Working Memory Training and Semantic Structuring Improves Remembering Future Events, Not Past Events

Kim Merle Richter1,2, Claudia Mödden2, Paul Eling, PhD3, and Helmut Hildebrandt1,4

Abstract
Objectives. Memory training in combination with practice in semantic structuring and word fluency has been shown to improve memory performance. This study investigated the efficacy of a working memory training combined with exercises in semantic structuring and word fluency and examined whether training effects generalize to other cognitive tasks. Methods. In this double-blind randomized control study, 36 patients with memory impairments following brain damage were allocated to either the experimental or the active control condition, with both groups receiving 9 hours of therapy. The experimental group received a computer-based working memory training and exercises in word fluency and semantic structuring. The control group received the standard memory therapy provided in the rehabilitation center. Patients were tested on a neuropsychological test battery before and after therapy, resulting in composite scores for working memory; immediate, delayed, and prospective memory; word fluency; and attention. Results. The experimental group improved significantly in working memory and word fluency. The training effects also generalized to prospective memory tasks. No specific effect on episodic memory could be demonstrated. Conclusion. Combined treatment of working memory training with exercises in semantic structuring is an effective method for cognitive rehabilitation of organic memory impairment.

Keywords
rehabilitation of memory impairments, randomized controlled study, stroke, long-term memory, prospective memory, working memory, word fluency

Introduction
The frequency of memory deficits following brain lesions such as toxic or metabolic encephalopathy, temporal lobe epilepsy, and stroke is estimated to lie between 40% and 60%.1 Previous studies on memory rehabilitation have shown that a partial recovery from amnesia may be possible and that a reorganization of memory functions might be achieved.2-4 In their meta-analysis, Cicerone et al,2 for example, recommend strategy training for mild memory impairments including the use of internalized strategies as practice standards. Most previous studies, however, focused on patients with traumatic brain injury (TBI), and there is still a need for evidence-based therapy methods, especially for other clinical populations like stroke patients.5,6

The effectiveness of a therapy is determined by various factors such as the severity of impairment, the type of task used to assess its effectiveness, as well as the relevance of the therapy for daily life situations. There is consensus that patients with severe amnesia should be taught to use external memory aid devices. In patients with limited memory impairments, the prospects to improve their performance are better and therefore the use of internal memory strategies may be helpful. In the study of Berg et al,7 patients with closed head-injury received either a strategy training or 18 hours of drill and practice tasks; a third group received no treatment (passive control group). The results showed that only the strategy training yielded a significant improvement on objective memory tests. The effects of the training were particularly clear at the 4-month follow-up assessment, indicating a strong positive effect of the strategy training.

An important issue in the literature concerns the generalization of cognitive rehabilitation therapy to other tasks and to daily life situations. It has been shown that the transfer of
training effects to other tasks than those trained is often limited. A central objective of recent approaches is therefore to reduce hyper-specificity of the training and to enhance effortful processing in order to facilitate generalization and flexible learning. Hyper-specificity of training can be reduced by increasing stimulus variation; effortful processing is stimulated through strategy components like verbal elaboration or semantic structuring. Focusing on simple internal strategies can also be useful in fostering generalization of treatment effects. Hildebrandt et al. showed that training of semantic structuring and word fluency, in combination with exercises in which patients have to learn and reproduce word lists, can be effective in improving verbal memory in general (word list learning, reproduction of text). In another study involving patients with multiple sclerosis, a training of executive functions was effective in increasing verbal memory performance, which also showed generalization of treatment-induced improvement between different cognitive functions.

In this study, working memory training is combined with exercises in semantic structuring and word fluency. Working memory is essential for the semantic processing of sequentially presented pieces of information and is involved in encoding and recall. The efficacy of certain memorizing strategies, such as structuring information, depends on intact working memory performance. Available working memory resources can then be used for the application of a proper strategy and therefore for a deeper encoding. Working memory is assumed to fulfill 2 kinds of functions: active maintenance (keeping information available) and active organization. It might therefore be assumed that especially an improvement in these components can contribute to an improved memory performance. On the other hand, Westerberg et al. studied the efficacy of a working memory training on stroke survivors and found an improvement in working memory and attentional performance, but no transfer to other cognitive functions like long-term memory.

Westerberg et al. focused exclusively on working memory and did not include exercises in semantic structuring in their treatment. Training in semantic structuring should improve the ability to detect and construct semantic relations. Our assumption was that combining exercises in semantic structuring and word fluency with working memory training might facilitate the processing of information, by enabling participants to better hold information online. Moreover, it should help construct semantic relations between the to be remembered bits of information (eg, semantic relations for better subsequent retrieval of information), thereby improving performance in episodic memory tasks. With regard to word fluency, previous studies have shown an association between improved recall and word fluency, suggesting a change in frontal and temporal memory functions. We assumed that such an effect might be obtained by directly training word fluency. Furthermore, unlike most studies before, we did not include any episodic memory tasks in our treatment, unlike, for example, Berg et al. and Hildebrandt et al. Therefore, any improvement in memory performance cannot be ascribed to a hyper-specific task effect and also not to similarity between training and evaluation procedures. In other words, because our training did not involve episodic memory tasks, in the case of a significant improvement of memory performance, generalization to other tasks and to daily life situations is much more plausible. We therefore assumed that (a) patients who have suffered a brain lesion should benefit from such a training in working memory function and that (b) the combination of working memory training and exercises in word fluency and semantic structuring leads to an improvement in episodic memory performance (encoding and delayed recall of information; prospective memory performance).

### Methods

#### Patients

The study was conducted at a rehabilitation center, which is specialized in patients in rehabilitation phase D, that is, the treatment aimed at further reducing impairments for patients who have a Barthel Index of at least 70 and are independent in daily life activities. Inclusion criteria were a brain lesion and an outcome score of at least 1 standard deviation below the age and education corrected average on the California Verbal Learning Test (CVLT) on either the first or fifth learning trial, on the short delay free recall or long delay free recall score. Mean time since brain lesion was 60 days. Exclusion criteria were a severe amnesia, impaired visual functions, and insufficient knowledge of the German language.

With regard to sample size we have oriented at a previous study, using similar neuropsychological tests and showing that with group sizes of about 15 patients significant treatment effects can be found.

Following an elaborate diagnostic examination, patients were allocated randomly, using a traditional lottery procedure with tickets, to either the experimental or the control group (Figure 1). Allocation was done blinded with respect to the results of the neuropsychological assessment. The study was approved by the ethics committee at the University of Oldenburg, and participants provided a written consent before participating in the study.

#### Procedure

Every patient was tested before and after the therapy. To avoid learning effects, parallel versions of the individual tests were used; order of the test versions was counterbalanced. The test battery was presented in a predetermined order in 2 sessions, lasting 1.5 hours in total. The first part
of the initial testing, in particular the CVLT, was performed by one of the psychologists from the rehabilitation center. Then patients were allocated to either the intervention or control group. Subsequently, the assessment was completed by 2 psychologists blind to group membership of the participant.

The patients of both groups received a therapy of at least 9 hours.

**Experimental Therapy**

In the experimental group, patients received 9 one-to-one therapy sessions. In every session, the patient was trained for half an hour on computer exercises with the new WOME software package17 (Hasomed GmbH). The program has been developed to enhance working memory performance. The patient was trained on a card game with different levels of difficulty; degree of difficulty can be adapted to the performance level of the patient. The patient’s task, for instance, was to remember a short series of cards and reproduce it in reverse order, or he had to memorize only the cards of a certain suit. In total there are 28 levels of difficulty. The numbers of points obtainable per level as well as the display time of the cards are adjustable. During the training, the patients constantly received feedback on their performance.

Subsequently, in each session various letter and category word fluency exercises and exercises in semantic structuring (15 minutes) were performed with the patients. One of the exercises was to create categories: the patient was presented with a list of 8 words that could be divided into 2 groups; for each of the 2 groups a generic term or heading had to be provided.

In another task, for instance, the patient had to produce as many examples as possible of a certain category. The exercises started with “easy” categories for which many examples can be found, such as “professions.” The exercises with smaller categories like “crafts” or “creative professions” were considered as more difficult and were presented at a later stage.

Moreover, some exercises were carried out aimed at training word fluency as well as working memory. The patient was requested to find for every letter of the alphabet an exemplar of a certain category like “animals.” Additionally, an exercise with a partner (always the therapist) was conducted, in which a word was produced by one participant and another word had to be produced by the second participant, beginning with the penultimate letter of the last mentioned word.

The number of exercises in a session was determined by the performance of a patient and the time needed to complete these exercises. For well-performing patients, additional exercises were available.

**Standard Therapy**

The patients of the control group received the standard memory treatment, which is usually provided at the rehabilitation center. The standard memory treatment consisted of three 1-hour sessions of group therapy per week.

Different exercises were presented aimed at improving memory performance. Verbal material was offered...
in a variety of modes such as lists of words, short stories, current news, or information about group members. The focus was on rehearsal of learning strategies and free and cued recall. Moreover, in every session also information presented in previous sessions had to be recalled, prompting patients to use the spaced retrieval method for better encoding of information.

The standard treatment took place on fixed days of the week. It was not possible to provide 9 hours of memory therapy to all patients due to the short duration of stay in the rehabilitation center or to other obligatory rehabilitation appointments. In that case, one-to-one sessions were offered with attention therapy as a substitute. During this computer-supported attention therapy, a patient had, for instance, to stop a clock at a specified time or control the speed of a virtual train. The exercises performed in the attention therapy were intended to enhance attention rather than memory encoding strategies.

**Evaluation of Therapy Effects**

**Composite Scores.** The patients were assessed on different cognitive functions by means of a neuropsychological test battery, and the scores were combined into 6 composite scores. The following composite scores (CS) were used to evaluate the treatment effects, all encompassing 3 different tests (except for attention).

*Working memory* was based on the sum of the well-known digit span forward and backward (from Wechsler Memory Scale\(^ {18,19} \)), as well as the results (hits minus false positives) of a 2-back task with numbers between 11 and 99.\(^ {20} \) The 2-back task is commonly used as a measure of executive components of the working memory since it requires online monitoring, updating, and active manipulation.\(^ {21,22} \)

*Immediate memory* was based on the scores of the CVLT. Learning trials 1 to 5 were added to the score obtained on the Rivermead Behavioural Memory Test (RBMT subtest 6 [Story], immediate recall\(^ {23} \)). The CVLT is often used in neuropsychological assessment and it predicts success in occupational rehabilitation.\(^ {24} \)

*Delayed memory* performance was determined by adding the scores of short-term free recall and long-term free recall on the CVLT and the scores of RBMT 6 delayed recall.

*Prospective memory* performance was established by adding the results of RBMT items 1 + 2 (Names), 3 (Belongings), and 4 (Appointment). The RBMT is one of very few tests with norms for prospective memory performance, and it has been shown to be ecologically valid in predicting everyday memory problems.\(^ {25} \)

Word fluency consisted of the sum of the 3 subtests of the Regensburger Wortflüssigkeitstest (RWT\(^ {26} \)), comprising semantic, phonologic word fluency, and the continuous shift between 2 initial letters.

For *Attention*, the 2 subtests of the alertness test Testbatterie für Aufmerksamkeit (TAP\(^ {27} \)) were combined into one measure by taking the average of the reaction times. These tasks were used because alertness is a core function of attention contributing to other dimensions of attentional performance.\(^ {28} \)

Treatment effects were also evaluated using the Cognitive Failures Questionnaire (CFQ\(^ {29} \)). The CFQ is a self-rating scale used to rank cognitive failures in daily life, like attention lapses and memory problems. The CFQ consists of 25 items, and for each item the participant is asked to indicate how often he has noticed problems, scores ranging from 0 (never) to 4 (very often). The final score is the sum for all items (with a maximum of 100).\(^ {15} \)

**Statistical Analysis**

To evaluate the therapy effects, an ANOVA for repeated measures was used. Time (before and after therapy) represented the within-subjects factor, and Group (experimental vs. control) represented the between-subjects factor. Significant interaction effects were analyzed with post hoc \(t\) tests for dependent measurements comparing the mean scores on pre- and posttherapy assessments. Post hoc \(t\) tests for dependent measurements were conducted to establish whether there were significant differences between the 2 groups with regard to the mean of the 6 CSs on the pretherapy assessment.

**Results**

The patients in the experimental and control groups did not differ in age, education, duration of rehabilitation, or time passed since brain lesion (Table 1).

Table 2 gives an overview of the treatment results for both groups.

**Working Memory**

For the composite score of the working memory performance, a significant main effect for Time, \(F(1, 34) = 25.509, P < .000\), and a marginally significant Time by Group interaction was found, \(F(1, 34) = 3.970, P = .054\). There was no main effect for Group. Post hoc \(t\) tests for dependent measurement, comparing the mean scores at pre- and posttherapy testing, indicated that both groups improved after treatment. However, as shown by the significant interaction effect and the mean scores, the improvement of the experimental group was significantly larger than that of the control group (Figure 2).
Immediate and Delayed Verbal Memory

The composite scores of immediate and delayed memory revealed no significant differences between the groups with regard to improvements in verbal memory. There was also no difference between the groups before the treatment. For immediate memory there was a significant main effect for Time, $F(1, 34) = 70.233, P < .001$, but not for Group, and there was also no significant Time by Group interaction. Similarly, for delayed memory there was a significant main effect for Time, $F(1, 34) = 40.440, P < .001$, but no main effect for Group and no significant Time by Group interaction. The composite scores of immediate and delayed memory therefore do not reveal a differential improvement for the groups on episodic memory.

Prospective Memory

There was no main effect for Group, but there was a significant main effect for Time, $F(1, 34) = 25.509, P < .001$, and a significant Time by Group interaction, $F(1, 34) = 6.377, P = .016$. A post hoc $t$ test showed that the groups did not differ before the treatment, and $t$ tests for dependent measures indicated a significant improvement for both groups between pre- and posttesting. As indicated by the significant interaction effect, however, the improvement of the experimental group was larger than that of the control group (Figure 3).

Word Fluency

There was no significant main effect for Group, but we did find a significant main effect for Time, $F(1, 34) = 30.645, P < .001$, and a significant Time by Group interaction, $F(1, 34) = 10.200, P = .003$. A post hoc $t$ test showed that there

### Table 1. Patient Characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Experimental Group</th>
<th>Active Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etiology and number of patients</td>
<td>Stroke = 12, TBI = 1, others = 5</td>
<td>Stroke = 13, TBI = 3, others = 2</td>
</tr>
<tr>
<td>Sex (male–female)</td>
<td>13:5</td>
<td>14:4</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>10.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Age in years*</td>
<td>50.0 ± 8.2</td>
<td>50.8 ± 10.4</td>
</tr>
<tr>
<td>Time since lesion*</td>
<td>64.4 ± 90.0</td>
<td>56.6 ± 75.4</td>
</tr>
</tbody>
</table>

Abbreviation: TBI, traumatic brain injury.

*Mean ± standard deviation.

### Table 2. Effects of the Therapeutic Interventions on the Different Composite Scores (CS).

<table>
<thead>
<tr>
<th></th>
<th>Experimental Group</th>
<th>Active Control Group</th>
<th>Pre Post</th>
<th>Pre Post</th>
<th>$F$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS Word fluency</td>
<td>36.3 ± 8.5</td>
<td>50.0 ± 10.6*</td>
<td>38.4 ± 11.4</td>
<td>42.1 ± 8.33*</td>
<td>10.200</td>
<td>.003</td>
</tr>
<tr>
<td>CS Working memory</td>
<td>28.9 ± 8.6</td>
<td>36.6 ± 7.1*</td>
<td>29.6 ± 9.7</td>
<td>34.1 ± 9.1*</td>
<td>3.970</td>
<td>.054</td>
</tr>
<tr>
<td>CS Prospective memory</td>
<td>6.9 ± 2.0</td>
<td>9.1 ± 1.5*</td>
<td>7.7 ± 1.9</td>
<td>8.4 ± 1.8*</td>
<td>6.377</td>
<td>.016</td>
</tr>
<tr>
<td>CS Immediate memory</td>
<td>17.1 ± 4.2</td>
<td>24.7 ± 4.6*</td>
<td>18.9 ± 3.8</td>
<td>25.1 ± 4.9*</td>
<td>0.799</td>
<td>.378</td>
</tr>
<tr>
<td>CS Delayed memory</td>
<td>15.6 ± 5.9</td>
<td>23.4 ± 6.3*</td>
<td>17.8 ± 6.7</td>
<td>24.7 ± 9.1*</td>
<td>0.130</td>
<td>.721</td>
</tr>
<tr>
<td>CS Attention (milliseconds)</td>
<td>288 ± 100</td>
<td>242 ± 34</td>
<td>269 ± 74</td>
<td>256 ± 63</td>
<td>1.550</td>
<td>.223</td>
</tr>
</tbody>
</table>

*Mean ± standard deviation.

$F$ and $P$ values represent the interaction effects between Group and Treatment for the ANOVA for repeated measurement.

*Significant ($P < .05$) intragroup improvement (dependent $t$ tests).

![Composite score for Working memory](image)

**Figure 2.** Performance on working memory before and after treatment.

Statistical evaluation showed that both groups improved but the experimental group improved significantly more (error bars indicate standard deviations of mean).
was no difference between the groups before the treatment, and \( t \) tests for dependent measures indicated that both groups improved after treatment. As shown by the significant interaction effect and the mean scores, the improvement of the experimental group was larger than that of the active control group.

Attention

For the composite score of attention, there was a significant effect for Time, \( F(1, 29) = 4.715, P = .038 \), but not for Group and the Time by Group interaction was also not significant. A post hoc \( t \) test showed that the groups did not differ before the treatment, and \( t \) tests for dependent measures showed no significant improvement in attention for the 2 groups.

Cognitive Failures Questionnaire

There was no main effect for Group or Time and no significant Time by Group interaction for the CFQ. Post hoc \( t \) tests revealed no significant improvements after treatment for the 2 groups.

Discussion

In this randomized controlled study, we analyzed the effects of a working memory training combined with exercises in semantic structuring and word fluency. The results show a significant improvement in working memory, prospective memory, and word fluency compared to an active control group, receiving a standard group therapy for memory disorders, whereas episodic memory and attention improved to a similar degree for both groups.

Word fluency and working memory performance has improved following treatment in the experimental condition as indicated by the significant interaction effects. Several previous investigations have shown that word fluency may play a role in improving memory performance in patients with a stroke or other organic brain disorders.1,10 Working memory, on the other hand, is important in various situations, especially those that require the simultaneous holding and manipulation of information or complex reasoning.30 Working memory performance also influences learning capabilities,31 seems to predict the ability to control thoughts and behavior,32 and is of crucial importance for the rehabilitation of other cognitive functions.15

The assumption underlying this study was that an enhanced working memory performance in combination with improved skills in word fluency and semantic structuring might lead to an improvement in general memory performance. We could not demonstrate a specific effect of our training on episodic memory. However, the training did yield significant improvements in prospective memory.

The fact that there was no significant difference in improvement between the 2 groups on immediate and delayed memory performance does not necessarily imply that the combined training had no effect at all on episodic memory. Both groups did improve in immediate and delayed memory as indicated by the significant main effect for Time. It only shows that our training was not better in improving episodic memory than a standard group memory therapy. Considering that the control group received a standard memory therapy, memory improvement in this group could be expected. We need to take into consideration, however, that the improved performance in episodic memory could be due to spontaneous recovery. This seems plausible considering the overall improved test performances in both groups and the relatively short period since onset (~60 days). Unfortunately, our design does not allow distinguishing between treatment induced and spontaneous recovery induced improvement of long-term memory.

With regard to prospective memory, the results are more convincing. Improvements in prospective memory might be expected for the experimental group, considering that the group has improved in working memory. Prospective memory presumably represents a multitask situation: it refers to the ability to carry out an intention after a certain amount of time while being involved in other ongoing activities.33 Involvement of working memory appears likely. Obviously, the patients managed to use their improved skills like the enhanced working memory capacity to execute the prospective memory tasks.

The clinical importance of successfully performing prospective memory tasks is obvious. The ability to keep an intention in mind and constantly keep an eye on the environment, while being involved in other activities, is crucial.
for accomplishing various everyday tasks. Even though prospective memory has not been studied very extensively until now, it might be of even greater importance than retrospective memory for daily coping and the ability to live independently.\textsuperscript{34} It has been shown that prospective memory disorders are related to a reduced quality of life, and the increasing number of studies on prospective memory suggests that it will probably become a central topic for future investigations.\textsuperscript{33}

The results of this study may also be relevant for the question whether prospective memory and episodic memory depend on the same underlying brain structures. Schacter et al\textsuperscript{35} found evidence that there is a considerable overlap between brain regions that are active when remembering the past and those that are active when imagining the future. According to the constructive episodic simulation hypothesis, both operations make use of information stored in episodic memory and depend on similar underlying processes. Since we found an effect of our therapeutic interventions on prospective but not on episodic memory, our results may indicate a greater independence of the episodic and prospective memory system than suggested by the Schacter et al\textsuperscript{35} findings. In this study, selective improvement in prospective memory may have been due to the nature of the training tasks, which relied on executive components like working memory and reorganization of semantic relations and might have limited generalization to episodic memory, which was the main goal of our study.

Attention performance has improved in both groups to a similar extent as indicated by the significant main effect for Time and the absence of the Group by Time interaction. This improvement may be due to spontaneous recovery; alternatively, it may be a global treatment effect. The fact that both groups improved in attention performance, however, rules out the possibility that this improvement in attention is responsible for the significant improvements in working memory and prospective memory tasks that were found in our experimental group.

Considering the fact that our experimental group shows significant improvements in working memory performance, the question arises why these improvements are not reflected in the scores on the Cognitive Failures Questionnaire. In our study, the validity of this test, in particular for the scores collected after therapy, might have been limited by the short time span between pre- and posttesting and by the fact that the patients did not return to daily life during this period but stayed in the rehabilitation center. Additionally, several patients commented that after having answered the questionnaire at the pretherapy assessment, they paid more attention to whether certain problems or situations occurred, like dropping objects. This may have resulted in reporting more failures and may explain why some patients even had higher scores at posttesting.

In evaluating our study there are some weaknesses to be considered. First, a relatively small number of patients was included (18 per group). Besides, with 9 hours of therapy the intensity of our treatment was relatively low. A more intensive treatment, as it is provided in rehabilitation studies for other domains (e.g., motor rehabilitation with about 40 hours of treatment within a few weeks\textsuperscript{4}), might have resulted in even larger therapy effects. Unfortunately, the short stay in the rehabilitation center did not allow for a more intensive treatment. Additionally, for ethical reasons we abstained from a placebo treatment. However, all patients received specific treatment and were not informed about the goals of the study. About two thirds of the patients in both groups suffered from stroke. This might make generalization to patient groups with other etiologies difficult.

Finally, a follow-up assessment would have been interesting to establish whether our treatment had long-lasting effects.

However, it should be emphasized that our study meets the most important methodological criteria to guarantee high validity of the results: The study was randomized (with concealment), controlled, double-blinded, and included an active control group.

To conclude, this study demonstrated the efficacy of the combined treatment of working memory performance and word fluency to improve prospective memory. Since our training did not involve prospective memory tasks, improvement on these tasks contradicts the assumption that only hyper-specific trainings effects can be achieved in cognitive rehabilitation due to a focus on specific cognitive functions. Similar generalization effects have been demonstrated in other studies focusing specifically on working memory.\textsuperscript{15,36,37} These results are encouraging, considering the fact that working memory performance as well as prospective memory performance can significantly influence coping in everyday life. They are also encouraging because they open a window for internal compensation in addition to external compensation of memory deficits after organic brain dysfunctions. Additionally, our results are striking as they show that already within 9 hours of therapy significant improvements in different cognitive functions can be yielded.

Declaration of Conflicting Interests
The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Kim Merle Richter received funding of Hasomed, Inc (Magdeburg), which sells the RehaCom cognitive training battery.

Funding
Kim Merle Richter received funding of Hasomed, Inc (Magdeburg), which sells the RehaCom cognitive training battery.
References


