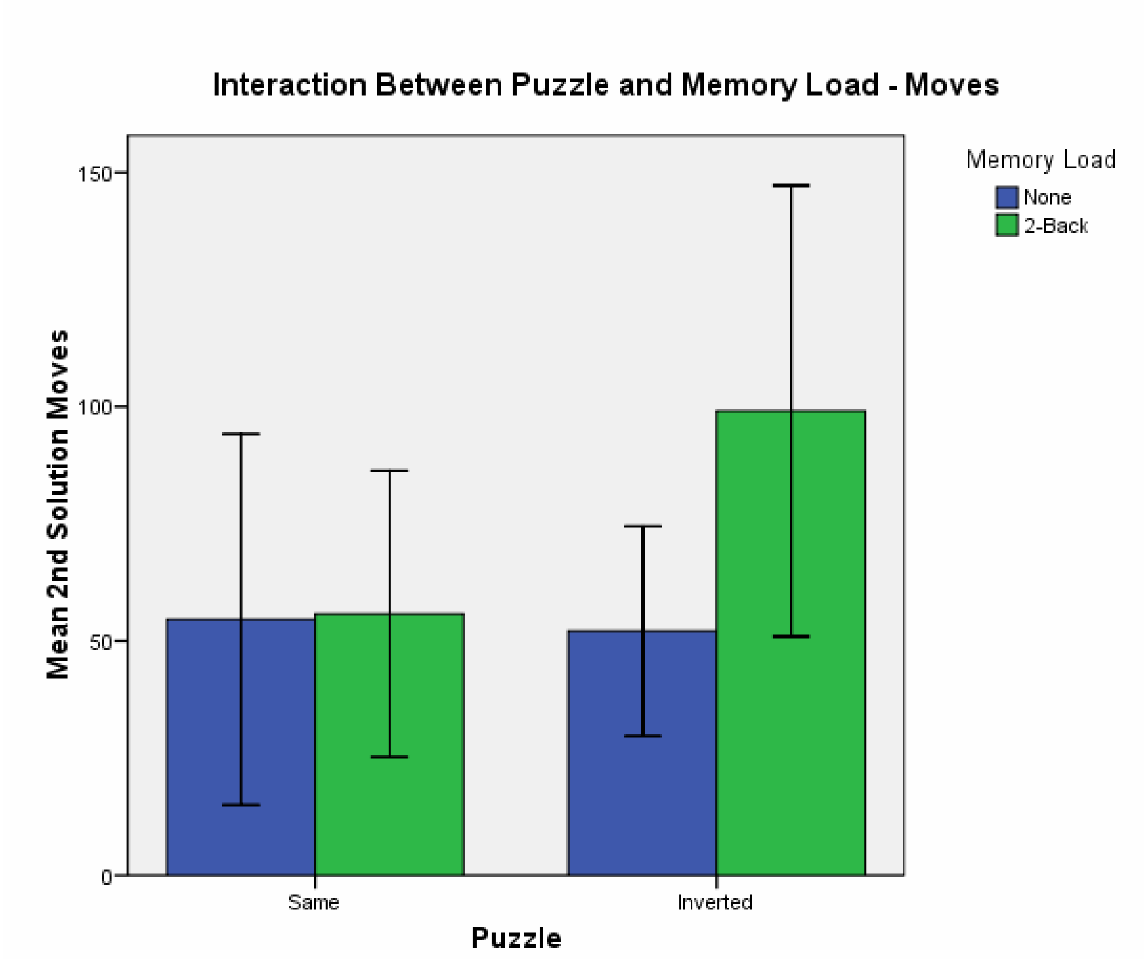

Figure 7: Interaction between puzzle and memory load for moves made on trial two

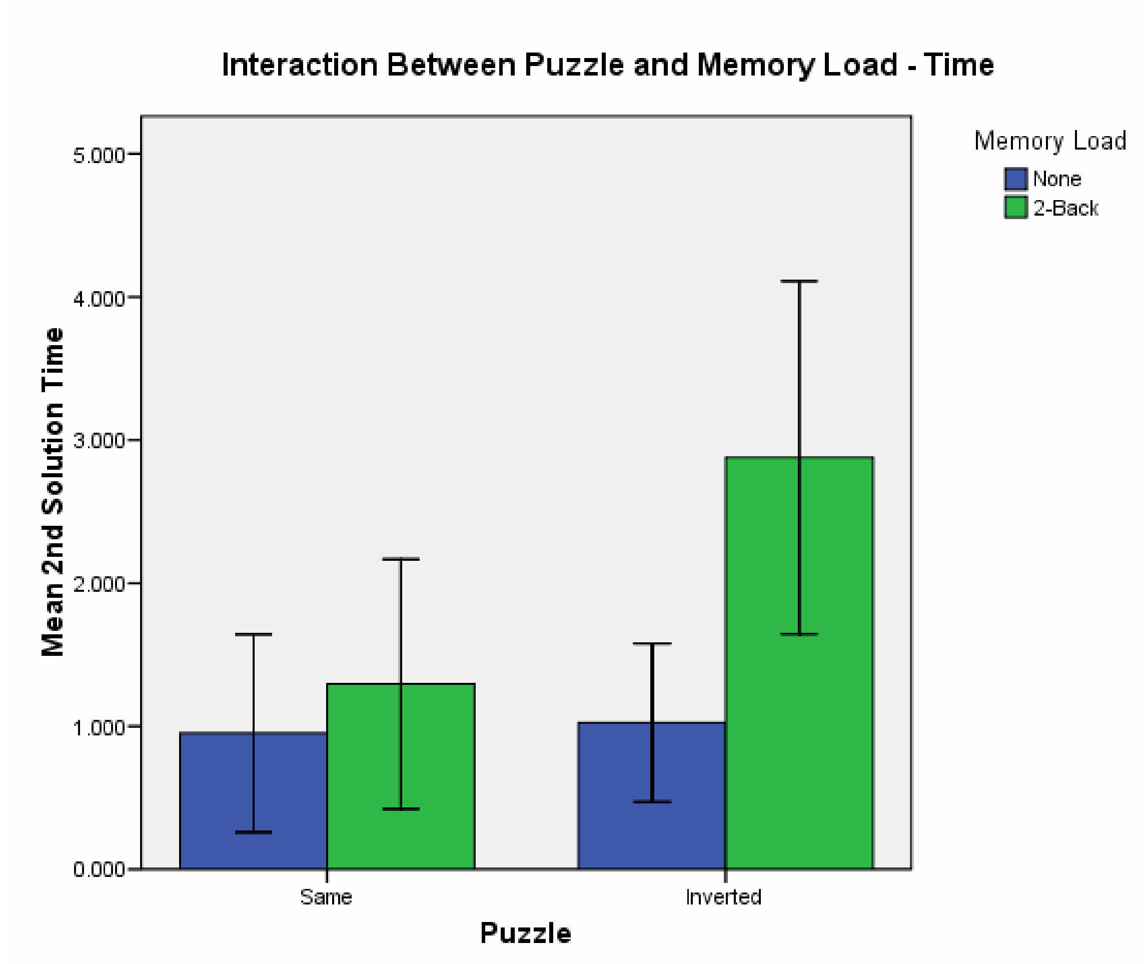

Figure 8: Interaction between puzzle and memory load for time to solution on trial two

References

Baddely, A. (1992). Working Memory: The interface between memory and cognition. *Journal of cognitive neuroscience*, *4*, 281-288.

Baddeley, A. D., & Hitch, G. (1974). *Working memory.* In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journals of Experimental Psychology*, *36A*, 209-231.

Dienes Z., Broadbent D., & Berry D. Implicit and explicit knowledge bases in artificial grammar learning. (1991). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 875-887.

Kotovsky, K., & Simon, H. A.(1990). Why are some problems really hard: Explorations in the problem space of difficulty; *Cognitive Psychology*, *22*, 143-183.

Lewicki, P. (1986). Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 135-146.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1083-1100.

Miller J. (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception and Psychophysics*, *41*, 419-434.

Musen, G., & Squire, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 789-798.

Poletiek, F. H. (2002). Implicit learning of a recursive rule in artificial grammar. *Acta Psychologica*, *111*, 323-325.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behaviors*, *6*, 855-863.

Reber, A. S., & Allen, R. (1978). Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition*, *6*, 189-221.

Reber, P., & Kotovsky, K. (1997). Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental Psychology, General*, *126*, 178-203.

Seger, C. A. (1994). Implicit Learning. *Psychological Bulletin*, *115*, 163-196.

Squire, L. R. & Frambach, M. (1990). Cognitive skill learning in amnesia. *Psycholbiology*, *18*, 109-117.

Acknowledgements

Infinite thank-yous are due to Dr. Ken Kotovsky whose gentle guidance and friendly nudging have helped me find my path in cognitive science. Without his advising this thesis would never have gotten off the ground. Special thanks also to Jarrod Moss who walked me through IRB negotiations and shared his lab space with me.