Long-term memory disorders: measurement and modeling
Meeter, M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 9
PARAHIPOPCAMPAL CONTRIBUTIONS TO MEMORY AND SCHIZOPHRENIA

Episodic memory impairment is one of the most reliable and well-characterized neuropsychological findings in schizophrenia. It is argued here that episodic memory impairment is a trait characteristic of the disorder and likely based on mediotemporal dysfunction. However, it is not clear from the literature how medial temporal lobe (MTL) abnormalities may produce the pattern of deficits observed in schizophrenia. Nor is it understood how MTL circuits regulate episodic memory function in general. The present study attempts to shed light on these issues using a computational approach. A MTL model was developed that could simulate many episodic memory tasks, including free recall, cued recall, recognition and source monitoring. In this model, schizophrenia neuropathology was simulated by decreasing the density of connections within the parahippocampal cortex and to the hippocampus. Findings suggest that connectivity within the parahippocampal region subserves integration of cortical inputs into an episodic trace, while the projections to the hippocampus proper contribute to autoassociative storage and retrieval. The reduction of both levels of connectivity thus leads to a mild encoding deficit and a superimposed difficulty in retrieving information, particularly when only contextual cues are provided. This resulted in reproduction of many of the memory deficits observed in schizophrenia, including mild impairments in recognition and a more pronounced deficit in free recall. These findings highlight the importance of parahippocampal neuropathology in schizophrenia, showing that reduced connectivity in this region may underlie the episodic memory problems associated with the disorder.

9.1 INTRODUCTION

Schizophrenia was described for the first time more than a century ago. Though since then much progress has been made, its underlying pathology remains as contentious as ever. One cause of this is its complex clinical manifestation. Characteristic symptoms of the disorder are positive symptoms, such as hallucinations, delusions, disturbances of thought and speech, bizarre behavior, and negative symptoms, such as flattened or inappropriate affect, apathy and disorders of movement. Theories of schizophrenia have up to date never succeeded in explaining the whole palette of symptoms of schizophrenia. Many theories centered around dopamine abnormalities, for example, have concentrated on the positive symptoms, while showing relative neglect for negative symptoms. Indeed, these are generally unresponsive to medication targeting the dopaminergic system (Talamini, 2000).

It may well be that the quest for a single underlying deficit schizophrenia is misguided. Not only is schizophrenia a complex disorder, there is also strong heterogeneity in the symptomatology of individual patients, which also tends to shift over the course of Work on this chapter was done in collaboration with Lucia Talamini, Jaap Murre, Brita Elvevåg, and Terry Goldberg. It is a draft of Talamini, Meeter, Murre, Elvevåg, & Goldberg (in prep.).
illness. Similar variability can be observed at the level of brain pathology, which differs quantitatively and qualitatively over patients. Insight into the relation between brain dysfunction and behavior in this disease may best be achieved by focusing on single symptoms or symptom clusters (Bell, Lysaker, Beam Goulet, Milstein, & Lindenmayer, 1994), rather than approaching the syndrome as a whole. Indeed, while neither brain pathology, cognitive profile, nor symptomatology are consistent over patients, it may be possible to relate specific dysfunctions to particular neuropathology. Following this approach, working memory deficits and predominant negative symptomatology have been related to dysfunction of the dorsolateral prefrontal area (Tamminga et al., 1992), while severity of thought disorder and auditory hallucinations have been correlated with decreased superior temporal gyrus size (Barta, Pearlson, Powers, Richards, & Tune, 1990; Marsh et al., 1997; M. E. Shenton et al., 1992).

Memory deficits in schizophrenia

One set of symptoms that seems very appropriate for such an approach are the episodic memory impairments frequently associated with schizophrenia. Detailed characterization of the deficits has revealed a coherent profile, with variable degrees of impairment on some memory paradigms and spared function on others. Moreover, deficits in this domain are highly prevalent in the schizophrenic patient group (Aleman, Hijnman, De Haan, & Kahn, 1999; Rushe, Woodruff, Murray, & Morris, 1999), and have to large extent been quantified. Data from the memory domain can thus effectively constrain a model of underlying brain dysfunction. After reviewing these findings, we will argue that episodic memory deficits are likely based on mediotemporal dysfunction. We will then present a computational model that suggests how medial temporal lobe abnormalities associated with the disorder may produce the particular set of episodic memory deficits observed in schizophrenia. Finally, we will discuss how the specific circuit abnormalities suggested by the model might affect other aspects of cognition in patients with schizophrenia.

While cognitive impairment in schizophrenia is broadly based (Bilder, 1996; Elvevåg & Goldberg, ; Pantelis, Nelson, & Barnes, 1996), studies that have compared neuropsychological measures have found the largest impairments in episodic memory (Gold, Randolph, Carpenter, Goldberg, & Weinberger, 1992; Heinrichs & Zakzanis, 1998; McKenna et al., 1990; Saykin et al., 1991; Saykin et al., 1994). The memory deficit profile in schizophrenia includes moderate to severe impairments in free recall, a lesser deficit in cued recall and a small but significant deficit in recognition (Aleman et al., 1999; Gold et al., 1992; Goldberg, Weinberger, Pliskin, Berman, & Podd, 1989; Paulsen et al., 1995). Importantly, the observed memory deficits are not dependent on delays interposed between encoding and retrieval (typically in the order of 1-30 minutes), suggesting that the forgetting rate in the short to intermediate term is normal in patients with schizophrenia (Aleman et al., 1999; Brebion, Amador, Smith, & Gorman, 1997; Gold et al., 2000; Paulsen et al., 1995; Rushe et al., 1999; Shoqirat & Mayes, 1988). Schizophrenic patients do not appear to be particularly susceptible to proactive interference (Elvevåg, Egan, & Goldberg, 2000; Paulsen et al., 1995).

Although recognition deficits may reflect impairment in either encoding or consolidation, in schizophrenia they likely reflect an encoding deficit, in view of normal forgetting. In accordance with this interpretation, the learning curves associated with the disorder are slightly slowed (Aleman et al., 1999). Moreover, the preferential deficit in recall versus recognition points to an additional impairment in retrieval, as subjects improve disproportionately in the condition that is less taxing on retrieval mechanisms.
Notably, the preferential retrieval deficit was observed in studies that matched tasks for difficulty level (Calev, 1984a, 1984b). Thus, these combined findings suggest mildly to moderately decreased encoding and a moderate to severe retrieval deficit in schizophrenia.

This pattern of memory deficit appears to be relatively independent of clinical state (acute or remitted), severity of psychopathology, patient status (in- or out patient), age and duration of illness (Aleman et al., 1999). It may, therefore, be considered to be a trait rather than a state characteristic of the disorder. Notably, the episodic memory impairments in schizophrenia cannot be accounted for solely by anticholinergic medication (Goldberg & Weinberger, 1996; Mortimer, 1997) and do not respond consistently to either typical antipsychotics or clozapine (Aleman et al., 1999; Lee, Jayathilake, & Meltzer, 1999; Paulsen et al., 1995).

The aforementioned pattern of episodic memory impairment can be clearly distinguished from profiles observed in other neuropsychiatric disorders. For instance, the amnesic syndromes are characterized by an abnormal transition from short-term to long-term memory (see chapter 2), while normal forgetting suggests that this is not the case in schizophrenic patients (e.g., Paulsen et al., 1995). In contrast to patients with frontal damage, schizophrenic patients do not show notable dependence of episodic impairment on 'executive' components of learning. That is, memory performance does not preferentially decrease with the attentional demand of the task, increasing difficulty of instructions, or in tasks that require organization of the information to be remembered for efficient performance (Aleman et al., 1999; Gold et al., 1992; Kenny & Meltzer, 1991; Saykin et al., 1994). Finally, schizophrenic patients can be distinguished from depressed individuals by the severity of memory impairments (Goldberg, 1999); two recent meta-analyses found effect sizes for global memory performance of 1.21 in schizophrenia versus 0.56 in depression (Burt, Zembar, & Niederehe, 1995). Furthermore, the memory deficits in depression correlate with depressive symptoms (Goldberg et al., 1993), while in schizophrenia correlations of memory deficits with symptomatology are small (Aleman et al., 1999; Goldberg et al., 1993; Rushe et al., 1999; Schmand, Kop, Kuipers, & Bosveld, 1992).

The profile of impairment in the episodic memory domain thus appears to be specific for schizophrenia. Crucially, this suggests that the underlying neuropathology may also be distinct from that in other neuropsychiatric disorders.

**Underlying brain pathology**

While evidence concerning the relation between memory impairment and brain pathology in schizophrenia is not yet conclusive, the pattern of memory impairment may be indicative of specific brain dysfunction. For instance, the lack of correlation with executive tasks renders a frontal origin of the memory deficits unlikely. As the deficits do not appear to be modality specific, an origin in a unimodal cerebral area is equally unlikely (Aleman et al., 1999; Gold et al., 1992; Kenny & Meltzer, 1991; Saykin et al., 1994). Furthermore, the stability of memory impairments argues against an underlying neurotransmitter imbalance. Such imbalances, for instance in dopamine transmission, would tend to fluctuate with clinical state and with age. As a consequence, so would related impairments, but memory impairments in schizophrenia remain relatively stable over the course of illness.

Memory deficits may instead be related to dysfunction of the mediotemporal lobe (MTL), which is thought to underlie many fundamental aspects of episodic memory.
function (Eichenbaum, 2000; Squire, 1992; Vargha-Khadem et al., 1997). The MTL has been implicated repeatedly in the neuropathology of schizophrenia and shows various structural abnormalities that appear to be of developmental origin (Harrison, 1999; Talamini, Louwerens, & Korf, 1995). These structural abnormalities in this region are more pronounced than elsewhere in the schizophrenic brain (Pakkenberg, 1993; Shenton, Kikinis, Jolesz, & al., 1992; Suddath, Christinson, Torrey, Casanova, & Weinberger, 1990; Wright, 2000). Moreover, two MRI studies have reported a correlation between the volume of medial temporal lobe divisions in schizophrenic patients and neuropsychological measures of memory performance (Goldberg, Torrey, Berman, & Weinberger, 1994; Gur et al., 2000). Also, functional neuroimaging studies have demonstrated abnormal levels of hippocampal activity, or blood flow, during the performance of memory retrieval tasks in schizophrenic patients (Crespo-Facorro et al., 2001; Heckers, 2001; Heckers et al., 1999; Heckers et al., 1998).

While volume reduction of the MTL is one of the most consistent findings in neuropathological research in schizophrenia (Harrison, 1999; Harrison & Eastwood, 2001; Nelson, Saykin, Hashman, & Riordan, 1998; Wright, 2000), the morphological changes underlying these volume reductions are not clear. For instance, there has been little consensus regarding neuronal numbers (reviewed in Harrison, 1999). However, neuronal numbers were unaltered in the so far sole stereological study performed on hippocampal specimens (Heckers, Heinsen, Geiger, & Beckmann, 1991). The data in this study are consistent with a hippocampal volume reduction through a reduction of the white matter compartment. With respect to the entorhinal cortex, one study that took into account the considerable intrinsic cytoarchitectural diversity of the region found only subtle reductions of cell number and density in some entorhinal subdivisions (Krimer et al., 1997). Considering that most studies do not report substantial changes in overall neuronal density (Arnold, 2000; Harrison, 1999; Harrison & Eastwood, 2001; Heckers & Konradi, 2002), one might tentatively conclude that any neuronal loss is, at the most, scaled with the loss of volume in these regions. In a previous study, we investigated consequences of such neuronal loss on memory performance (Meeter et al., 2002) and found that cell loss would hamper memory function only if combined with secondary changes in either wiring or levels of network activity.

A fairly recent approach to schizophrenia neuropathology has been to evaluate density of synaptic and dendritic molecules in schizophrenic hippocampi. The pertinent studies have generally reported reductions of these markers, particularly in the in- and output regions of the hippocampus (Harrison & Eastwood, 2001) and in the EC/parahippocampal gyrus (Hemby et al., 2002; Arnold, 2000). Although the data have not always been conclusive as to the cellular origin of the reduced neurites and synapses, the evidence suggests a loss of cortico-cortical connectivity within the parahippocampal gyrus, and between the parahippocampal gyrus and the hippocampus proper. There is less evidence for a reduction of intrahippocampal, recurrent fibers. This might be taken to suggest a distinct type of MTL pathology in schizophrenia, in which afferent, but not intrinsic hippocampal connections are preferentially affected. Such a pattern would be consistent with the lack of abnormalities in the transition from short- to intermediate-term memory in schizophrenia, which have been associated with damage to the hippocampus proper (see chapter 2).

In the present paper, we propose that a reduction of parahippocampo-hippocampal projections leads to impaired integration of sensory information, deficient episodic binding and, consequently, impaired memory performance. To test this hypothesis an MTL model was developed that could simulate performance in a variety of episodic memory paradigms, including list learning. The consequence of reduced
parahippocampal-hippocampal connectivity on network performance was assessed and compared with findings in schizophrenic patients. The objective of the study was to determine whether this particular type of miswiring could produce the pattern of memory deficits observed in the schizophrenic patient group. Model ‘anatomy’ and function will be described in the following sections. Later sections will present simulations of list learning, and various other episodic memory paradigms.

9.2 THE MODEL

Model architecture

As shown in various anatomical studies, the bulk of cortical input to the hippocampus is segregated over two streams (Burwell, 2000; Suzuki & Amaral, 1994; see Figure 30a). One of these streams runs over the perirhinal area and targets the anterior and lateral entorhinal cortex, while the other, via the parahippocampal cortex, projects mostly to the medial posterior portions of the entorhinal cortex. In primates, the perirhinal region mainly conveys information regarding meaningful objects (Aggleton & Brown, 1999; Murray, Bussey, Hampton, & Saksida, 2000), while the parahippocampal cortex is involved in memory for spaces and spatial relations (Bohbot, Allen, & Nadel, 2000; Vann, Brown, Erichsen, & Aggleton, 2000). Different aspects of an event may thus enter the hippocampus over dual input streams.

The dual input streams finally converge on the same cells in the dentate and CA3 region of the hippocampus. Before this point, however, they are interconnected at various levels within the parahippocampal region (Burwell, 2000; Suzuki & Amaral, 1994; Witter, Wouterlood, Naber, & Van Haeften, 2000). For example, the perirhinal cortex receives a considerable ‘horizontal’ projection from the parahippocampal region, while the lateral entorhinal cortex receives, next to a feedforward projection from the perirhinal region, a ‘horizontal’ one from the medial entorhinal cortex. These connections likely contribute to the integration of different kinds of information. Indeed, recordings of EC neurons during delayed-matching-to-place (DMTP) and delayed-matching-to-sample (DMTS) tasks indicate that the activity of entorhinal cells carries information about both objects and spatial locations (Suzuki, Miller, & Desimone, 1997). The entorhinal cortex may thus maintain transient representations of stimulus-context relationships (Dusek & Eichenbaum, 1997; Freeman, Weible, Rossi, & Gabriel, 1997; Suzuki et al., 1997; Young, Otto, Fox, & Eichenbaum, 1997). Context may here also consist of an endogenous event, or the behavioral significance of a stimulus.

By contrast, hippocampal neurons do not demonstrate evoked or maintained stimulus-specific codings in such tasks (Young et al., 1997). Moreover, hippocampal representations are known to be sparse and to hold poor feature representation. One of the contributions of the hippocampus proper to memory processing is to associate a compact code to the conjunction of cortical inputs, through the use of highly plastic connections. Thus, the hippocampus subserves fast auto-associative binding and storage of an episode (Eichenbaum, 2000; Squire, 1992; Vargha-Khadem et al., 1997; chapter 10 this dissertation).

The model used in the current simulations captures this basic organization. Four modules were used: Link, ParaLink and two input modules labeled ‘Item’ and ‘Context’ (see Figure 30b). The two input modules represent the cortical sources providing item and visuospatial information, respectively, to the MTL. Both input modules had a
random projection to a higher order module, ParaLink, which represents regions of the MTL where cortical input streams converge and start to be integrated. ParaLink in turn had full feedback connections to the Item and Context modules, which produced the output of the model. The highest module in the model hierarchy was Link. This module received a dense, random connection from ParaLink and sent a full projection back to the same module. Link represents the hippocampal subregions involved in compact coding for episodic information. These subregions may include the dentate and Ammon’s horn, as all these regions display sparse firing and poor specificity for stimulus modality or other stimulus features.

All connections in the model were fanning, as is the case for most of the known connections within the hippocampal and parahippocampal regions (some notable exceptions to this rule, including the reciprocal pathways between CA1 and the entorhinal cortex, are not implemented in the model). Learning was implemented with the Oja rule, a variant of the Hebb rule (see appendix to this chapter). The reciprocal connections between the EC and hippocampus proper are highly plastic, as suggested by numerous electrophysiological studies and by high concentrations of NMDA receptors in these pathways. Accordingly, reciprocal connections between the Link and ParaLink layers had a high learning parameter. Conversely, plasticity of the connections from the two input layers to the ParaLink module was relatively low, so that synaptic weights changed negligibly on the time scale of the present simulations.

Linear threshold nodes were used for the simulations (see appendix to this chapter). In both higher-level modules global inhibition was mimicked by k-Winner-Take-All dynamics, which limits activity in a layer to a predetermined number of nodes (k) receiving the largest input. This ‘k’ was relatively large in ParaLink, and small in Link. This is in accordance with findings showing sparse firing in the hippocampus proper and less sparse firing in parahippocampal areas (Amaral, Ishizuka, & Claiborne, 1990). The model was built using the Walnut software developed in our group (NeuroMod, 2000). Parameter settings were obtained through optimization procedures for free recall. All simulations were performed with the parameter settings listed in figure 1b and the appendix, except for manipulations explicitly stated in the paper.

**Functional model properties**

Episodic memories are generally seen as the concatenation, product, or convolution of the bits of information that together make up the episode (Humphreys, Bain, & Pike, 1989; Murre, 1996; Shiffrin & Steyvers, 1997). Similarly, episodic traces are formed in the model by allowing two input patterns to establish an integrated representation of their co-occurrence in Link and ParaLink. This occurs in the following manner: first, the patterns in the two input layers, representing item and context information, stimulate a set of nodes in ParaLink. Out of this set, the k ParaLink nodes with the largest input become active and form the ParaLink representation of the co-occurrence of item and context. Some of the activated nodes will represent information from just one input layer, but most of the k winners will be nodes receiving both types of input (see Figure 34a). In a similar way, the activated ParaLink nodes select a group of Link nodes. Through the plastic bi-directional connections between ParaLink and Link, these two representations are then bound together. On a more abstract level, Link thus enables
the ParaLink representation to be auto-associatively stored. Together, the ParaLink and Link pattern form the episodic trace.

The model can now retrieve patterns when presented with a partial cue in one of the input layers. A partial cue will not activate the entire associated pattern in ParaLink, but if the set of activated ParaLink nodes sufficiently resembles a stored representation, it will activate the associated Link nodes. These will then complete the original pattern in ParaLink, which can, in turn, reinstate associated information in the uncued input layer. The system can thus complete patterns in ParaLink through the recurrence between ParaLink and Link, while the feedback projections to the input layers enable the reinstatement of the uncued features.

The use of two layers, ParaLink and Link, for the formation of an episodic trace ensures that patterns with similar ParaLink representations do not interfere with one-another and can be retrieved separately. Similar ParaLink patterns are separated by their associated Link patterns, which usually resemble one-another less than the ParaLink patterns. This pattern separation occurs through sparse firing in Link and through LTD in the connections from ParaLink to Link, which both lead to decorrelation of patterns (O’Reilly & McClelland, 1994). Without such orthogonalization between ParaLink and Link, pattern overlap leads to the formation of spurious patterns that are always recalled. Indeed, we have shown previously that free recall scores drop dramatically when Link and ParaLink are fused into a single recurrent layer (Meeter et al., 2002).

Figure 30: (a) Simplified anatomy of the medial temporal lobe. The gray overlay depicts the four modules of the model. (b) Structure of the model. For each module, the number of nodes and the global inhibition parameter, $k$, (number of active nodes) is shown. For each connection, the percentage of nodes connected to each other over that connection is given.
Implementing schizophrenia neuropathology

Schizophrenia was modeled as a decreased connectivity from the input layers (Item and Context) to ParaLink, and from ParaLink to Link. Both sets of connections were decreased by 50%; whereas in the normal model, each ParaLink node received connections from on average 3% of Item nodes and 6% of Context nodes, this dropped to 1.5% of Item nodes and 3% of Context nodes in the schizophrenic model. Connectivity from ParaLink to Link dropped from 50% to 25%. The size of the reduction was chosen so as to be able to clearly show the effects of such a lesion.

To analyze the effects of individual pathways, we also considered models in which just one of these two levels of connectivity was reduced. As an alternative manipulation to miswiring, we evaluated whether increasing the noise in the system would lead to the memory deficit profile observed in schizophrenia. This manipulation was chosen because of speculations that some of the cognitive symptoms in schizophrenia may be due to a decreased signal-to-noise ratio in regions underlying semantic processing (Spitzer, 1997). In all, four manipulations were thus tested:

- the simulated schizophrenia manipulation, with a reduced number of connections from the two input layers to ParaLink, and also from ParaLink to Link.
- the reduced inputs to ParaLink manipulation, in which only connections from input layers to ParaLink were reduced
- the reduced ParaLink outputs manipulation, in which only connections from ParaLink to Link were reduced
- the increased noise manipulation

Noise was simulated as the firing of random nodes, uninvolved with normal memory processes, in either the input layers or ParaLink. From the normal level (1 extra input node active, no extra ParaLink node), we increased noise until it was equivalent to 8 extra input nodes and 3 extra ParaLink nodes.

9.3 SIMULATIONS

List learning

In list learning experiments participants are typically presented with a list of items to learn, consisting of single words or word pairs. Usually the items are presented one at a time. At some time after learning the participants are asked to retrieve the learned material. Hereby, they may be asked to 'free recall' the list; that is, to reproduce the entire list without the aid of item cues. In free recall, participants are assumed to only use context information as a retrieval cue. 'Context' is here taken to consist of all information that remains stable over the course of the learning trial, such as the room in which the learning takes place. In cued recall, participants are provided at test with a semantic or phonetic cue related to the learned items, for instance a superordinate category or word stem which they are asked to complete forming a previously learned item. The availability of item cues typically increases participants' recall performance.

A participant may also be asked to recognize previously learned items among a list of learned items and foil items. Foils are items that did not occur on the study list and which the participant must identify as new. In this case, maximal cueing is provided and
performance increases still further. Indeed, with the short lists and relatively long presentation times typically used in recall paradigms, healthy participants perform near perfect on this task. As the retrieval component in recognition tasks is reduced to a minimum, the participant's recall score is thought to closely reflect encoding performance.

Procedure: All three tasks discussed above were simulated in the model. List learning was implemented by presenting the model with a list of 10 items during the learning phase of the simulation. The item representations were activated one at a time, together with one stable context representation. Item and context representations consisted of sets of nodes (n=8) in the Item and Context layer, respectively. Item representations were non-overlapping. The model was allowed to cycle during three iterations and then learn the pattern emerging in the system. Transmission in the feedback connections of Link was dampened during learning, so that the activity in the network was largely determined by the 'on-line' inputs (Hasselmo, Bradley, Wyble, & Wallenstein, 1996; Meeter, Talamini, & Murre, subm.).

Importantly, the model did not have undifferentiated connection weights at the start of simulations. We initialized the weights in most connections in our network to what they would be after learning many random patterns (see 'Weight initialization' in the appendix to this chapter). Furthermore, prior to the actual list learning, all items used in the simulation (including foil items) were learned with a random context to simulate recent exposures to these items. This was done with exponentially varying learning rates to simulate heterogeneity in item frequency and recency. The same was not done for list contexts, since the assumption was that a context represents a unique configuration of stimuli. These combined network initialization procedures allowed us to simulate the learning of new patterns against the background of a 'full memory', and test retrieval under competitive circumstances.

Following the learning session, transmission in the feedback connections was restored. Retrieval was then tested under various cueing conditions representing free recall, cued recall and recognition (Table 8). Cueing consisted in partial activation of input patterns (see Humphreys et al., 1989 for a similar approach in an abstract model). A context cue was set by activating part (75%) of the context pattern that had been active during learning. This reflects the common assumption in memory modeling that following the task-instruction from the experimenter, a participant will partially reconstruct the original experimental context (e.g., Hasselmo & Wyble, 1997; Humphreys et al., 1989). In cued recall and recognition, the model was additionally provided with item cues. During recognition, this cue consisted in 75% of an entire item pattern (not 100%, because aspects of item presentation such as fonts or illumination may be different at study than at test). The cued-recall item cue was equal to half the recognition cue, reflecting partial presentation of the item. The recognition paradigm included presentation of foils (unstudied items) at test.

After setting the appropriate cues, the network was allowed to cycle for 150 iterations. Some noise was present at every iteration, in the form of random context nodes being active (on average 1 per iteration). As performance measure, we used feedback from ParaLink to the Item representations. The assumption underlying this mechanism was that episodic memory only stores co-occurrences of patterns that are themselves stored elsewhere. Episodic retrieval just serves to reactivate these patterns in the brain areas where they are encoded.
Table 8: Cues in the three list learning tests, free recall, cued recall and recognition. Cues are given as the number of activated nodes, relative to the number of nodes in a pattern.

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>Item cue</th>
<th>Context cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free recall</td>
<td>0</td>
<td>6/8</td>
</tr>
<tr>
<td>Cued recall</td>
<td>3/8</td>
<td>6/8</td>
</tr>
<tr>
<td>Recognition</td>
<td>6/8</td>
<td>6/8</td>
</tr>
</tbody>
</table>

If during an iteration the feedback input to an item pattern crossed a threshold, it was counted as retrieved, and the ParaLink and Link layers were reset (see appendix for how thresholds were set). In free recall, the number of different list items retrieved over the 150 test iterations was used as performance measure. Retrieval of any specific item was counted only once. In cued recall, only the first pattern retrieved was examined. A hit was scored if it was the cued item, an intrusion if it was another item, and an omission if no pattern reached threshold in the 150 test iterations. If any pattern reached threshold during recognition, it was counted as an ‘old’ response by the model; when no pattern was retrieved a ‘new’ response was counted.

Each simulation was repeated at least 50 times, with random initial settings. Results presented below are averages of these replications. Simulations were kept on a semi-quantitative level – the parameter space was not searched to produce a best quantitative fit for experimental data.

Normal performance: Free recall, cued recall and recognition performance in normal human participants could plausibly be simulated with this model. In free recall, the ‘normal’ model performs as a sample-and-complete network (similar in vein to the mathematical model, SAM; Raaijmakers & Shiffrin, 1981). Hereby, context cues activate parts of stored ParaLink patterns. The activated nodes in a sampled ParaLink pattern will activate Link nodes belonging to the same representation. This representation is subsequently completed in an interactive loop between Link and ParaLink pattern nodes. The competition between various sampled ParaLink patterns tends to be won by the pattern that, due to random fluctuations, receives the strongest activation at the onset of the retrieval trial. Feedback from the ParaLink pattern to the item layer leads to the item pattern crossing threshold and being counted as retrieved.

Figure 31a shows how in one run of free recall, activation of the ten stored patterns varies until activation of one given pattern crosses threshold (circles). Then the layers are reset, and the context cue again samples the network as before. Usually retrieved patterns were learned list items. In the run here, eight patterns were retrieved at least once (black circles). The thick black line tracks one such pattern, which was retrieved in the 37th iteration, and then again in the 62nd (gray circle; this retrieval does not count). Sometimes an item not part of the list was retrieved, as here in the 69th iteration (thick gray line). In recognition and cued recall, the item cue guided retrieval, and the cued pattern was retrieved rapidly (see Figure 31b for a run of cued recall; the thick black line is for the cued pattern).

The model was able to retrieve on average 49% of list items in the given time. Around half of the patterns were thus encoded well enough to be sampled and completed. As the
sampling occurred with replacement, sometimes a few strongly encoded patterns would be sampled persistently. In other replications many patterns had similar strength and would be sampled in turn. In recognition and cued recall the partly activated item representation guided the search process, leading to higher recall scores of 62% and 87.5%, respectively (see black line in Figure 32a).

False alarms were scored in recognition when a foil item was cued and reached threshold. As shown in Figure 32b, the model produced a small percentage of false alarms. From the hit rate and false alarm rate in recognition, a $d'$ of 3.2 was calculated. Intrusions were scored in free recall, when an item not on the list was retrieved, and in cued recall, when an uncued item reached threshold. Again, there were not exceptionally many intrusions; most errors in both conditions were omissions. The relative performance over recall conditions, both with respect to correct and false positive responses, is in line with data obtained from healthy human participants (Figure 33).

To investigate the context sensitivity of retrieval, all tests were repeated with a random new context. Upon cueing with 75% of a new random context, the model retrieved few items in the recognition and cued recall tests (on average 22% and 7% respectively), and free recall was virtually eliminated. This implies that retrieval in the normal system was highly sensitive to context, with only those patterns being retrieved that had been learned in conjunction with the test context.

Effects of miswiring: In Figure 33b, the effects of simulated schizophrenia through miswiring on memory performance in the model are depicted. This manipulation resulted in a preferential reduction of free recall performance and a smaller deficit on recognition. A similar profile has often been found in schizophrenia, of which an
example is given in Figure 33a (unpublished data, see caption). As has been repeatedly found (e.g. Paulsen et al., 1995), the deficit in cued recall was also smaller than that in free recall (Figure 32a). Moreover, there was an increase in false alarms in recognition (Figure 32b), and a large increase in retrieval when the system was cued with a random context, implying a drop in the context sensitivity of retrieval (Figure 32c).

Figure 32 shows how miswiring at each level of connectivity (Inputs-to-PL, and PL-to-Link) contributes to the current results. Reduction of inputs to ParaLink severely impaired free recall, without a notable loss of cued recall performance, and only a small reduction in the recognition hit rate. On the other hand, reduction of ParaLink outputs led to mild deficits on all three memory paradigms. Both the increase in false alarms and the increase in retrieval in a random context were seen with reduced inputs to ParaLink, not with a reduction of ParaLink outputs to Link. These deficits induced by the two levels of miswiring more or less add up in our simulated schizophrenia condition.

These results can be understood in the following manner. Successful free recall depends on the formation of ParaLink patterns in the intersection of projections from the context and the active Item (see Figure 34a). However, the decreased density of the input projections reduces the probability of ParaLink nodes to receive input from both input sources. This favors the inclusion of nodes receiving only a single type of input in

---

**Figure 32**: Model performance on the three memory paradigms (free recall, cued recall and recognition) for the control condition, simulated schizophrenia, and the two conditions with miswiring in just one set of connections (fewer input-ParaLink connections, or fewer ParaLink-Link connections). Maximal S.E.M. in the figure is 0.015 (with binomial distribution, 100 repetitions and 10 learned patterns). (a) Proportion of retrieved items in recall, or hits in recognition. (b) Proportion of intrusions in recall, and false alarms in recognition. (c) Proportion of retrieved items/hits when the model is cued with a random context.
ParaLink patterns (see Figure 34b). ParaLink nodes receiving input only from Item nodes cannot be activated by the context cue during free recall. Hence, free recall is compromised. The impairment is less pronounced when item cues are provided, since with cues from both input sources, the ParaLink nodes receiving input from a single source also have a probability of being activated.

The reduction of Input connections also shifts the balance between the bottom-up and top-down inputs to ParaLink (from Input layers and Link, respectively), in favor of the Link-to-ParaLink inputs. This is the cause of the increase in false positives seen with the reduced Inputs-to-ParaLink manipulation (Figure 32b): there is a tendency to ‘complete’ patterns towards already stored ones, even when the recalled pattern does not match the input cues. The same mechanisms underlie the decreased context sensitivity of the system shown in the large number of inappropriate retrievals after cueing with a random new context (Figure 32c).

The reduction of ParaLink outputs to Link creates a different problem for memory function (Figure 34c). As a single ParaLink node now contacts fewer Link nodes, the correlation between firing of specific ParaLink and Link nodes increases. Thus, mean pattern overlap between Link patterns increases (from 6% to 13%). With increased pattern overlap in Link, its ability to separate similar patterns decreases. Consequently, Link may activate overlapping and competing representations in ParaLink. The crucial

---

**Figure 33:** Comparison of real and simulated data on free recall and recognition. (a) Proportion of items retrieved by free recall after one learning trial, and hit rate for item recognition after a x minute delay. The data, unpublished, is from 21 patients with schizophrenia and 28 controls, tested on the CVLT. (b) Retrieved items in free recall, and hit rate in recognition in the simulation (same as in Figure 32). (c) Sensitivity measure, d’, for the schizophrenic patients and normal controls and in the corresponding simulated groups (d’ from (Calev, 1984a) also added).
function of the Link layer is to keep patterns in the system segregated, and this function thus becomes compromised. These problems in pattern segregation result in less ParaLink patterns being completed by their Link companion, and less items reaching threshold. This deficit does not involve the sampling of memories, but their completion. Retrieval therefore suffers irrespective of cueing condition. Performance in all three memory tasks drops to an equal extent, and there is no rise in false positives or recall in random new contexts.

In essence, the reduction of inputs to ParaLink compromises the integration of cortical inputs into an episodic trace, leading to inadequate activation of memories when few cues are used. The reduction of ParaLink outputs to Link leads to a deficit in pattern segregation, hampering autoassociative storage and retrieval. The memory deficits seen in schizophrenia result from the concatenation of these two problems: the pattern segregation problem makes some pattern unretrievable, affecting all tasks equally and leading to what has been interpreted as an encoding deficit. The reduced integration problem leads to a lesser efficacy of single cues, affecting tasks with large retrieval demands and leading to what has been interpreted as a retrieval deficit.

Figure 34: Formation of intersection patterns. At any time, the ParaLink pattern consists of the k nodes that receive the highest amount of input. (a) In the normal model, most ParaLink nodes in a pattern receive inputs from both sources – in the figure, they reside in the intersection of the two projections. (b) Reduced input connections imply sparser projections from single input patterns to ParaLink and, thus, smaller intersections of the projections from the two concurrently active input patterns. As a result, only a small part of the ParaLink pattern comes to represent both inputs. (b) Reduced connections to Link lead to a higher overlap between Link patterns. This decreases the likelihood of correct pattern completion through the Link pattern.
Effects of noise: To investigate the specificity of the above deficits, we also investigated an alternative implementation of schizophrenia, involving increased noise (Spitzer, 1997). Figure 35 shows how increased levels of noise in ParaLink and Link affected memory performance. With increasing noise, cued recall but especially free recall performance became progressively lower. Recognition performance was notably resistant to noise: although maximal levels of noise provided more input than the context cue, the context and item cues together were still able to guide retrieval to the appropriate patterns. At levels of noise where recognition performance did start to decrease, free recall had already deteriorated to very low levels. There was thus no single noise level that produced the pattern observed in schizophrenia, namely a mild to moderate recognition deficit and moderate to severe free recall impairment. In sum, an increase in the noise level did not lead, in our model, to the memory profile observed in schizophrenia.

As a side note, although performance benefited from low levels of noise, it dropped significantly when noise was totally absent (Figure 35, left side). A certain degree of noise in the input turned out to be necessary for sampling to be successful. Otherwise, even small differences in encoding strength between patterns would lead to the same pattern being sampled continuously.

Proactive interference

As an additional test of the model, we decided to investigate whether the model would reproduce these findings. In a recent study, proactive interference effects were found to be much smaller than the general memory deficits of schizophrenic patients (Elvevåg et al., 2000).

In the model, proactive interference was evaluated through a typical design, in which participants learn two lists of paired associates (A-B, followed by A-C). The task is to generate the second word of a pair (the target word) when cued with the first word (the cue word). Cue words used in the first list (the AB list) are repeated in a second list (the AC list), but are associated with different target words (e.g. if dog-lamp is in the first list, then dog-church might be part of the second list). Participants are tested on the AB

![Figure 35: Influence of noise in the input to ParaLink (input noise) and in the input to Link (Link noise). Input noise was implemented by activating a number of random context nodes, Link noise by simulating the inputs that a number of randomly activated ParaLink nodes would send to Link. (a) Performance on the free recall test. (b) Performance on the cued recall test. (c) Performance on the recognition test.](image)
list after studying only that list, and on the AC list after studying both lists. A typical finding is that participants do worse on the AC list than on the AB list, presumably because the first response associated with a cue word (that on the AB list) interferes with storage or retrieval of the second response associated with the cue.

Procedure: The AB-AC paradigm was simulated as follows: the model was presented with pairs of items in the first list, which were activated one pair at a time together with the associated context. The activation patterns emerging in the model were learned over three iterations. In a subsequent retrieval test, the model was presented with the first item (or ‘cue word’) of each item pair, as well as a 75% context cue. The model was allowed to ‘search’ the associated target item during 50 iterations. Next, the second list of word pairs was learned with the second associated context. In the subsequent retrieval test, the cue items used in the first list, coupled with a context cue for the second list, should elicit the associated response items from the second list.

In this simulation items representations consisted of four nodes each. The items were ‘prelearned’, as in all other simulations. To simulate that the two lists are presented during the same experimental session with only a brief delay interposed, context patterns associated with the first and second list overlapped by 50%.

Results: In the simulation, both the control condition and the schizophrenia condition produced an interference effect, in that the number of hits was lower on the second list than on the first (Figure 36b). In raw scores, the interference effect was slightly larger in the schizophrenia condition. This was also the case in the Elvevåg et al. study, but that difference was not significant. We therefore subjected our data to a repeated measures analysis of variance (ANOVA) with model permutation (simulated schizophrenia, normal model) as the between-group factor and list (list 1, list 2) as the within-group factor. Not surprisingly, there was a main effect of group, due to the simulated patients’ inferior performance, $F(1,198)=1595.6; p<0.001$, and of list, with performance on list 1 being superior to that on list 2, $F(1,198)=13.8; p=0.001$. However, there was no significant interaction between group and list, $F(1,198)=3.05; p=0.08$, notwithstanding one hundred simulated participants per model condition. As in Elvevåg et al. (2000)’s analysis, the interference effect was dominated by the overall memory deficit shown in the patient group.

Elvevåg et al. (2000) also analyzed intrusions from the AB list in the AC list (e.g., if a participant generates ‘lamp’ to ‘dog’, while ‘church’ is the correct AC-list associate to ‘dog’). Correspondingly, we scored intrusions when the model produced a response from the first list during testing of the second list. In line with Elvevåg et al.’s (2000) results, we did not find a preferential increase in the number of AB-intrusions (from 10% of all answers in the intact model to 14% in the schizophrenia manipulation, which is 20.3% and 20.7%, respectively, of all errors).

Again, the data pattern was not replicated when schizophrenia was implemented as increased noise. Instead, interference effects disappeared when noise was increased. Figure 36b shows, for example, results from a simulation with the maximum noise, equivalent to 8 extra input nodes and 3 extra ParaLink nodes. Here, performance was even better for the second AC list than for the first AB list. Such effects were found in all simulations with increased input noise. Further analysis showed that this was a result of overlap between representations for AB and AC patterns in both ParaLink and Link. This overlap led to stronger weights (as nodes in the pattern were twice simultaneously
active during learning), which helped AC patterns overcome noise from which AB patterns suffered.

9.4 DISCUSSION

The model presented here provides a connectionist account of episodic memory and its dysfunctions in schizophrenia. In the model dual sensory processing pathways, conveying different aspects of an episode, converge on the hippocampal formation where a representation of their co-occurrence is stored. In schizophrenia, this integration may become compromised through reduced connectivity within the parahippocampal gyrus, and in the perforant path from entorhinal cortex to hippocampus proper. Before we discuss our work as a model of memory and a model of schizophrenia, several remarks on this architecture are in order.

A point that merits attention is the use of the term ‘context’ in the model. This term is usually denotes all features of the situation (of environmental or endogenous origin) other than the stimulus in the focus of attention. Here, context was identified more narrowly with the spatial configuration of stimuli at the time of an object presentation. Processing of this information is known to occur in the stream through the parahippocampal cortex (Bohbot et al., 2000; Vann et al., 2000). Neocortical inputs in that stream were identified with the context module. One might argue that such context information is in itself already episodic, and would require hippocampal contributions for its storage. In the present model, the difference between representations in the context input and higher order layers is the larger contribution of object information to the latter. In the real brain the hippocampus may, additionally, contribute affective and motivational components to the episodic representation.

In the list-learning paradigm input conveyed by the context-layer remains constant over the course of a simulation, while the input conveyed by the Item layer changes

Figure 36: Interference data from schizophrenic and healthy subjects (a) and corresponding simulated groups (b). Human data is from experiment 1 in Elvevag et al. (2000). See text for description of experiment. Retrieval rates for the first list (AB) and second list (AC) are shown.
CHAPTER 9

considerably. Accordingly, the connectivity characteristics of the two input layers are different, supporting associations of multiple items with one context representation. It will be interesting to see if the same circuitry can support paradigms in which the spatial configuration of stimuli changes quickly over time. It will also be noteworthy to investigate whether the same circuitry can be made to handle configurations of non-object as well as object information.

A model of memory

Although the current model is quite simple by the standards of the field, it does incorporate some realistic features as a model of human memory. Recall performance increments from free recall, to cued recall, to recognition. This profile arises because with increasing cue-information, the probability of retrieving a correct item increases. Secondly, retrieval is highly context-sensitive. Context can also discriminate between episodes incorporating the same item, such as in the AB-AC paradigm, where correct episodes are retrieved with only small interference effects.

However, the model does not account for a number of phenomena. Learning was modeled as involving only one item at a time, and leading to associations only between item and context. In reality inter-item associations are formed through concurrent activation of items in working memory, and modeling them would improve our account of free recall (Raaijmakers & Shiffrin, 1981). Retrieval was implemented as a process of sampling with replacement. If this were changed to sampling without replacement, performance would obviously improve as already retrieved patterns would not compete with still-to-be retrieved patterns. By introducing either adaptation or depressing synapses in the model, such sampling could be simulated. Moreover, forgetting was not explicitly modeled (though some overwriting of older patterns by newer ones did occur, especially in conditions in which pattern overlap was substantial).

Recognition is often thought of as a dual process (Mandler, 1980; Norman & O'Reilly, in press; Yonelinas, 2002). The two putative process components, recollection and familiarity, are not distinguished in the present model. The current implementation of recognition as a retrieval process would be a natural implementation of recollection. Familiarity could be the contribution of direct activation by input nodes of associated ParaLink nodes. Indeed, such direct activation sometimes led to crossing of the threshold of an item, leading to a positive recognition response. Preliminary simulations showed that this way of implementing a dual process recognition did reproduce some characteristics of the two putative processes, but familiarity did not add much to recollection, which was very efficient.

Another shortcoming of the model is its sensitivity to repeats of items. Repeats in the model make retrieval of the repeated information easier, even in situations where it would not help a human participant. An example is a two-list interference paradigm in which the second list consists of rearranged cue-target combinations from the first list are (A-B, A-Br). While in humans this causes massive confusion and a decrement in performance, in the model the repetition of target words counteracts interference effects and brings performance for the second list (A-Br) up to near-normal levels. These effects would probably disappear if items had more ‘prelearning’ trials, implementing a longer history of the item as a unit.
A model of schizophrenia

Other models may thus be more complete as theories of memory function. Via its explicit modeling of parahippocampal integration, however, the current work allows a principled investigation of the effects of malformations seen in the schizophrenic brain. Schizophrenia was modeled as a reduction of connectivity in the pathways subserving integration of information – those within the parahippocampal gyrus, and from the parahippocampal gyrus to the hippocampus. This is consistent with findings of reductions in dendritic markers in entorhinal cortex and parahippocampal gyrus (Arnold, 2000; Hemby et al., 2002), and in the in- and output regions of the hippocampus (Harrison & Eastwood, 2001).

These manipulations led to recall deficits that are progressively worse with increasing dependence of retrieval on the context cue. Proactive interference effects were not increased, as was found in schizophrenics. Another manipulation, an increase in noise, could not reproduce these findings. It should be noted that effects not modeled here may also partly explain memory deficits observed in schizophrenic patients (e.g., attentional deficits or abnormalities in sensory processing).

Although not the target of the modeling, amnesia-like memory deficits, with impaired memory performance across all tasks, were also reproduced by the model. They occur following perturbations of the ‘hippocampal’ layer, as was apparent with connectivity lesions restricted to the connections between ParaLink and link. This matter is investigated more thoroughly in a related model, in which simulated lesions of the autoassociative Link module resulted in a breakdown of storage and the emergence of Ribot gradients typically seen in retrograde amnesia (Meeter & Murre, in press; Murre, 1996).

Our modeling efforts have been concentrated on memory in schizophrenia. It is congruent, however, with a larger amount of data suggesting that some other cognitive dysfunctions in schizophrenic patients may be related to an abnormal use of context information (Cohen, 1999; Nestor et al., 1997). Examples are abnormalities in semantic categorization and other paradigms assessing the organization of ‘semantic space’. Briefly, some experiments suggest normal low-level representations of individual items in schizophrenia, while the ‘associational management’ of these representations appears to be impaired. It can be envisaged that the integration deficits investigated here may lead to formation of abnormal associations and to abnormal associational accessing of information. The present model might thus be extended to impairments in categorization tasks and other tasks assessing semantic space, as well as symptoms such as loose association, delusions and conversational drift. However, a number of these points have not been adequately quantified, which might lead to relatively unconstrained modeling (e.g., it is very easy to have a model ‘hallucinate’, as many changes in modeling architecture will produce spurious activity in computational models).

Closer to the present work is the modeling of other memory paradigms, with the goal of deriving predictions for memory performance in schizophrenia. Preliminary simulations have already suggested that source monitoring may be affected in schizophrenia, but that retroactive interference effects are smaller in the schizophrenia condition than in the normal model.

Although more research is clearly needed on memory and schizophrenia, the present findings support an important role of the parahippocampal region in memory deficits of patients with schizophrenia. Impaired integration of information and impaired
autoassociative binding may together explain the pattern of memory impairment seen in schizophrenia.

APPENDIX TO CHAPTER 9

Model neuron
The model is built with simple linearly summing nodes with k-Winner-Take-All dynamics. A node $i$ can either be active ($S_i=1$) or inactive ($S_i=0$). Whether or not neuron $i$ is active depends on its total input $H_i$, the weighted sum of the input it receives from all nodes $j$ with which it is connected:

Equation 2:  
$$H_i = \sum_j w_{ij} S_j$$

The weights $w_{ij}$ can vary between 0 and 1. In each module, the $k$ nodes with the highest total input ($H_i$) become active. If several nodes on the cut-off have an equal total input, a random selection of these nodes is activated so as to keep the total number of active nodes equal to $k$. The parameter $k$ is set separately for each module in the model (see Figure 30b).

Learning rule
The learning rule used is the Oja variant of the Hebbian learning rule (Levy et al., 1990; Oja, 1982) showed that this rule fitted LTP data well). The rule is given in Equation 3.

Equation 3:  
$$\Delta w_{ij} = \mu (S_i S_j (1-w_{ij}) - S_j (1-S_j) w_{ij})$$

As in the normal Hebbian rule (Hebb, 1949), the weight is strengthened whenever both the presynaptic and the postsynaptic nodes fire (i.e., whenever $S_i S_j = 1$). The weights do not grow without bounds: by multiplying weight change by $1-w_{ij}$, the Oja rule assures that with continuing learning weights asymptotically approach the implicit maximum value of 1. Weights decrease whenever the presynaptic node does not fire but the postsynaptic node does (i.e., whenever $S_i (1-S_j) = 1$), modeling heterosynaptic LTD. Again, weights do not decrease without bounds but asymptotically approach the implicit minimum weight value of 0. Learning is scaled by the learning rate, $\mu$.

Weight initialization
An important characteristic of the Oja rule is that the long-term expected value of a weight is equal to the average activity level of the presynaptic node given that the
postsynaptic node is active (Levy et al., 1990). For example, if the presynaptic node is active 60% of the times that the postsynaptic node is active, then the weight on the connection between the two nodes will tend to hover around 0.60. We use this characteristic to initialize the weights in most connections in our network. We initialize them to the value that they would have after many independent patterns are learned. This simulates the background of a “full memory”.

Using the binomial distribution, bottom-up determination of firing patterns during learning and independence of input nodes, we calculate the likelihood that, given postsynaptic firing, the presynaptic node also fires. This value is then used as the mean value in a weight distribution. The means for different connections are given in Table 9. Although in this way the expected weights can be calculated, it is impossible to analytically derive the full weight distribution. Using Monte Carlo simulations, we determined that the shape of the distribution around the mean depends both on the size of the mean and of the learning rate. If the mean was not too low and the learning rate not too high, the distribution was approximately normal. We estimated the standard deviation of this approximating normal distribution, and then added continuously distributed noise with that standard deviation and mean 0 to the weights.

One obvious determiner of the likelihood that presynaptic node \(x\) fires given firing of postsynaptic node \(y\), is whether there is a reverse connection between the two. If the presynaptic node gets input from the postsynaptic node (i.e., the connection is reciprocal), its likelihood of firing is greater than when no such reciprocal connection exists. Table 9 therefore lists weights of feedback connections separately for those node pairs that have a feedforward connection as well as one feedback, and those that do not.

The one exception to this initialization procedure are the connections from ParaLink to Link, which were initiated at a higher value to allow the formation of orthogonalized patterns. This alters the balance between LTP and LTD, favoring LTD over LTP. This was necessary as LTD strongly favors pattern orthogonalization, while LTP diminishes it (O'Reilly & McClelland, 1994).

Threshold setting
The input regions of the model also function as output regions. For example, activity in the item layer determines whether an item is retrieved or not. The output of the model

<table>
<thead>
<tr>
<th>connection</th>
<th>learning rate</th>
<th>Initial weight (SD betw. brackets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>item-paralink</td>
<td>0</td>
<td>.33 or 0</td>
</tr>
<tr>
<td>paralink-item</td>
<td>.03</td>
<td>.484 (.12) or .09 (.04)</td>
</tr>
<tr>
<td>context-paralink</td>
<td>0</td>
<td>.30 or 0</td>
</tr>
<tr>
<td>paralink-context</td>
<td>.03</td>
<td>.253 (.07) or .09 (.04)</td>
</tr>
<tr>
<td>paralink-link</td>
<td>.20</td>
<td>.50 or 0</td>
</tr>
<tr>
<td>link-paralink</td>
<td>.125</td>
<td>.035 (.02) or .015 (.01)</td>
</tr>
</tbody>
</table>

Table 9: Parameters of the connections in the model (listed with the originating layer and the target layer). In feedforward connections (marked in italics), a certain percentage of node-node connections ‘exists’ (see Figure 30); they are created with a certain weight; all others start out with a weight of 0 and do not change in the simulation. Weights in feedback connections are computed as explained in the appendix. As the number of existing connections is different in simulated schizophrenia, the computed feedback weights are also different.
is based on the feedback signal that the input regions receive from the parahippocampal module. We assume that the input nodes have adapted to the statistics of the feedback that they receive from the parahippocampal module. The threshold is therefore set at some standard location on the interval between the expected feedback elicited from the parahippocampal module if the pattern is retrieved, and the expected feedback elicited when a random collection of parahippocampal nodes is active.

An example of such criterion-setting is illustrated in Figure 37. In the standard situation, 32 parahippocampal nodes will be active, which is 10% of all parahippocampal nodes. As this is also the expected frequency that a given parahippocampal node is active given that a certain item node is active, the mean weight on the feedback connection of a random parahippocampal node to this input node is 0.10. If a random pattern is thus active in the parahippocampal module, the expected feedback signal to an item node is: 

$$32 \times 0.10 = 3.2.$$ 

If the item node is part of the active pattern in the item module, the pattern in the parahippocampal module will not be random anymore. Instead, many nodes in the parahippocampal pattern will be innervated by the active item node. To be exact, on average 15% of the active parahippocampal nodes will have a connection from the active item node (this percentage was calculated using the binomial distribution, and is dependent on the value of both k and connection parameters).

![Figure 37: Example of how feedback signals are used to distinguish old from new patterns. Plotted is a frequency plot of the feedback signal onto 1000 Item nodes. Three situations are plotted. The node belongs to a pattern that was learned in the context, and has been retrieved (old). It may also belong to a pattern that was not learned and that is now tested for free recall (new-recall), or for a recognition trial (new-recognition). Cueing condition does not matter for feedback given that an item is retrieved. For new items, it does: if item nodes are activated as a cue (as in recognition), it leads to more feedback to the not-activated nodes than when no item nodes are active (see text). Criteria for counting a pattern is as retrieved are set halfway the center of the old distribution and either the new-recall distribution (the recall criterion) or the new-recognition distribution (the recognition criterion).](image)
From the perspective of an individual parahippocampal node, if it receives a feedforward connection from the active item node, its likelihood to be part of the pattern jumps from 10% to 48%. Parahippocampal nodes that do not receive a feedforward connection from the active node, have a likelihood of only 9% to be part of the pattern. Since the expected value of the weights is equal to these likelihood, parahippocampal nodes that receive a feedforward connection from the active item node will send a feedback connection back with an average weight of 0.48. Nodes without the feedforward connection send a feedback connection back with an average weight of 0.09.

Now we can calculate the expected value of the feedback signal to an active item node: $15\% \times 32 \times 0.48 + 85\% \times 32 \times 0.09 = 4.74$ (we can also recalculate the feedback from a random pattern: $3\% \times 32 \times 0.48 + 97\% \times 32 \times 0.09 = 3.2$, as above). If an item node is active, this will thus shift its expected feedback-activation from 3.2 to 4.74.

If the association between the activated parahippocampal pattern and the item node is learned, the weights on the feedback connections to the item node will be strengthened, leading to a higher expected feedback signal. Using the Oja learning rule given in Equation 3 and the learning rate used in this paper (0.03), the expected feedback signal will become: $15\% \times 32 \times (0.48 + 0.03 \times (1 - 0.48)) + 85\% \times 32 \times (0.09 + 0.03 \times (1 - 0.09)) = 5.63$. If a stored pattern is active in the parahippocampal module, this will thus engender a feedback signal of 5.63 to an item node associated with this pattern.

The fact that patterns were also presented for what we have called a pre-learning trial in a random context, raises the expected feedback even more. It can be shown that with more and more learning trials in random contexts, feedback weights will exponentially approach the likelihood that the target node is in a pattern, given the number of feedforward inputs it gets from all of the pattern nodes. In the simulations presented here, there was one 'prelearning' trial with a learning parameter of an average 0.66 (as learning parameters were drawn from an exponential distribution with $\lambda=1.5$ for prelearning). Using the average prelearning trial in the calculations, the expected feedback for retrieved patterns becomes 5.68.

The criterion for retrieval is set at 50% of the interval between the feedback signal for a random pattern, and the feedback signal for a stored pattern. In this case: $3.2 + 0.50(5.68 - 3.2) = 4.44$. In other words, if the feedback signal to the item nodes belonging to one item exceeds 4.44, this is taken as evidence that a stored pattern has become active in the parahippocampal module.

Because activating item nodes already changes their feedback signal, the feedback signal is only measured in not-activated nodes. In recognition, where 6 item nodes are activated as cue, only the feedback signal to the two non-activated pattern nodes contributes to the output measure. However, even then the fact that the cued item nodes have been activated together with the not-cued item nodes changes the feedback signal to these latter nodes (in this case, from 3.2 to 3.44). The criterion in recognition is therefore set with respect to the baseline feedback signal when the item cue is activated: $3.44 + 0.50(5.68 - 3.44) = 4.56$.

In practice, we average feedback to all item pattern nodes, and count the item as retrieved if the average feedback exceeded criterion. Moreover, we divide the feedback signal by the number of ParaLink nodes, so as to get a number between 0 and 1. Expected feedback for unstored items is then simply the average feedback weight (i.e., 0.10), while the criteria become: $4.44 / 32 = 0.139$ (recall), and $4.56 / 32 = 0.142$ (recognition, see Figure 37).
Because of a different feedforward connectivity, feedback weights are different in the simulated schizophrenia model (see Table 9). Expected feedback for unstored items remains the same (3.2), but the expected feedback from stored patterns becomes bigger (from 4.56 it became 6.13). This leads to different criteria than in the normal model: 0.146 in recall, 0.1501 in recognition.