



Published in final edited form as:

*J Math Psychol.* 2010 February 1; 54(1): 73–89. doi:10.1016/j.jmp.2009.04.012.

## Modeling and Estimating Recall Processing Capacity: Sensitivity and Diagnostic Utility in Application to Mild Cognitive Impairment

**Michael K. Wenger,**  
The Pennsylvania State University

**Selamawit Negash,**  
Mayo Clinic College of Medicine

**Ronald C. Petersen,** and  
Mayo Clinic College of Medicine

**Lyndsay Petersen**  
Mayo Clinic College of Medicine

### Abstract

We investigate the potential for using latency-based measures of retrieval processing capacity to assess changes in performance specific to individuals with mild cognitive impairment (MCI), a reliable precursor state to Alzheimer's Disease. Use of these capacity measures is motivated in part by exploration of the effects of atrophy on a computational model of a basic hippocampal circuit. We use this model to suggest that capacity may be a more sensitive indicator of undelying atrophy than speed of processing, and test this hypothesis by adapting a standard behavioral measure of memory (the free and cued selective reminding test, FCSRT) to allow for the collection of cued recall latencies. Participants were drawn from five groups: college-aged, middle-aged, healthy elderly, those with a diagnosis of MCI, and a sample of MCI control participants. The measure of capacity is shown to offer increased classificatory sensitivity relative to the standard behavioral measures, and is also shown to be the behavioral measure that correlated most strongly with hippocampal volume.

---

Since the time of Ebbinghaus (1885), the study of memory has been guided by numerous conceptualizations and metaphors for what memory is and does (see, e.g., Kandel, Schwartz & Jessell, 2000; Neath & Suprenant, 2003; Roediger, 1980; Spear, 1978). Although there are numerous differences among these views, a common thread is that a comprehensive understanding of memory functioning requires consideration of four aspects of memory functioning: encoding, consolidation, retention, and expression (retrieval). Within that consensus view, however, there does exist a critical dichotomy. Specifically, aspects of encoding, consolidation, retention, and expression have typically been considered either in terms of the information involved or in terms of the processes that operate on that information (see, e.g., Kollers & Roediger, 1984; Massaro, 1998; O'Toole, Wenger & Townsend, 2001; Roediger, Gallo & Geraci, 2002; Wenger, 1999). Typically, the conceptual distinction has been

---

© 2009 Elsevier Inc. All rights reserved.

Correspondence: M. J. Wenger, Department of Psychology, 216 Moore Building, The Pennsylvania State University, University Park PA 16802, mjw19@psu.edu..

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

expressed in the choice of the dependent variable, with questions regarding information relying on frequencies of various types of outcomes, and questions regarding processing relying on reaction times (RTs).

This same dichotomy has also characterized much of the research on life-span changes in memory, with respect to both changes associated with normal development and aging, and changes associated with impairments or disease states (see in particular Cerella & Hale, 1994; Kail & Salthouse, 1994; Petersen, Smith, Ivnik, Tangalos, Schaid, Thibodeau, Kokmen, Waring & Kurland, 1995; Petersen, Smith, Kokmen, Ivnik & Tangalos, 1992; Petersen, 2007; Petersen & Negash, 2008). A central purpose of the work described here was to use a novel (in this context) measure of memory performance to provide a link between the focus on memory information and the focus on memory processes. That link is provided by adding a theoretically-motivated measure of a fundamental characteristic of processing—specifically, a measure of processing *capacity*—to a memory task that has been shown to be very sensitive in characterizing and detecting changes associated with mild cognitive impairment (MCI). MCI is a diagnostic category that has been shown to reliably distinguish the mild deficits associated with normal aging from the more profound deficits associated with Alzheimer's disease (AD, e.g., Petersen, 2007; Petersen & Negash, 2008; Petersen, Smith, Kokmen, Ivnik & Tangalos, 1992; Petersen, Smith, Waring, Ivnik, Tangalos & Komken, 1999). MCI status has also been shown to be predictive of a transition to AD (e.g., Dickerson, Sperling, Hyman, Albert & Blacker, 2007; Morris, Storandt, Miller, McKeel, Price, Rubin & Berg, 2001).

We begin with a general consideration of the distinction between processing speed and processing capacity, as this distinction will be crucial to understanding the potential issues for maximizing the sensitivity of behavioral changes associated with early neural pathology. We then use a computational model of hippocampal connectivity and integrity to investigate the extent to which the various behavioral measures of processing speed and capacity may be differentially sensitive to the types of degeneration and atrophy that have been shown to be characteristic of MCI and AD. Finally, we describe an experimental investigation of the utility of combining measures of processing speed and capacity with standard measures of recall. To our knowledge, the experimental work represents the first and only use of reaction times with frequencies of cued and free recall in application to this clinical issue.

## Processing Speed And Capacity in Memory

Studies concerned with the characteristics of memory processing have often focused on the speed or rate with which memory tasks can be performed. This reflects a long history (dating at least to work in the nineteenth century, such as that of Donders, Wundt, and Jastrow Boring, 1957; Jastrow, 1890; Townsend & Ashby, 1983) of relying on response latencies as a source of evidence for the temporal and logical relationships among mental processes (see Luce, 1986; Townsend & Ashby, 1983; Townsend & Wenger, 2004a, for historical reviews). With respect to changes across the life span, suggestions that response latencies may be critical to understanding both normal aging and disease states also have a reasonably long history (e.g., Birren, 1974; Salthouse, 1985; Welford, 1958).

Reviews of the literature (see in particular Cerella & Hale, 1994; Kail, 1991; Kail & Salthouse, 1994) have outlined a general pattern relating overall response time to age: an inverted-U shaped function (see, e.g., Kail & Salthouse, 1994, Figure 1, p. 200). That is, processing speeds increase from early childhood through adolescence, peaking in young adulthood, remaining stable through very early middle age, before beginning an uninterrupted decline. The tasks that have been used to provide evidence for this pattern of changes include a wide range of tests of perceptual, mnemonic, and cognitive abilities, such as simple and choice RT, same-different

and categorization judgments, short-term memory scanning, recognition, and recall (see, e.g., Cerella & Hale, 1994, Table 1, p. 130).

Kail, Salthouse, and colleagues (e.g., Kail & Salthouse, 1994; Salthouse, 1985, 2000, 2001) have argued that differences in processing speed can be interpreted in terms of global differences in mental capacity (e.g., as a general cognitive resource) or in terms of general executive control functions. There have been numerous criticisms of this and related conceptualizations of capacity (e.g., Fisher & Glaser, 1996; Kantowitz & Knight, 1976; Light & Burke, 1988; Schweickert & Boggs, 1984), most focusing on the extent to which the typically-used measures of response speed—particularly mean or median RT—can be interpreted in terms of processing capacity.

There are a number of conceptual problems associated with using mean or median RT as a measure of processing capacity. Two are of particular relevance for the present effort. The first concerns the manner in which capacity differences are revealed. As discussed in detail elsewhere (see in particular Townsend & Ashby, 1978, 1983; Wenger & Gibson, 2004; Wenger & Townsend, 2000, 2006), capacity is best understood in terms of the response of a system (e.g., memory) to variations in workload. In the work presented here, workload is operationalized as the number of retrieval cues presented on each trial of a cued recall task. A system that shows degraded performance with increases in workload is characterized as possessing a capacity limitation. A system that shows no changes in performance with increases in workload is characterized as possessing unlimited capacity (at least across a specified range of workload manipulations, see Fisher, 1984). Finally, a system that shows performance improvements with increases in workload is characterized as possessing unlimited capacity. A central characteristic of much of the work that has been used to support age-related differences in capacity is that manipulations of workload are not consistently employed, nor are contrasts made with respect to the effects of variations in workload. Consequently, differences in processing speed cannot be unambiguously interpreted in terms of differences in capacity.

The second conceptual problem concerns the information that is provided by, for example, the mean RT. By definition, the mean gives the expectation for task completion time (overall speed). In this sense, it is at best only coarsely reflective of the notion of processing capacity (see the detailed exposition in Townsend & Ashby, 1978). In contrast, the work reported here uses a measure on the entire RT distribution, one originally suggested by Townsend and Ashby in a critical and comprehensive consideration of the theoretical and statistical characterization of processing capacity (Townsend & Ashby, 1978, 1983). Specifically we make use of the hazard function of the RT distribution,

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)},$$

where  $f(t)$  is the probability density function (pdf) and  $S(t)$  is the survivor function of the RT distribution. This measure is one that is commonly applied in areas as diverse as reliability engineering (e.g., Cooper & McGillem, 1999) and epidemiology (e.g., Collett, 1994), where it is also known as the *intensity function* (see also Townsend & Ashby, 1978), as the variation in the hazard function characterizes the intensity with which (for example) a disease state is progressing. The integrated form of the hazard function has been of central importance in characterizing capacity in work focused on strong empirical tests of process characteristics of a variety of cognitive processes. For example, Townsend and Nozawa (1995), Wenger and Townsend (2001), and Ingvalson and Wenger (2005) used these functions to characterize

individual observers' performance in detection and identification tasks involving manipulations of target redundancy. Wenger and Gibson (2006) used a similar approach to measure effects in selective attention. (e.g., Ingvalson & Wenger, 2005; Townsend & Nozawa, 1995; Wenger & Gibson, 2004; Wenger & Townsend, 2006)

## Focusing the Question: A Model of Hippocampal Atrophy

At this point, we have established a general case for why and how capacity at a system level can be characterized, and how that characterization can allow for a principled distinction to be made between processing capacity and overall speed. We now turn to consideration of how one of the more prominent changes associated with MCI and AD can provide a concrete motivation for using these measures. Accumulating data suggests that measures of atrophy across a restricted set of brain regions is highly predictive of the conversion from MCI to AD (e.g., Fleisher et al., 2008; Jack et al., 1999; Vemuri et al., 2008; Whitwell et al., 2007, 2008a,b). Of particular interest for present purposes is hippocampal atrophy, given the critical role that hippocampal circuits play in a range of memory tasks (see, e.g., Davachi & Wagner, 2002; O'Reilly & Rudy, 2001, among others).

We created a computational representation of a basic hippocampal circuit,<sup>1</sup> shown in schematic form in Figure 1, in order to contrast measures of capacity with measures of overall speed, given varying levels of structural integrity. The model architecture was based on the established intrinsic connectivity of the hippocampus (see Shepherd, 2004). We assumed the output of the entorhinal cortex as the input to model, and assessed the output of the model at the level of the subiculum. The major pathway in the circuit was from entorhinal cortex to dentate gyrus, then to CA3, CA1, and the subiculum. In addition, CA3 and CA1 also took input directly from the entorhinal cortex.

Each of the three regions in the circuit was modeled as a network of discrete-time integrate-and-fire neurons (Amit & Brunel, 1995, 1997a,b; Dayan & Abbott, 2001). The sub-threshold dynamics of the membrane potential  $V$  for each cell  $i$  in any given region evolved over time ( $t$ ) according to

$$V_i(t+dt) = \left(1 - \frac{dt}{\tau_\alpha}\right) V_i(t) + \frac{dt}{\tau_\alpha} [I_i^{\text{rec}}(t) + I_i^{\text{ext}}(t)] \quad (1)$$

with  $\tau_\alpha$  being the membrane time constant,  $I_i^{\text{rec}}$  being the recurrent (within-region) input, and  $I_i^{\text{ext}}$  being the external input. The state of the recurrent input to each cell  $i$  at each point in time is modeled as a combination of excitatory ( $E$ ) and inhibitory ( $I$ ) inputs from all other cells  $j \neq i$ , in that region,

$$I_i^{\text{rec}}(t) = \frac{dt}{\tau_\alpha} \left\{ \sum_{j \in E} \rho_j(t) J_{ij} S_j(t - \delta_{ij}) - \sum_{j \in I} \rho_j(t) J_{ij} S_j(t - \delta_{ij}) \right\} \quad (2)$$

Here,  $\rho_j(t)$  is a reliability variable, set to 1 with probability  $r$  and 0 otherwise (see also Tsodyks & Sejnowski, 1995);  $J_{ij}$  is the relative efficiency of the synapse between cells  $j$  and  $i$  (0 for cells that are not connected);  $\delta_{ji}$  is the time required to transmit a signal from cell  $j$  to cell  $i$  and  $S_j$  indicates if a pre-synaptic spike has (= 1) or has not (= 0) occurred in cell  $j$  during the interval from  $t - \delta_{ji}$  to  $t$ .

<sup>1</sup>The modeling approach used here follows a very general algorithm outlined by Ashby and Valentin (2007), and as such could be applied to a range of questions in computational cognitive neuroscience.

Each cell in each region was assumed to possess a threshold value for the membrane potential  $V_i(t)$ , such that if  $V_i(t) \geq \gamma$  at time  $t$ ,  $V(t + dt)$  was forced to an upper limit  $V_{\max}$  (simulating a spike), and then returned to a lower limit  $V_{\min}$ . Each cell was assumed to have a resting value,  $V_0$ , which was also used as the default value for each cell at the beginning of each simulated trial. The sub-threshold dynamics of cells in CA1 were also governed by a waiting time function

$$w(t) = \lambda \left\{ 1 - \exp \left[ -\frac{s(t)}{\beta} \right] \right\} \quad (3)$$

where  $w(t)$  is the minimum number of time steps ( $dt$ ) that needed to elapse before the cell could spike,  $\lambda$  is an upper asymptote,  $\beta$  determines the slope of the function, and  $s(t)$  is the cumulative number of spikes generated by the cell by time  $t$ . This allowed for the type of spike-rate adaptation observed for cells in CA1 (Shepherd, 2004). The spiking behavior of CA3 neurons was modified such that successive spikes diminished the magnitudes of  $V_{\max}$  and  $V_{\min}$  linearly as a function of the total number of spikes generated up to time  $t$ ,

$$V_{\max}(t+dt) = V_{\max}(t) \{1 - [s(t) \times p]\}, \quad (4)$$

in order to mimic the known behavior of CA3 cells (Shepherd, 2004).

Specific values for each of the simulation parameters were fixed according to published data (see, e.g., Amit & Brunel, 1995, 1997a,b; Shepherd, 2004, and see Appendix A for the simulation methods). The number of cells in each region, and the relative number of excitatory and inhibitory connections within and between regions were also set according to published biological and simulation data.

The critical dependent variable in the simulations was the time required to reach a criterion number of action potentials at the output end of the circuit (see Gold & Shadlen, 2007; Smith & Ratcliff, 2004, for reviews of similar approaches). We then assumed that observable response times (RTs) were proportional to the time to criterion for the circuit. Our critical independent variables were those manipulated in order to simulate the effects of varying processing load on a circuit at differential levels of integrity (atrophy). We simulated processing load by varying the criterion number of action potentials (this corresponding to an assumption of limited capacity for the circuit). We simulated atrophy by reducing the total number of cells in each region.

Figure 2 plots the outcomes of all of the simulations of the model circuit, showing the relationship between the relative increase in mean RT and the relative change in the proportional hazards estimates. The open symbols show the effects on these two measures of increasing atrophy at the lowest level of processing load, and the filled symbols show the effects of load at each level of atrophy. The dashed diagonal line indicates equality of proportional change at the levels of the hazard function and the mean. The data clearly show that proportional change in the estimated hazard functions is always greater than proportional change in the mean, with this being true both for changes in load and changes in circuit integrity. These results suggest that the hazard function analyses may be superior to analyses of mean RTs in detecting changes due to deterioration of a critical portion of the overall set of circuits responsible for memory performance.

## Connecting Theory and Data

Following from the results of the simulation, and from earlier work on similar issues in other domains (e.g., Townsend & Ashby, 1978; Wenger & Gibson, 2004; Wenger & Townsend,

2006), our hypotheses regarding capacity take the form of orderings on hazard functions for individual observers in two or more conditions, defined in terms of some manipulation of workload. For example, imagine that we have the hypothesis that individuals in the early stages of Alzheimer's disease show a pronounced limitation in their capacity to process memory information, with variation in the number of retrieval cues (one vs. two) being used as the manipulation of workload. We can state this hypothesis at the level of the hazard function as

$$h_2(t) < h_1(t)$$

for all  $t > 0$ , where the subscripts indicate the number of retrieval cues presented. Note that this hypothesis, by extension, implies that

$$\frac{h_2(t)}{h_1(t)} < 1$$

with the corresponding null hypothesis being

$$H_0: \frac{h_2(t)}{h_1(t)} \geq 1$$

for all  $t > 0$ .

The advantages of this measure of capacity, relative to mean RT, are both conceptual and statistical. As discussed earlier the hazard function characterization maps onto the notion of processing capacity in a way that is much more direct than the mapping allowed by, for example, the mean RT (Townsend & Ashby, 1978; Townsend & Nozawa, 1995). There is also an inferential (statistical) advantage associated with having information from the entire RT distribution, rather than a point-estimate of central tendency (see particular Townsend, 1990; Townsend & Ashby, 1978). Specifically, Townsend (1990) has demonstrated that an ordering on hazard functions always implies an ordering on means, whereas an ordering on means does not of necessity imply an ordering on hazard functions. Finally, and pragmatically, framing hypotheses regarding capacity at the level of the hazard functions allows one to make use of a robust and very general semi-parametric regression model known as the Cox proportional hazards model (e.g., Allison, 1984, 1995; Collett, 1994; Cox, 1972; Therneau & Grambsch, 2000).

The standard Cox proportional hazards model is based on the assumption that the hazard function for an observer in any experimental condition  $i$  can be written as

$$h_i(t) = \lambda_0(t) \exp [\beta_1 x_{i1} + \dots + \beta_k x_{ik}] \quad (5)$$

where  $x_{i1}, x_{i2}, \dots, x_{ik}$  are a set of  $k$  independent variables or covariates (with corresponding weights  $\beta_{i1}, \beta_{i2}, \dots, \beta_{ik}$ ), and  $\lambda_0(t)$  is an underlying but unspecified baseline hazard. Consequently, it is not necessary to make any assumptions about the parametric form of the baseline hazard (thus getting past many of the technical concerns discussed by Luce, 1986; Van Zandt, 2000, 2002, in estimating the form of individual hazard functions). For this reason, proportional hazards models are sometimes referred to as "semi-parametric" (see, e.g., Allison, 1995; Collett, 1994; Therneau & Grambsch, 2000). All that is required of this baseline hazard is that it not be negative.



If it can be assumed that the hazard functions for the observer in any two experimental conditions  $i$  and  $j$  are proportional for all values of  $t$ , then the model for the effects of the independent variables and covariates can be simplified as follows:

$$\frac{h_i(t)}{h_j(t)} = \frac{\lambda_0(t) \exp [\beta_1 x_{i1} + \dots + \beta_k x_{ik}]}{\lambda_0(t) \exp [\beta_1 x_{j1} + \dots + \beta_k x_{jk}]} \quad (6)$$

$$= \exp [\beta_1 (x_{i1} - x_{j1}) + \dots + \beta_k (x_{ik} - x_{jk})] \quad (7)$$

Further, taking the log of this relationship results in a linear model whose parameters can be estimated using well-understood statistical methods (e.g., maximum likelihood).

The robustness of this statistical model in the face of the known challenges associated with analyzing RT distributions (e.g., Luce, 1986; Van Zandt, 2000, 2002) was initially explored in a set of simulation studies (reported in detail in Wenger et al., 2006). That work involved generating data that possessed a known proportionality relationship at the level of the hazard function, across a set of distributional forms commonly associated with RT data, given varying levels of sample size, and examining the ability of the proportional hazards model to both detect reliable orderings and recover the true magnitude of the proportionality relationship. The Weibull (1951) distribution was used as the basis for these simulations, given its ability to be parameterized so as to produce distributional shapes that are commonly used to model and analyze RT distributions (as in, e.g., Logan, 1992; Spieler & Balota, 1997; Spieler et al., 2000). In addition, these simulations were performed so as to produce two types of data: one from contexts in which enough observations are taken to allow distributional analyses at the level of the individual observer (e.g., as in Townsend & Nozawa, 1995; Wenger & Townsend, 2001), and one from contexts in which distributional analyses consider data from multiple observers (as is generally the case in studies of memory across the life span). This second context allowed us to further examine extensions to the proportional hazards model (in particular Allison, 1995, 1996; Klein, 1992; McGilchrist, 1993; Therneau & Grambsch, 2000) that take into account the unobserved heterogeneity due to multiple observations on multiple observers.

However, these initial simulations did not address a question that is of critical interest in any clinical application. Specifically, it may be the case that it is possible to obtain only a limited number of observations on any single individual. Consequently, we performed an additional set of numeric simulations to investigate the effects of varying sample size on the ability of the proportional hazards model to recover known levels of proportionality on the hazard functions. We generated data at four levels of proportionality or ordering (representing a range of experimental effect sizes), for sample sizes between five and 200 observations; details of the numeric simulations are presented in Appendix B. The results are plotted in Figure 3. Here it can be seen that there is a liberal bias in the estimated proportionality that is amplified at small sample sizes and with small effect sizes. In addition, the bias drops off rapidly and approaches asymptotic levels for all effect sizes when the sample sizes exceed 30 observations. The presence of a consistent liberal bias implies that any deleterious effects of load will be *underestimated*, meaning that estimates of capacity drawn from inferred orderings will of necessity be conservative.

## Experiment: Applying the Measures to Memory Performance

Our approach to investigating the potential for this latency-based measure of capacity involved a modification of a well-established experimental paradigm that has demonstrated significant utility in both basic and applied research, the *free and cued selective reminding test* (FCSRT, Bushke, 1984; Grober & Bushke, 1987). Before describing the specific modification to the FCSRT used in the present study, it is worth considering the nature of that paradigm and the advantages it provides in the study of memory deficits.

Studies of the information involved in encoding, consolidation, retention, and expression often focus on the relative frequency of specific types of observable events, such as items correctly recalled or recognized. However, when such measures are used in the study of changes in memory performance as a function of age and/or impairment, interpretation of these measures can be complicated by a number of potentially confounding factors. In particular, as discussed by Grober and Bushke (1987), any observed deficiencies in performance may be due to at least two contributing effects: “apparent” memory deficits, or performance decrements that are revealed in a test of memory but that are attributable to difficulties with processes outside of those associated with successful encoding, consolidation, retention, and retrieval (e.g., perceptual or motoric processes); and “genuine” memory deficits, or performance decrements that can be isolated to some aspect of encoding, consolidation, retention, and expression.

In order to distinguish between apparent and genuine deficits, it is necessary to show that a decrement in performance is observable in conditions which are optimal on all possible dimensions of processing (Bushke, 1984; Petersen et al., 1992, 1995, 1999). The necessity of these conditions is indicated by documented age-related differences in abilities related to encoding and immediate memory (e.g., Eysenck & Eysenck, 1979; Salthouse & Babcock, 1991; Waugh & Barr, 1982), in the extent to which both general (e.g., Larrabee, Youngjohn, Sudilovsky & Crook, 1993) and specific information (e.g., episodic features) are retained over time (e.g., Johnson, Leonardis, Hashtroudi & Ferguson, 1995), and in the relative effectiveness of various types of cuing conditions (e.g., Craik, 1994; Faust, Balota & Spieler, 2001; Micco & Masson, 1992; Smith, 1977).

The FCSRT was developed by Bushke and colleagues (Bushke, 1984; Grober & Bushke, 1987) in order to provide the conditions that would allow for discriminating between apparent and genuine deficits in memory performance. Figure 4 presents a flow-chart summary of the task. Initial encoding (learning) in this task is controlled by requiring semantic processing of each of the to-be-remembered (TBR) items, for example, by requiring participants to search for and verbally identify a picture of a particular object (e.g., “sweater”) among a set of pictures in response to a categorical cue (e.g., “item of clothing”). Following presentation of the entire set of TBR items, participants freely recall as many items as possible during a fixed period (e.g., 2-3 min). After the free recall period, retrieval cues (e.g., the category labels that were used to guide initial encoding) are provided for words that were not recalled. The total number of items that are retrieved using both free and cued recall provides a minimum estimate of the number of items available for retrieval. Following this, an additional set of encoding trials and free and (if necessary) cued recalls for the remaining set of items is performed. At the outset of each of these iterations, the orienting cues for items not recalled either freely or with cues on the preceding iteration are provided, with participants re-performing the search for and identification of these items, a process that acts as a form of selective reminding (Bushke, 1973).

Since the introduction of the FCSRT, its sensitivity and effectiveness in detecting age- and pathology-related changes in memory performance have been documented in a number of studies (e.g., Grober et al., 1988; Grober & Kawas, 1997; Grober et al., 1998, 2000; Petersen



et al., 1995; Tuokko & Crockett, 1989; Tuokko et al., 1991). The test is easy to administer, provides a high degree of diagnostic accuracy with respect to memory impairment (e.g., Grober & Bushke, 1987; Grober et al., 1988; Grober & Kawas, 1997; Grober et al., 2000; Petersen et al., 1999), and (due to the iterative nature of the procedure) allows for a reasonable number of samples to be obtained from each participant. While it is possible to obtain measures of response latency on the cued trials, latency data have not (to our knowledge, prior to the present study) been systematically collected and/or reported in the literature. Given the known sensitivity of latency measures for revealing important aspects of cognitive functioning that are often not detectable with frequency data (see, e.g., Luce, 1986; Townsend & Ashby, 1983; Van Zandt, 2002, for general reviews), and given the potential of the latency-based capacity measures to be sensitive to the effects of hippocampal atrophy, we augmented the FCSRT by collecting RTs to the retrieval cues in the cued portion of the paradigm.

## Method

**Participants**—Five groups of participants were recruited, three groups from the area in and around South Bend, Indiana, and two groups in and around Rochester, Minnesota. The first group (to be referred to as the CA group) included 103 healthy undergraduate students from the University of Notre Dame; data from three of these individuals were discarded prior to analysis for reasons unrelated to performance (e.g., equipment malfunctions, etc., with the most common being difficulties with the lapel microphone). The second group included 107 healthy middle-aged (25–65 years of age, to be referred to as the MA group) individuals from the South Bend community; data from three of these individuals were discarded prior to analyses. The third group included 104 healthy elderly individuals (over the age of 65, to be referred to as the HE-SB group) from the South Bend community; data from seven of these individuals were discarded prior to analyses. The fourth and fifth groups were recruited from a registry of individuals participating in the Mayo Clinic Study of Aging (see Roberts et al., 2008, for methodological and demographic details of this effort). Each of these two groups included 50 individuals. Individuals in one of these two groups had been previously assigned a diagnosis of MCI (to be referred to as the MCI group), and individuals in the other were age-matched controls (no diagnosis of MCI, to be referred to as the MCI-C group); data from two individuals in the MCI-C group was discarded prior to analyses. All participants reported fluency with spoken and written English, and all received financial compensation for their participation.

**Materials**—The items used in the memory task were 24 line-drawings from the Snodgrass and Vanderwart (1980) norms (see Table 1). For each of these drawings, a category label and a highly-associated instance were selected from the McEvoy and Nelson (1982) category and instance norms. The images were drawn in black on white squares, 6.4 cm on each side, so that at an unconstrained viewing distance of approximately 70 cm, the squares subtended approximately 5.2° of visual angle. Stimulus presentation and data recording were done using either a desktop (for the CA, MA, and HE-SB participants) or laptop (for the MCI-C and MCI participants) computer. Screen resolutions in both cases were set to 800 × 600 pixels, at a refresh rate of 17 ms. Manual responses were collected and timed using the computer keyboards, and verbal responses were timed using a lapel microphone attached to the computers; timing of responses in all cases was accurate to ±1 ms (Forster & Forster, 2003).

## Design

The overall design was cross-sectional, with the groups (CA, MA, HE-SB, MCI-C, MCI) defined by age and MCI status. With the exception of the FCSRT, none of the experimental tasks—visual threshold estimation, naming time, FCSRT—involved any other between-subjects manipulation. In the case of the FCSRT, the independent variables were number of cues in the cued recall portion (one or two, manipulated within-subjects), and the specific type of cue used on the single-cue trials (category label or instance, manipulated between-subjects).

Half of all subjects in each group were randomly assigned to each of the single-cue-type conditions.

**Procedure**—Testing sessions for the CA, MA, and HE-SB participants took place in a quiet, dimly-lit room in the first author's laboratory on the campus of the University of Notre Dame. Testing sessions for the MCI-C and MCI participants took place in the Alzheimer's Disease Research Center, Mayo Clinic College of Medicine, in Rochester, MN. Sessions lasted between 45 and 90 min, with variations in duration due to longer sessions for the older participants.

**Background data:** Each session began with participants completing a background questionnaire consisting of a total of 21 questions, covering basic demographic and health information, along with quality-of-life items from the Health and Retirement Survey (Gustman, Mitchell & Steinmeier, 1995). Items were presented sequentially in a fixed order in plain text on the computer screen, and participants were given the option of either reading and entering their responses on their own, or having the items read to them with responses entered by the experimenter.

**Visual identification thresholds:** Following this, participants' visual identification thresholds (as a function of exposure duration) were estimated using the method of constant stimuli. This was done in order to provide measures that could help distinguish latency-based effects specific to memory processing from general age-related slowing. The range and specific values for the set of seven exposure durations varied for each of the groups, with the specific values selected on the basis of pilot testing. The exposure durations for the CA and MA participants ranged from 17 to 119 ms in 17 ms steps. The exposure durations for the HE-SB and MCI-C participants ranged from 34 to 238 ms in 34 ms steps. The exposure durations for the MCI participants ranged from 68 to 476 ms in 68 ms steps. A total of 20 trials were presented at each level of exposure duration, and order of presentation was determined randomly for each participant.

Trials began with a fixation stimulus, presented centrally for between 510 and 1020 ms, with the specific duration determined by a uniform distribution over this interval (rounded to the closest multiple of 17 ms). This was followed by the presentation of either the target ([ + ]) or non-target ([ - ]) stimulus for the appropriate exposure duration. Participants signaled their perception of the target stimulus by pressing either the "z" or "/" key on the computer keyboard, using the index finger of their dominant hand, and the non-target stimulus using the middle finger of their dominant hand. Feedback (the word "Correct" or the word "Incorrect," presented centrally) was provided on each trial.

**Naming time:** Participants then performed a simple naming-time task. This task, like the visual identification threshold task, was included in order to help differentiate latency-based effects specific to memory processing from generalized age-related slowing. A set of 20 three- or four-letter words was chosen from the labels for items in the Snodgrass and Vanderwart (1980). These words were chosen on the basis of pilot testing to be minimally confusable with the set of items used in the FCSRT. Each word was presented once only, and the order of presentation was determined randomly for each participant.

Trials began with a fixation screen presented centrally for between 510 and 1020 ms, with the specific duration determined by a uniform distribution over this interval (rounded to the closest multiple of 17 ms). This was followed by the presentation of one of the words, in upper-case letters, with the word remaining visible until the participant responded. Participants responded verbally, speaking at a normal (conversational) volume, triggering the software voice key. The experimenter logged each trial as correct, incorrect, or artifact (cough, etc.). No feedback was provided.

**FCSRT:** The final task was the FCSRT. The encoding portion of this task required participants to find a target image located in a display involving three randomly selected non-target images on the basis of both a category cue and a related instance cue. The cues were presented in sentence form at the top of each display of four images, and took the form “Find the *category label* that is related to *instance*.” The location of the target image was determined randomly on each encoding trial. Participants responded by pointing to the target item, and the display remained available until the participant responded. Each encoding trial was initiated and terminated by the experimenter.

Once the entire list of target items had been presented, participants were asked to recall, verbally, as many of the encoded items as possible, with no requirement for serial recall. The experimenter logged the items correctly recalled, then initiated the cued recall trials. Each cued recall trial began with the presentation of either one or two cues at the top of the screen. On the double-cue trials, the instance and category cues were separated by eight spaces and a pair of upright brackets centered in those eight spaces. The specific location of each of the cues was determined randomly on each trial, and the number of single- and double-cue trials was set to be approximately equal for each participant on the basis of the number of words available for cued recall following the first free recall. The cue(s) were available until the participant responded. Participants responded verbally, triggering a software voice key. The RT for the trial was the time from the onset of the retrieval cue(s) to the onset of the participant's vocal response. Following each response, the experimenter logged the response as being correct, incorrect, or an artifact (cough, etc.). Each of the items that had not been recalled in the free recall trials was cued once, with order determined randomly. If items remained not-recalled following the cued recall trials, those items were used in another set of encoding, free recall, and cued recall trials. This sequence was repeated either until all of the items were recalled, or a total of six replications of the sequence were completed.

## Results

Unless otherwise indicated, all analyses used an  $\alpha$  of 0.05. We begin with a summary of the basic characteristics of individuals in each group (including visual identification thresholds and mean naming times), then consider the frequency- and latency-based measures of memory performance, including (for a subset of the participants) a consideration of the relationship between these measures and measures of hippocampal integrity.

### Participant characteristics

Table 2 presents a summary of a relevant subset of the information reported on the background questionnaire. At least 93% of all participants in each group reported being native speakers of English and, overall, the percentage of total age during which participants reported speaking English was at least 97% for all groups. In addition to the results reported here and in the sections following, we performed analyses in which we excluded (a) participants reporting head injury, (b) participants reporting depression, and (c) all females. The qualitative results did not differ in any way from those obtained using all of the participants' data. In addition, we estimated pairwise correlations among a subset of the demographic and individual difference variables (age, years of education, years of spoken English, visual identification threshold) and a subset of the critical dependent measures (cumulative free recall, cumulative cued recall, total change in recall, mean RT on cued recall, and the individual estimate of capacity) within each group. None of these correlations reached criterion for statistical reliability.

### Visual identification thresholds

Thresholds for each individual were estimated by fitting a two-parameter Weibull cdf

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\beta}\right)^\alpha\right] \quad (8)$$

to identification accuracy as a function of exposure duration ( $t$ ), then finding the value of  $t$  for which the parameterized function was equal to 0.79. Figure 5(a) and (b) present the group means and individual estimates as a function of age for the entire sample. Individual estimates were analyzed using a one-way analysis of variance (ANOVA), with age as the single factor. This analysis revealed a reliable effect due to group status,  $F(4, 397) = 34.70$ ,  $MSE = 5994$ , with the temporal thresholds for the CA (64 ms) and MA (87 ms) participants being reliably shorter than those for the HE-SB (177 ms), MCI-C (143 ms), and MCI (165 ms) participants, and there being no reliable differences (based on Tukey post-hoc tests) between the CA and MA participants, or among the HE-SB, MCI-C, and MCI participants.

### Naming times

Panels (c) and (d) of Figure 5 present the group means and individual means on the naming as a function of age for the entire sample. Individual means were analyzed using a one-way ANOVA, with group as the single factor. This analysis revealed a main effect for group,  $F(4, 397) = 29.77$ ,  $MSE = 208008$ , with naming latencies being reliably shortest for the CA participants (489 ms), followed by the MA (562 ms) and HE-SB (617 ms) participants, with mean latencies for the MCI-C (716 ms) and MCI (731 ms) being the longest, with there being no reliable differences within each grouping (MA vs. HE-SB, and MCI-C vs. MCI).

### FCSRT: Free and cued recall

Participants in all five groups completed recall by, at the latest, the fourth pass through the procedure (see panel (a) of Figure 6), with the majority of the improvement occurring between the first and second pass. Consequently, we analyzed all of the data from the FCSRT with respect to differences between initial and asymptotic levels of performance.

The data for change in cumulative free recall (initial to asymptotic) were analyzed using a 5 (group: CA, MA, HE-SB, MCI-C, MCI)  $\times$  2 (single-cue condition: category, instance) ANOVA, collapsing across total number of cues. The only effect to reach the criterion level for reliability was the main effect of group  $F(4, 389) = 10.85$ ,  $MSE = 13.96$ : participants in the MCI group on average increased cumulative free recall by an amount (1.8 items) that was reliably greater than that for individuals in the CA (0.6 items), MA (1.0 items), HE-SB (1.1 items) or MCI-C (0.6 items) groups, with there being no reliable differences among these latter four groups. Post-hoc examination of the group means showed all of the increases in cumulative free recall to be reliably different from 0.0, all  $t_s > 4.74$ , all  $p_s < .001$ .

A similar result was obtained in the analysis of the data for change in cumulative cued recall (collapsing across number of cues), which were analyzed using a 5 (group: CA, MA, HE-SB, MCI-C, MCI)  $\times$  2 (single-cue condition: category, instance) ANOVA. As with the cumulative free recall data, the only effect to reach criterion for reliability was the main effect for group,  $F(4, 389) = 14.19$ ,  $MSE = 0.35$ . Participants in the MCI group increased their cumulative cued recall by 0.7 items (on average), reliably more than was the case for participants in the CA (0.1 items), MA (0.2 items), HE-SB (0.2 items), or MCI-C (0.2) items, with there being no reliable differences among the latter four groups. In addition, only the increase for the MCI participants was reliably different from 0,  $t = 4.78$ ,  $p < .001$ , all other  $t_s < 1.77$ , all other  $p_s > .09$ .

Unsurprisingly, then, the analysis of the data for change in cumulative total (free plus cued) recall revealed the same basic result. These data were analyzed using a 5 (group: CA, MA,

HE-SB, MCI-C, MCI)  $\times$  2 (single-cue condition: category, instance) ANOVA. As in the preceding analyses, only the main effect of group reached the criterion level for reliability,  $F(4, 389) = 16.36$ ,  $MSE = 20.52$ , with the overall increase for participants in the MCI group (2.5 items) being reliably greater than was the case for participants in the CA (0.7 items), MA (1.2 items), HE-SB (1.3 items), or MCI-C (0.8 items) groups, with there being no reliable differences among the latter four groups, and with all five of the means being reliably greater than 0. Our data are, then, at this level of analysis, consistent with established findings in the literature with respect to differences in performance on the FCSRT—particularly deficits on the initial pass with associated increases on subsequent passes—being reliable indicators of MCI (e.g. Grober et al., 2000; Kelley & Petersen, 2007; Petersen, 2007; Petersen & Negash, 2008).

### FCSRT: Cued recall, mean RT

Systematic effects at the level of overall speed of processing were examined by way of mean RTs on the cued recall trials. Prior to analyses, all RTs less than 200 ms were deleted, as were all RTs associated with incorrect responses. Median RTs were calculated for each participant at each level of cuing (one vs. two), and the data (means of those medians) were then analyzed using a 5 (group: CA, MA, HE-SB, MCI-C, MCI)  $\times$  2 (single-cue condition: category, instance)  $\times$  2 (number of cues: one vs. two) ANOVA, with group and single-cue condition being between-subjects variables, and number of cues being a within-subjects factor. The data are summarized in the three panels of Figure 7.

The analyses of these data are summarized in Table 3. Critically, the three-way interaction of group, single-cue condition, and number of cues was reliable. To understand the form of this interaction, we calculated the following interaction contrast (Maxwell & Delaney, 2004) for the interaction of single-cue condition and number of cues, separately for each group:

$$IC = \left( \bar{RT}_{c_1} - \bar{RT}_{c_2} \right) - \left( \bar{RT}_{i_1} - \bar{RT}_{i_2} \right) \quad (9)$$

The subscripts indicate the type of cue used on the single-cue trials (category, C, or instance, I) and the number of cues present on the trial (one or two). Thus,  $\bar{RT}_{c_1}$  and  $\bar{RT}_{c_2}$  are the mean RTs for single- and double-cue trials for those individuals who were presented with category cues on the single-cue trials. The values of the interaction contrast are plotted in panel (d) of

Figure 7. The first difference on the right-hand side of Equation 9 ( $\bar{RT}_{c_1} - \bar{RT}_{c_2}$ ) was generally small and negative, reflecting little in the way of cost of adding an instance cue to a category cue (see panel (a) of Figure 7). The exception was the MCI group, which showed a reliable benefit of adding an instance cue to the category cue. The second difference on the right-hand side of Equation 9 ( $\bar{RT}_{i_1} - \bar{RT}_{i_2}$ ) was generally negative, but larger in magnitude than the first difference (see panel (b) of Figure 7), indicating an overall cost of adding a category cue to an instance cue. In contrast, the difference for the MCI group was quite large. Taken together the interaction contrast indexes the extent to which the cost of adding a category cue to an instance cue exceeded either the cost or the benefit adding an instance cue to a category. As can be seen in panel (d) of Figure 7, this cost was largest for the MCI group.<sup>2</sup>

### FCSRT: Effect of load on capacity

To estimate the effects of varying the number of cues (workload) on cued recall RT, we fit the proportional hazards model to the correct RTs for each observer for each combination of group status and single-cue condition separately, with number of cues as the predictor.<sup>3</sup> The results



of these analyses are presented in Table 4 and Figure 8. Consistent with the analyses of the mean RTs, adding a category cue to an instance cue produced a reliable reduction in capacity for participants in all five groups. However, only the MCI group showed a reliable effect of adding an instance cue to a category cue, in this case producing a reliable increase in capacity. Thus, at this level of analysis, participants in the MCI condition appear to be much more sensitive to the effect of varying processing load, showing both negative and positive effects for increasing the number of cues.

### Relative diagnostic utility

The analyses of the mean RTs and the capacity estimates suggest that the latency information may be of equal if not more importance in distinguishing the MCI group from the others. Specifically, it was the case that the RT data revealed an effect of cuing that was not evidenced in the free and cued recall frequency data. Consequently, we sought to evaluate the relative diagnostic utility of the latency measures by using discriminant function analysis (DFA, e.g. Rao, 1973). DFA uses one or more predictor (independent) variables to estimate a function that is used to assign each observation to one or more categories (values of the dependent variable). The specific goal of this analysis was to determine how well the capacity measures would do, relative to the acknowledged ability of the free and cued recall measures, in correctly classifying participants as to their MCI status.

We considered the data only for the HE-SB, MCI-C, and MCI groups for this analysis.<sup>4</sup> We evaluated the relative ability of models formed by five different combinations of three predictor variables to correctly categorize individuals as to their MCI status. The three predictor variables were each observer's (a) total change in free recall (initial to asymptotic), (b) mean RT on the cued recall trials (collapsing across cue type and number), and (c) absolute value of the change in capacity as a function of adding a second cue. Preliminary analyses indicated that the overall mean RT produced model fits that were statistically no different from those obtained using means specific to cue number and type. The absolute (versus signed) value of the estimate capacity change was used given the differential effect of the two types of cues.

These three predictor variables were used to parameterize five different DFA models, with the relationships among the models illustrated schematically in Figure 9. The models were fit to the data assuming a linear form for the discriminant function, and the quality of the fit of each obtained model was assessed in two ways, first with respect to a general set of measures of model performance, and second using a set of measures of diagnostic sensitivity and utility. Both methods of assessment used the predicted set of categorization frequencies and probabilities as their basis.

The quality of each model's fit to the data was first quantified by calculating the log likelihood

$$\ln(L) = \sum_{i=0}^1 \sum_{j=0}^1 f_{ij} \ln \left[ P(R_i | S_j) \right] \quad (10)$$

<sup>2</sup>A reviewer astutely noted that the two cuing conditions differ in one potentially critical way, other than in terms of workload. Specifically, it is possible that presentation of the instance cues recruited processing of memory for sensory features, while presentation of the category cues recruited processing of memory for functional or abstract features (following, e.g., Rosch, Mervis, Gray, Johnson & Boyesbraem, 1976). Consequently, it is possible that adding a demand to process abstract features has more of an impact than does adding a demand to process sensory features. Critically, this possibility allows for a very easy experimental test (also suggested by the reviewer), involving two category cues that vary in the extent to which they tap instance-, basic-, or superordinate-level representations.

<sup>3</sup>We also estimated models that included visual identification thresholds, mean naming times, and age as predictors. However none of these variables produced parameter estimates that were reliably different from 0 in any of the cases. The relative estimates for the effect of number of cues did not change in their orderings with and without the additional predictors.

<sup>4</sup>We also ran the DFA analyses using only the data for the MCI-C and MCI groups, and obtained qualitatively identical results.



where  $f_{ij}$  was the observed frequency of individuals with known MCI status  $j$  being assigned status  $i$  ( $i, j = 0, 1$ ), and  $P(R_i|S_j)$  was the overall probability of the model assigning MCI status  $i$  to an individual with MCI status  $j$ . The log likelihood was used to compare the relative performance of the five models, following a strategy outlined by Thomas (2001). Models that differed in terms of the total number of predictor variables were compared using

$$\chi^2 = -2 \ln \left( \frac{L_r}{L_g} \right) = 2 \left[ \ln(L_g) - \ln(L_r) \right] \quad (11)$$

where  $L_r$  is the log likelihood for the model with the smaller number of predictor variables (the restricted model) and  $L_g$  is the log likelihood for the model with the larger number of predictor variables (the general model). The degrees of freedom for the comparison is given by the difference in the number of predictor variables for the two models under consideration. Models that had the same number of predictor variables were compared using the Akaike Information Criterion

$$\text{AIC} = -2 \ln(L) + 2v \quad (12)$$

(Akaike, 1983; Ashby & Lee, 1991), where  $v$  is the total number of predictor variables for the model under consideration. The model with the smallest value of the AIC was selected as providing the best account of the data. Table 5 summarizes this analysis. Among the single-variable models, the one based on the capacity measure provided the superior account of the data, while between the two two-variable models, the one that combined the capacity measure with total change in recall provided the best account of the data. Finally, the model that combined the capacity and total recall measures provided a statistically superior account of the data, relative to the models that relied on either of these two measures alone.

We also contrasted the performance of the five sets of predictor variables using a set of measures reflecting diagnostic (or choice) accuracy, sensitivity, and bias (see Table 6). For the purposes of these comparisons, we coded the rate of correct and incorrect categorizations for each predictor as follows. A “hit” was defined as correctly classifying an individual with a diagnosis of MCI, with a “miss” being the associated error. A “correct rejection” was defined as correctly classifying an individual who did not have a diagnosis of MCI, and a “false alarm” was the associated error.

The results of the DFA are summarized in panels (a)-(c) of Figure 10, and show that, generally, the capacity measure when used either alone or in combination with total change in recall, performed well with minimal bias. With respect to diagnostic sensitivity and specificity (Figure 10(a)), the capacity measure alone performed as well as the total change in recall in terms of sensitivity (true positive), and outperformed total change in recall in terms of specificity (true negative). The best performance on both measures was obtained when the capacity and recall measures were combined. This same pattern held for the relationship between the measures of positive and negative predictive value (Figure 10(b)). Finally, the superior categorization performance of the model that combined change in total recall with capacity can be seen in the values of  $d'$  and  $c$  (Figure 10(c)): this model gives the highest value for the sensitivity parameter with minimal levels of bias.

### Capacity estimates and hippocampal volumes

Structural MRI scans, obtained within one month of the behavioral testing, were available for total of 36 individuals in the MCI-C (17) and MCI (19) groups. Of these, data for two individuals (both in the MCI-C group) were excluded, as their behavioral data was incomplete

(see the “Participants” section above). For the remaining 34, we obtained estimates of total hippocampal volume in order to determine the potential relationship of each of our dependent measures with a measure of hippocampal atrophy.

Hippocampal volumes were estimated using voxel-based morphology methods, as part of a larger ongoing research effort; parameters for image acquisition and analysis are provided in reports pertinent to that work (e.g., Whitwell et al., 2008b), with a brief summary reported here. In order to quantify total hippocampal volume, the borders of the left and right hippocampi were traced sequentially from posterior to anterior. In-plane hippocampal anatomic boundaries were defined to include the CA1 through CA4 sectors of the hippocampus proper, the dentate gyrus, and subiculum. The posterior boundary was determined by the oblique coronal anatomic section on which the crura of the fornices were identified in full profile. The inferior boundary of the hippocampus was determined by the gray-white interface formed by the subiculum and underlying parahippocampal gyrus.

Simple linear correlations were calculated between total hippocampal volume, each participant's age, and each of the predictor variables considered in the DFA. The results are summarized in Table 7. Here it can be seen that the only correlations to be reliably different from 0 were those for total change in recall and for the capacity estimate, and then only in the case of the individuals in the MCI group. In addition, the magnitude of the correlation between capacity and hippocampal volume was reliably greater than the correlation between total recall and hippocampal volume ( $Z = 2.12, p < .05$ ). The correlation between change in recall and hippocampal volume is consistent with the findings, here and in a range of other studies, suggesting that changes in FCSRT performance can reliably distinguish individuals with a diagnosis of MCI from those without. And the correlation between the capacity measure and hippocampal volume, particularly given the lack of a reliable correlation for the mean RT measures, suggests the potential importance of considering the effects of atrophy in terms of capacity rather than speed of processing.

## Discussion

The motivation for the present investigation was the question of whether a measure of processing capacity (Townsend & Ashby, 1978) could offer advantages in distinguishing individuals with a diagnosis of MCI from those without. We considered this question in the context of known findings regarding hippocampal atrophy, and used a computational model of the hippocampal circuit to demonstrate that measures of capacity showed more change with degeneration than did measures of speed of processing.

In order to test the ability of measures of capacity to identify individuals with a diagnosis of MCI, we modified a behavioral task that has repeatedly demonstrated its utility in distinguishing memory impairment from normal age-related changes in memory, the FCSRT. Specifically, we collected RTs on cued recall trials, using a voice key to allow for verbal responses. In addition, we also collected visual identification thresholds (as a function of exposure duration) and mean naming times, in order to provide a set of covariates that could help us distinguish deficits in recall from more general slowing of processing.

At the level of free and cued recall, we replicated standard findings in the literature. Participants in the MCI group showed a deficit in recall on the initial pass of the task, and consequently showed the largest change from initial to final levels of recall. The analysis of mean RTs in the cued recall task also showed differences specific to participants in the MCI group, with the impact of increasing recall cues from one to two being most pronounced (in terms of both facilitation and inhibition) for participants in the MCI group. Importantly, this effect of cuing

was not observed in the recall frequency data, suggesting that the latency data may be relatively more sensitive in refining the characterization of specific deficits in MCI.

The effects of adding recall cues (increasing load in cued recall) with respect to processing capacity was assessed by estimating proportional changes in the hazard function of the RT data for each participant. These analyses documented that the manipulation of load had reliable positive and negative impacts on processing capacity, selective to participants in the MCI group. Having demonstrated that the capacity measures can be used in this context, and that they are selectively sensitive to differences in performance, we then compared the capacity measures to a set of alternative predictors of MCI status, and demonstrated the general superiority of the capacity measures, in terms of sensitivity, specificity, and minimal selection bias. Finally, we also showed that the capacity measures alone showed a reliable relationship to total hippocampal volume, with this being specific to the individuals diagnosed with MCI.

We find these outcomes to be very promising, though admittedly preliminary. The approach taken in the present study was easily implemented on a lap-top computer, and as a consequence could be taken into the community. The capacity estimates were done at the level of the individual participant, with that participant providing their own baseline data. This allows the effect of load to be distinguished from age-related slowing. This also allows the potential for positively-biased estimates of proportionality due to small sample size to be in some sense “canceled out” at the level of the individual. Finally, the diagnostic sensitivity and specificity of the capacity estimates, and the reliable relationship between the capacity measures and hippocampal volume, relative to the effects at the level of mean RT, are consistent with the hypothesis drawn from the computational simulations, that hippocampal atrophy is expressed most strongly in behavior in terms of processing capacity (“bandwidth,” or information processed per unit time) rather than in terms of overall speed of processing.

Much, of course, can be done to follow-up on these preliminary findings. The numerical simulations reported in the introduction suggest the need to create experimental and diagnostic situations in which larger sample sizes can be obtained. This can easily be done by increasing list length. The sample size (for the MCI-C and MCI) groups were small for this initial investigation, and the results should be replicated with much larger groups. The differential effects of cuing condition need to be systematically investigated, with particular stress on cuing conditions that differentially require episodic information (e.g., Davachi & Wagner, 2002; Davachi, Mitchell & Wagner, 2003). Finally, it seems important to validate the relationship between changes in hippocampal volume and changes in processing capacity in a longitudinal design.

Characterizing processing capacity in a disciplined, theoretically-motivated manner, has been repeatedly suggested as being foundational to understanding the foundational characteristics of human information processing (Townsend & Ashby, 1978, 1983; Townsend & Nozawa, 1995; Townsend & Wenger, 2004a,b). We hope to have demonstrated here the critical role that characterizing capacity can have in one specific application, and as a consequence demonstrated how foundational theory in human information processing can naturally inform applied research.

## Acknowledgments

This work was supported by funding from the National Institute on Aging, 1 R03 AG022680-01. Sincere thanks are due to Karyn Dionisio for her incredible efforts in coordinating data collection, and to two reviewers for their detailed and helpful comments.

## Appendix A: Simulation Methods, Model of the Hippocampal Circuit

The model circuit represented in Figure 1 was parameterized using published values in the literature (specifically Shepherd, 2004), including those used in previous simulations of similar architectures (specifically Amit & Brunel, 1995, 1997a,b). Thus, the only free parameters were the total number of cells in each of the regions (dentate gyrus, CA3, and CA1) and the criterion number of spikes at the output end of the circuit needed to infer completion of processing. Table 8 lists all of the parameters and their values. Each trial of the simulation involved a 200 ms presentation of input, followed by 200 ms of no input. Equations 1 and 2 describe the (discrete-time) sub-threshold dynamics for each neuron in each region, and Equations 3 and 4 describe spike-dependent modifications for CA3 and CA1 neurons, respectively.

Each trial in each replication, at each level of load and each level of atrophy, began by first setting the values of the synaptic efficiency matrix  $J$  using independent samples from normal distributions with means  $J_{ij}$  and standard deviations  $\sigma(j)$ . Sparseness was introduced in each matrix by setting all values less than  $s$  to 0. External inputs (from entorhinal cortex to each of the other regions) were randomly set on each trial using independent samples from normal distributions with mean  $I_i^{\text{ext}}$  and standard deviation  $\sigma(I_i^{\text{ext}})$ . The input was stepped from 0 to the specific value of  $I_i^{\text{ext}}$  for each cell  $i$  at the beginning of each trial, and remained on for  $t_+$  ms, with the trial continuing for another  $t_-$  ms. Cumulative counts of total spikes at the output end of the circuit (subiculum) were recorded at each value of  $t$ , and the time required to reach the criteria  $\gamma_L$  and  $\gamma_H$  were recorded.

## Appendix B: Numeric Methods, Effects of Sample Size and Effect Size on Proportional Hazards Estimates

In order to assess the level of bias in estimates of proportionality obtained using the Cox proportional hazards model, we simulated samples RT measurements that varied in size and in the magnitude true proportional relationship ( $r$ ) between the underlying hazard functions

$$r = \frac{h_1(t)}{h_2(t)}$$

We assumed a two-parameter Weibull distribution in all cases, with cumulative distribution function

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\beta}\right)^\alpha\right]$$

and hazard function

$$h(t) = \frac{\alpha}{\beta} \left(\frac{t}{\beta}\right)^{\alpha-1}$$

(Blumenfeld, 2001), where  $\alpha$  and  $\beta$  are shape and scale parameters, respectively. For each level of proportionality

$$r = \frac{h_1(t)}{h_0(t)}$$

where  $h_0(t)$  was parameterized as the reference distribution, and  $h_1(t)$  was parameterized to give the desired value of  $r$ , we generated  $n$  observations, and repeated this  $N$  times. The simulated data were then submitted to the Cox proportional hazards analysis to obtain an estimate of the proportionality  $\hat{r}$ . Bias of the estimate was calculated as  $\hat{r} - r$  for each replication, and the bias estimates were averaged over replications. Table 9 lists all simulation parameters and their values.

## References

- Akaike H. Information measures and model selection. *Bulletin of the International Statistical Institute* 1983;50:277–290.
- Allison, PD. *Event history analysis*. Sage; Beverly Hills, CA: 1984.
- Allison, PD. *Survival analysis using the SAS system: A practical guide*. SAS Institute, Inc; Cary, NC: 1995.
- Allison PD. Fixed-effects partial likelihood for repeated events. *Sociological Methods and Research* 1996;25:207–222.
- Amit DJ, Brunel N. Learning internal representations in an attractor neural network with analogue neurons. *Network: Computation in Neural Systems* 1995;6:359–388.
- Amit DJ, Brunel N. Dynamics of a recurrent network of spiking neurons before and after learning. *Network: Computation in Neural Systems* 1997a;8:373–404.
- Amit DJ, Brunel N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex* 1997b;7:237–252. [PubMed: 9143444]
- Ashby FG, Lee WW. Predicting similarity and categorization from identification. *Journal of Experimental Psychology: General* 1991;120:150–172. [PubMed: 1830609]
- Ashby, FG.; Valentin, VV. Computational cognitive neuroscience: Building and testing biologically plausible computational models of neuroscience, neuroimaging, and behavioral data. In: Wenger, MJ.; Schuster, C., editors. *Statistical and process models for cognitive neuroscience and aging*. Erlbaum; Mahwah, NJ: 2007. p. 15–58.
- Birren JE. Translations in gerontology—from lab to life: Psychophysiology and the speed of response. *American Psychologist* 1974;29:808–815. [PubMed: 4440935]
- Blumenfeld, D. *Operations research calculations handbook*. CRC Press; Boca Raton, FL: 2001.
- Boring, EG. *History of experimental psychology*. Appleton-Century-Crofts; New York: 1957.
- Bushke H. Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior* 1973;12:543–550.
- Bushke H. Cued recall in amnesia. *Journal of Clinical Neuropsychology* 1984;6:433–440. [PubMed: 6501581]
- Cerella J, Hale S. The rise and fall of information processing rates over the life span. *Acta Psychologica* 1994;86:109–197. [PubMed: 7976466]
- Collett, D. *Modeling survival data in medical research*. Chapman and Hall; London: 1994.
- Cooper, GR.; McGillem, CD. *Probabilistic methods of signal and system analysis*. Oxford; New York: 1999.
- Cox DR. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society* 1972;B34:187–220.
- Craik FIM. Memory changes in normal aging. *Current Directions in Psychological Science* 1994;3:155–158.
- Davachi L, Mitchell JP, Wagner AG. Multiple routes to memory: Distinct medial temporal lobe processes build item and source memory. *Proceedings of the National Academy of Sciences* 2003;100:2157–2162.

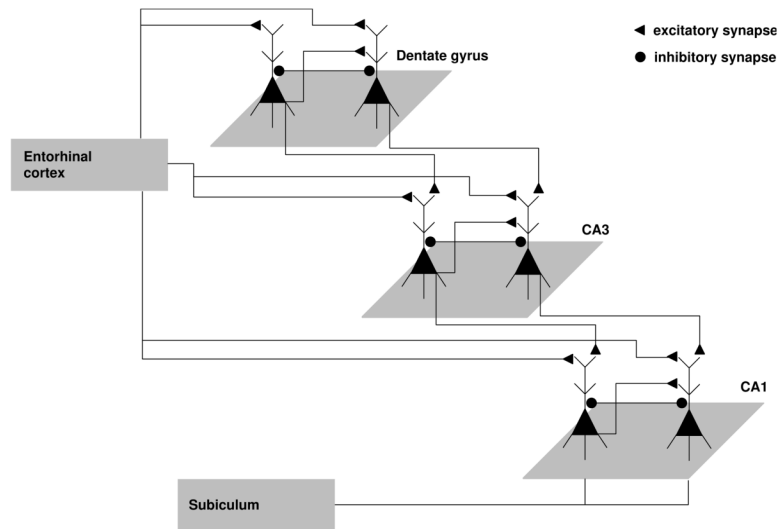
- Davachi L, Wagner AG. Hippocampal contributions to episodic encoding: Insights from relational and item-based learning. *Journal of Neurophysiology* 2002;88:982–990. [PubMed: 12163547]
- Dayan, P.; Abbott, LF. *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT; Cambridge, MA: 2001.
- Dickerson BC, Sperling RA, Hyman BT, Albert MS, Blacker D. Clinical prediction of Alzheimer disease dementia across the spectrum of mild cognitive impairment. *Archives of General Psychiatry* 2007;64(12):1443–1450. [PubMed: 18056553]
- Ebbinghaus, H. *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Duncker and Humblot; Leipzig: 1885.
- Eysenck MW, Eysenck MC. Processing depth, elaboration of encoding, memory stores, and extended memory capacity. *Journal of Experimental Psychology: Human Learning and Memory* 1979;5:472–484.
- Faust ME, Balota DA, Spieler DH. Building episodic connections: Changes in episodic priming with age and dementia. *Neuropsychology* 2001;15:626–637. [PubMed: 11761052]
- Fisher DL. Central capacity limits in consistent mapping, visual search tasks: Four channels or more? *Cognitive Psychology* 1984;16:449–484. [PubMed: 6509895]
- Fisher DL, Glaser R. Molar and latent models of cognitive slowing: Implications for aging, dementia, depression, development and intelligence. *Psychonomic Bulletin & Review* 1996;3:458–480.
- Fleisher AS, Sun S, Taylor C, Ward CP, Gamst AC, Petersen RC, Jack CR, Aisen PS, Thal LJ. Volumetric MRI vs clinical predictors of alzheimer disease in mild cognitive impairment. *Neurology* 2008;70(3):191–199. [PubMed: 18195264]
- Forster KI, Forster JC. DMDX: A windows display program with millisecond accuracy. *Behavioral Research Methods, Instruments, and Computers* 2003;35:116–124.
- Gold JJ, Shadlen MN. The neural basis of decision making. *Annual Review of Neuroscience* 2007;30:535–574.
- Grober E, Bushke H. Genuine memory deficits in dementia. *Developmental Neuropsychology* 1987;3:13–36.
- Grober E, Bushke H, Crystal H, Bang S, Dresner R. Screening for dementia by memory testing. *Neurology* 1988;38:900–903. [PubMed: 3368071]
- Grober E, Kawas C. Learning and retention in preclinical and early Alzheimer's disease. *Psychology and Aging* 1997;12:183–188. [PubMed: 9100279]
- Grober E, Lipton RB, Hall C, Crystal H. Memory impairment on free and cued selective reminding predicts dementia. *Neurology* 2000;54:827–832. [PubMed: 10690971]
- Grober E, Lipton RB, Katz M, Sliwinski M. Demographic influences on free and cued selective reminding in older persons. *Journal of Clinical and Experimental Neuropsychology* 1998;20:221–226. [PubMed: 9777476]
- Gustman A, Mitchell OS, Steinmeier TL. Retirement measures in the health and retirement study. *Journal of Human Resources* 1995;30(Supplement S):S57–S83.
- Ingvalson EM, Wenger MJ. A strong test of the dual mode hypothesis. *Perception & Psychophysics* 2005;67:14–35. [PubMed: 15912870]
- Jack CR, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, Boeve BF, Waring SC, Tangalos EG, Kokmen E. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology* 1999;52(7):1397–1403. [PubMed: 10227624]
- Jastrow, J. *The time relations of mental phenomena*. Hodges; New York: 1890.
- Johnson MK, Leonardis DMD, Hashtroudi S, Ferguson SA. Aging and single versus multiple cues in source monitoring. *Psychology and Aging* 1995;10:507–517. [PubMed: 8749578]
- Kail R. Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin* 1991;109:490–501. [PubMed: 2062981]
- Kail R, Salthouse TA. Processing speed as mental capacity. *Acta Psychologica* 1994;86:199–225. [PubMed: 7976467]
- Kandel, ER.; Schwartz, JH.; Jessell, TM. *Principles of neural science*. 4th edition. McGraw-Hill/Appleton & Lange; New York: 2000.
- Kantowitz BH, Knight JL. On experimenter-limited processes. *Psychological Review* 1976;83:502–507.



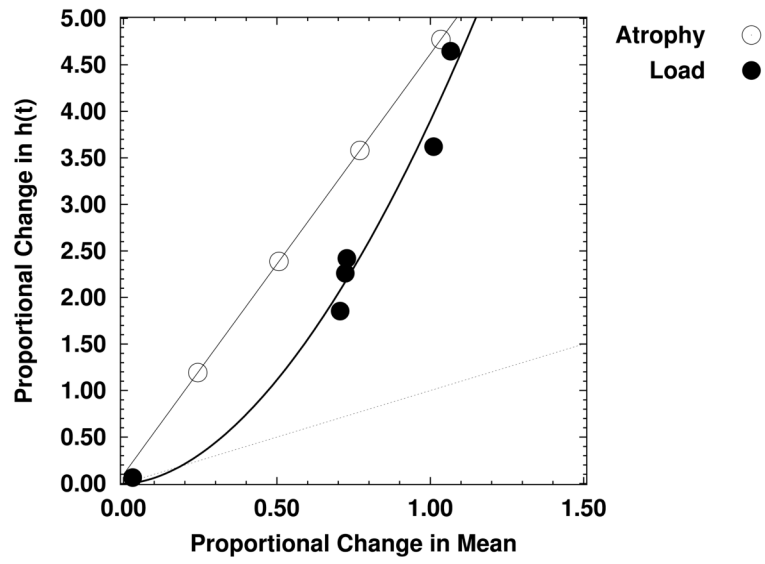
- Kelley BJ, Petersen RC. Alzheimer's disease and mild cognitive impairment. *Neurologic Clinics* 2007;25(3):577. [PubMed: 17659182]
- Klein JP. Semiparametric estimation of random effects using the cox model based on the em algorithm. *Biometrics* 1992;48:795–806. [PubMed: 1420842]
- Kolers PA, Roediger HL. Procedures of mind. *Journal of Verbal Learning and Verbal Behavior* 1984;23:425–449.
- Larrabee GJ, Youngjohn JR, Sudilovsky A, Crook TH. Accelerated forgetting in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology* 1993;15:701–712. [PubMed: 8276930]
- Light, L.; Burke, D. Patterns of language and memory in old age. In: Light, LL.; Burke, DM., editors. *Language, memory, and aging*. Cambridge; New York: 1988. p. 244-271.
- Logan GD. Shapes of reaction time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1992;18:883–914.
- Luce, RD. *Reaction times: Their role in inferring elementary mental organization*. Oxford University Press; New York: 1986.
- Massaro, DW. *Perceiving talking faces: From speech perception to a general principle*. Bradford; Cambridge, MA: 1998.
- Maxwell, SE.; Delaney, HD. *Designing experiments and analyzing data: A model comparison perspective*. second edition. Erlbaum; Mahwah, NJ: 2004.
- McEvoy CL, Nelson DL. Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology* 1982;95(4):581–634.
- McGilchrist CA. REML estimation for survival models with frailty. *Biometrics* 1993;49:221–225. [PubMed: 8513103]
- Micco A, Masson MEJ. Age-related differences in the specificity of verbal encoding. *Memory & Cognition* 1992;20:244–253.
- Morris JC, Storandt M, Miller JP, McKeel DW, Price JL, Rubin EH, Berg L. Mild cognitive impairment represents early-stage Alzheimer disease. *Archives of Neurology* 2001;58(10):1705–1706.
- Neath, I.; Suprenant, AM. *Human memory: An introduction to research, data, and theory*. Second edition Ed.. Brooks/Cole; Pacific Grove, CA: 2003.
- O'Reilly RC, Rudy JW. Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review* 2001;108:311–345. [PubMed: 11381832]
- O'Toole, AJ.; Wenger, MJ.; Townsend, JT. Quantitative models of perceiving and remembering faces: Precedents and possibilities. In: Wenger, MJ.; Townsend, JT., editors. *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges*. Erlbaum; Mahwah NJ: 2001. p. 1-38.
- Petersen RC. Mild cognitive impairment: Current research and clinical implication. *Seminars in Neurology* 2007;27(1):22–31. [PubMed: 17226738]
- Petersen RC, Negash S. Mild cognitive impairment: An overview. *CNS Spectrums* 2008;13(1):45–53. [PubMed: 18204414]
- Petersen RC, Smith G, Kokmen E, Ivnik RJ, Tangalos EG. Memory function in normal aging. *Neurology* 1992;42:396–401. [PubMed: 1736173]
- Petersen RC, Smith GE, Ivnik RJ, Tangalos EG, Schaid DJ, Thibodeau SN, Kokmen E, Waring SC, Kurland LT. Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *Journal of the American Medical Association* 1995;273:1274–1278. [PubMed: 7646655]
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Komken E. Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology* 1999;56:303–308. [PubMed: 10190820]
- Rao, CR. *Linear Statistical Inference and Its Applications*. Second Edition. Wiley; New York: 1973.
- Roberts RO, Geda YE, Knopman DS, Cha RH, Boeve VSPBF, Ivnik RJ, Tangalos EG, Petersen RC, Roca WA. The Mayo Clinic Study of Aging: Design and sampling, participation, baseline measures and sample characteristics. *Neuroepidemiology* 2008;30(1):58–69. [PubMed: 18259084]

- Roediger HL. Memory metaphors in cognitive psychology. *Memory & Cognition* 1980;8:231–246.
- Roediger HL, Gallo DA, Geraci L. Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory* 2002;10:319–332. [PubMed: 12396644]
- Rosch E, Mervis CB, Gray WD, Johnson DM, Boyesbraem P. Basic objects in natural categories. *Cognitive Psychology* 1976;8(3):382–439.
- Salthouse, TA. A theory of cognitive aging. North-Holland; Amsterdam: 1985.
- Salthouse TA. Aging and measures of processing speed. *Biological Psychology* 2000;54:35–54. [PubMed: 11035219]
- Salthouse TA. A research strategy for investigating group differences in a cognitive construct: Application to aging and executive processes. *European Journal of Cognitive Psychology* 2001;13:39–46.
- Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. *Developmental Psychology* 1991;27:763–776.
- Schweickert R, Boggs GJ. Models of central capacity and concurrency. *Journal of Mathematical Psychology* 1984;28:223–281.
- Shepherd, GM., editor. *The synaptic organization of the brain*. fifth Ed.. Oxford University Press; Oxford: 2004.
- Smith AD. Adult age differences in cued recall. *Developmental Psychology* 1977;13:326–331.
- Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. *Trends in Neuroscience* 2004;27:161–168.
- Snodgrass JG, Vanderwart MA. A standardized set of 260 pictures: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* 1980;6:174–215. [PubMed: 7373248]
- Spear, NE. *The processing of memories: Forgetting and retention*. Erlbaum; Hillsdale, NJ: 1978.
- Spieler DH, Balota DA. Bringing computational models of word naming down to the item level. *Psychological Science* 1997;8:411–416.
- Spieler DH, Balota DA, Faust ME. Levels of selective attention revealed through analyses of reaction time distributions. *Journal of Experimental Psychology: Human Perception and Performance* 2000;26:506–526. [PubMed: 10811160]
- Therneau, TM.; Grambsch, PM. *Modeling survival data: Extending the Cox model*. Springer-Verlag; New York: 2000.
- Thomas, RD. Characterizing perceptual interactions in face identification using multidimensional signal detection theory. In: Wenger, MJ.; Townsend, JT., editors. *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges*. Erlbaum; Mahwah, NJ: 2001. p. 193–228.
- Townsend JT. Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin* 1990;108:551–567. [PubMed: 2270240]
- Townsend, JT.; Ashby, FG. Methods of modeling capacity in simple processing systems. In: Castellan, J.; Restle, F., editors. *Cognitive theory*. Vol. 3. Erlbaum; Hillsdale, NJ: 1978. p. 200–239.
- Townsend, JT.; Ashby, FG. *Stochastic modeling of elementary psychological processes*. Cambridge University press; Cambridge: 1983.
- Townsend JT, Nozawa G. On the spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology* 1995;39:321–359.
- Townsend JT, Wenger MJ. The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review* 2004a;11:391–418. [PubMed: 15376788]
- Townsend JT, Wenger MJ. A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review* 2004b;111:1003–1035. [PubMed: 15482071]
- Tsodyks M, Sejnowski T. Rapid state switching in balanced cortical network models. *Network: Computation in Neural Systems* 1995;6:111–124.
- Tuokko H, Crockett D. Cued recall and memory disorders in dementia. *Journal of Clinical and Experimental Neuropsychology* 1989;11:278–294. [PubMed: 2925836]
- Tuokko H, Vernon-Wilkinson R, Weir J, Beattie BL. Cued recall and early identification of dementia. *Journal of Clinical and Experimental Neuropsychology* 1991;13:871–879. [PubMed: 1779027]

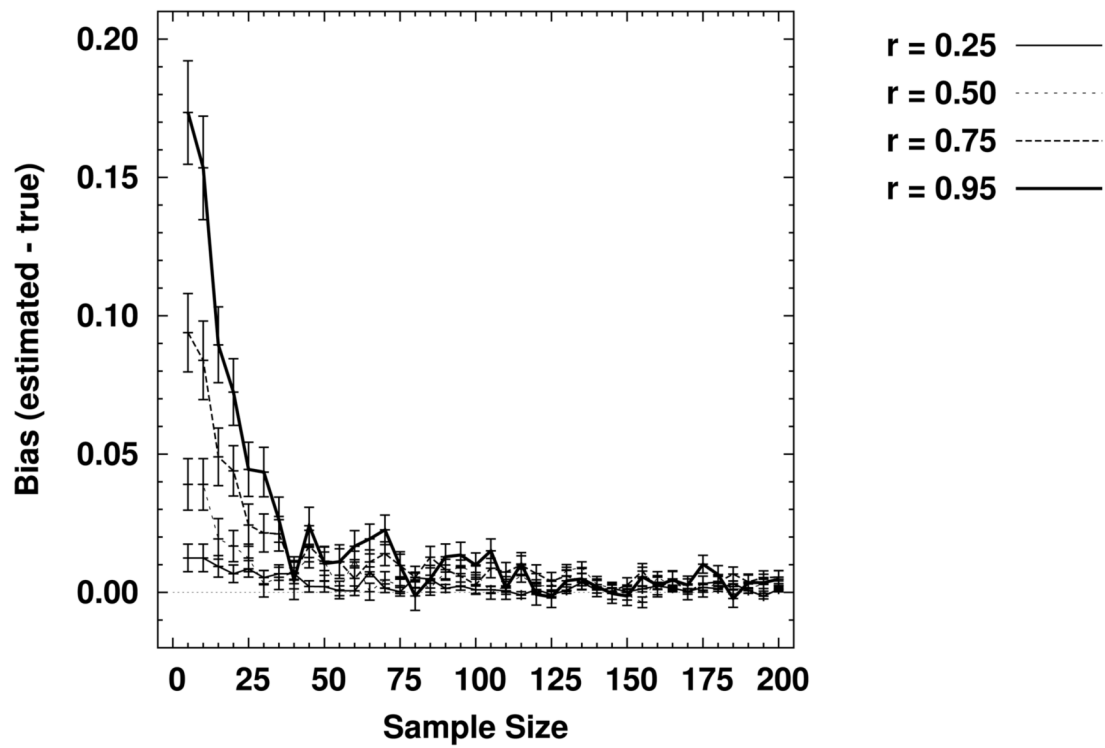
- Vemuri P, Gunter JL, Senjem ML, Whitwell JL, Kantarci K, Knopman DS, Boeve BF, Petersen RC, Jack CR. Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies. *Neuroimage* 2008;39(3):1186–1197. [PubMed: 18054253]
- Waugh, NC.; Barr, RA. Encoding deficits in aging. In: Craik, FIM.; Trehub, S., editors. *Aging and Cognitive Processes, Volume 8, Advances in the study of communication and affect*. Plenum; New York: 1982. p. 183-190.
- Weibull W. A statistical distribution function of wide applicability. *Journal of Applied Mechanics* 1951;18:293–297.
- Welford, AT. *Ageing and human skill*. Oxford University Press; London: 1958.
- Wenger MJ. On the whats and hows of retrieval in the acquisition of a simple skill. *Journal of Experimental Psychology: Learning, Memory and Cognition* 1999;25:1137–1160.
- Wenger MJ, Gibson BS. Using hazard functions to assess changes in processing capacity in an attentional cuing paradigm. *Journal of Experimental Psychology: Human Perception and Performance* 2004;30:708–719. [PubMed: 15301619]
- Wenger, MJ.; Schuster, C.; Petersen, LM.; Petersen, RC. Application of proportional hazards models to reaction time data. In: Bergeman, C.; Boker, S., editors. *Notre Dame Series on Quantitative Methodologies*. Erlbaum; Mahwah, NJ: 2006. p. 231-268.
- Wenger MJ, Townsend JT. Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology* 2000;127:67–99. [PubMed: 10695952]
- Wenger, MJ.; Townsend, JT. Faces as gestalt stimuli: Process characteristics. In: Wenger, MJ.; Townsend, JT., editors. *Computational, geometric, and process perspectives on facial cognition*. Erlbaum; Mahwah, NJ: 2001. p. 229-284.
- Wenger MJ, Townsend JT. On the costs and benefits of faces and words. *Journal of Experimental Psychology: Human Perception and Performance* 2006;32:755–779. [PubMed: 16822136]
- Whitwell JL, Jack CR, Pankratz VS, Parisi JE, Knopman DS, Boeve BF, Petersen RC, Dickson DW, Josephs KA. Rates of brain atrophy over time in autopsy-proven frontotemporal dementia and Alzheimer disease. *Neuroimage* 2008a;39(3):1034–1040. [PubMed: 17988893]
- Whitwell JL, Petersen RC, Negash S, Weigand SD, Kantarci K, Ivnik RJ, Knopman DS, Boeve BF, Smith GE, Jack CR. Patterns of atrophy differ among specific subtypes of mild cognitive impairment. *Archives of Neurology* 2007;64(8):1130–1138. [PubMed: 17698703]
- Whitwell JL, Shiung MM, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, Petersen RC, Jack CR. MRI patterns of atrophy associated with progression to AD in amnesic mild cognitive impairment. *Neurology* 2008b;70(7):512–520. [PubMed: 17898323]
- Van Zandt T. How to fit a response time distribution. *Psychonomic Bulletin & Review* 2000;7:424–465. [PubMed: 11082851]
- Van Zandt, T. Analysis of response time distributions. In: Wixted, JT., editor. *Stevens' Handbook of Experimental Psychology, Volume 4: Methodology in Experimental Psychology*. 3rd Ed.. Academic; San Diego: 2002. p. 461-516.



**Figure 1.** Schematic representation of the architecture of the computational model of the hippocampal circuit.

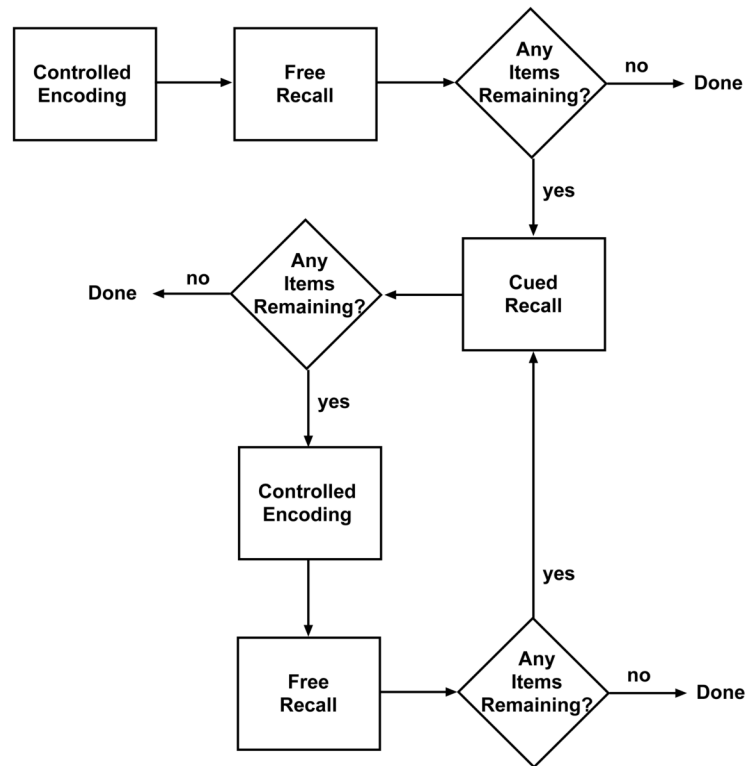


**Figure 2.** Effects of increasing atrophy and load (at each level of atrophy) on proportional change in the mean and the hazard function, for the computational model of the hippocampal circuit.

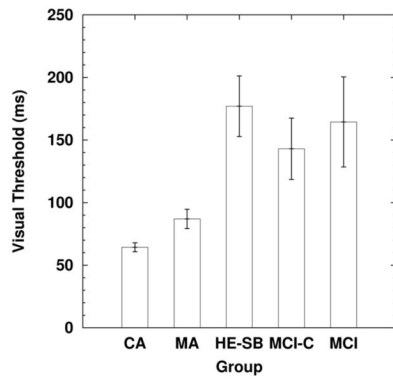


**Figure 3.** Bias in estimates of proportionality at the level of the hazard function, as a function of sample size, for four levels of known proportionality ( $r$ ).

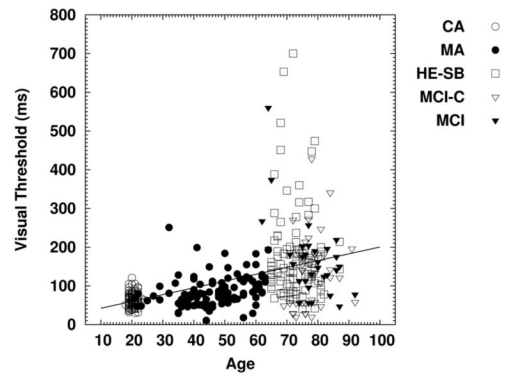




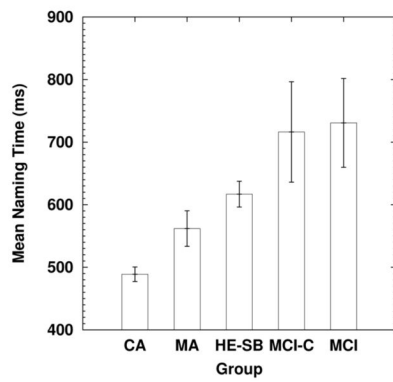
**Figure 4.**  
Flow-chart representation of the free and cued selective reminding task (FCSRT).



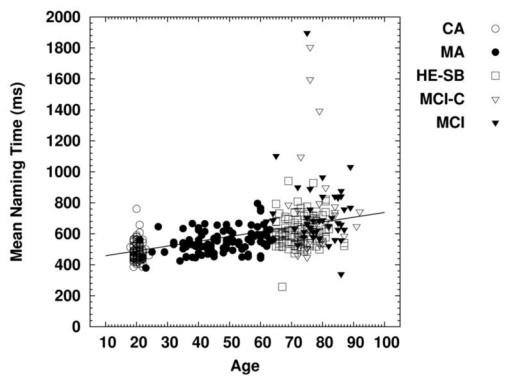
(a)



(b)

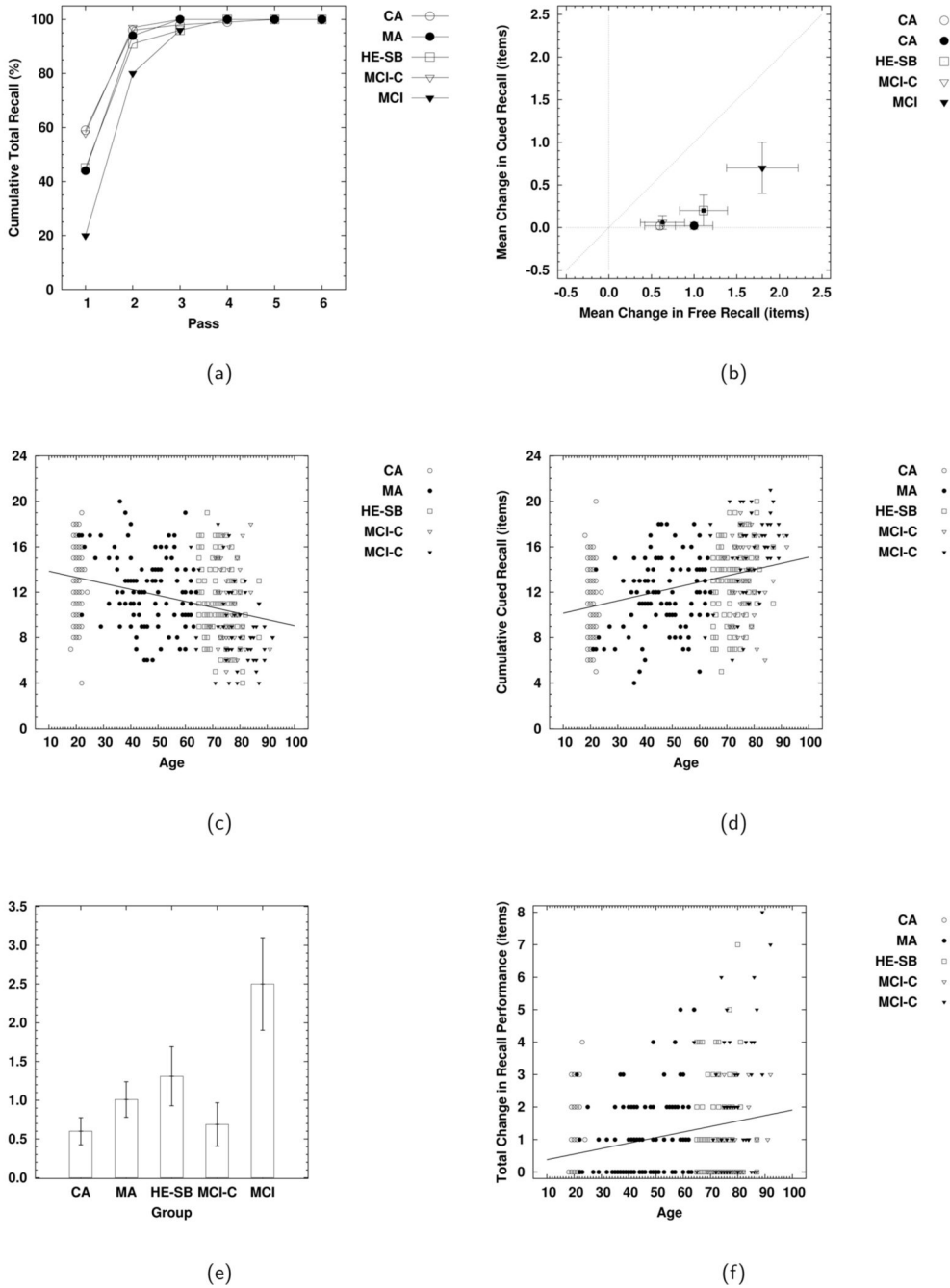


(c)

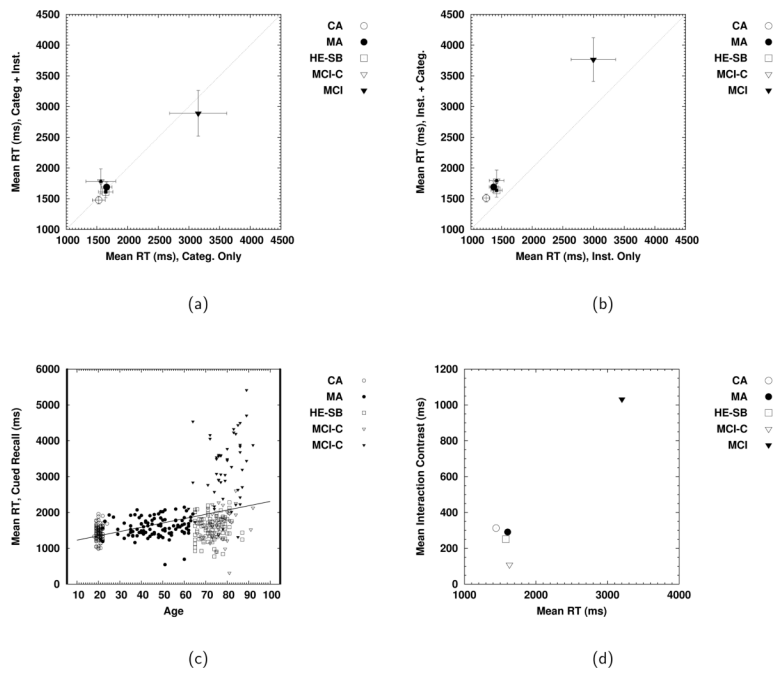


(d)

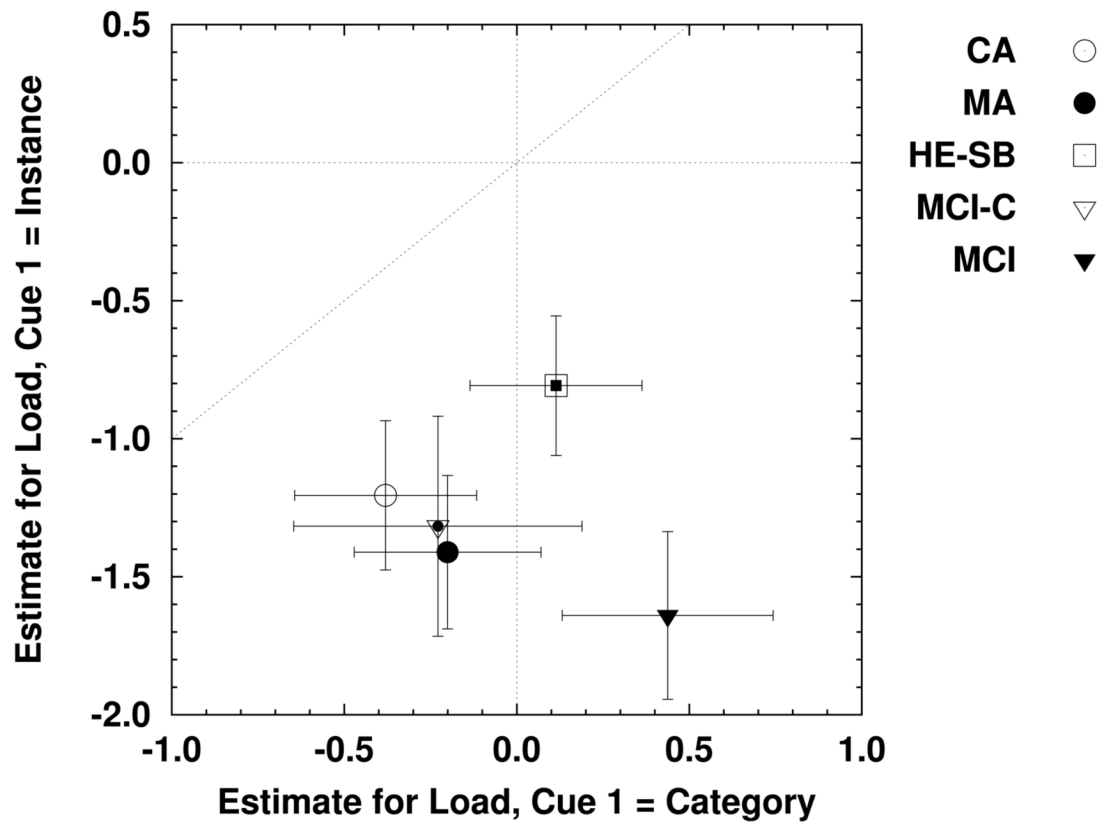
**Figure 5.** Group means (a) and individual estimates (b) of visual identification threshold. Group means (c) and individual means (d) for naming times. Error bars indicate the width of the 95% confidence intervals.



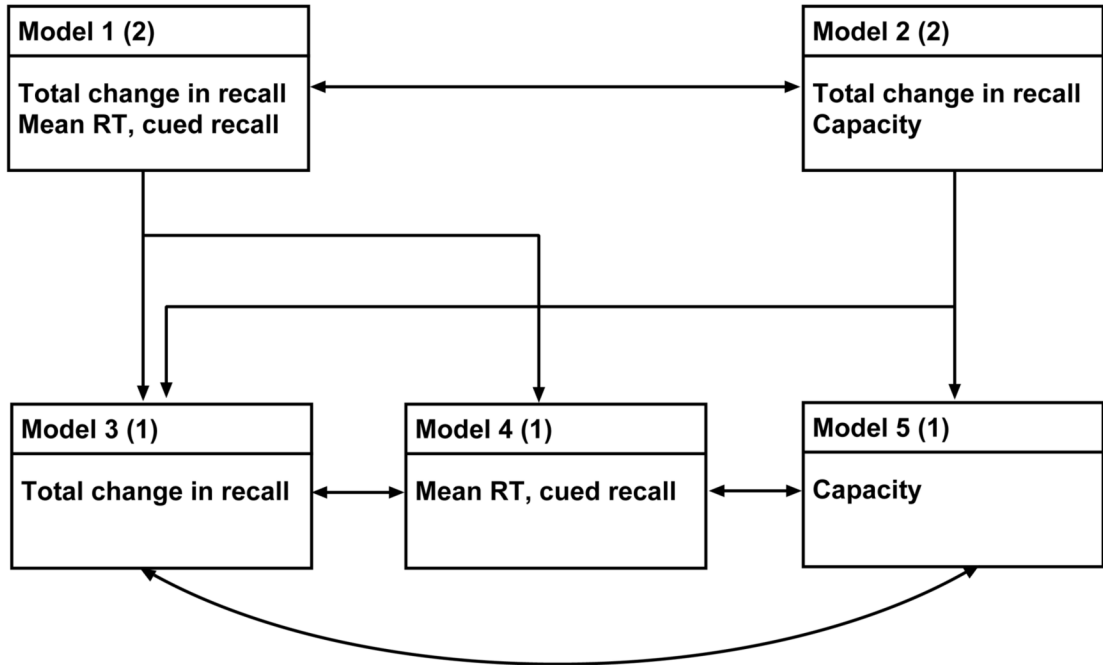
**Figure 6.** (a) Cumulative total recall by pass for each of the five groups; (b) mean change (final - initial) in free and cued recall for each of the five groups; (c) final free recall for all participants as a function of age; (d) final cued recall for all participants as a function of age; (e) mean total change in recall for each of the five groups; (f) total change in recall for all participants as a function of age. Error bars indicate 95% confidence intervals.



**Figure 7.** (a) Mean RTs for single- and double-cue cued recall trials, given that the single cue was a category cue; (b) mean RTs for single- and double-cue cued recall trials, given that the single cue was an instance cue; (c) mean RTs overall per participant as a function of age; (d) interaction contrast (see Equation 9) for the interaction of single-cue condition and number of cues for each group, relative to the overall mean RT for that group.

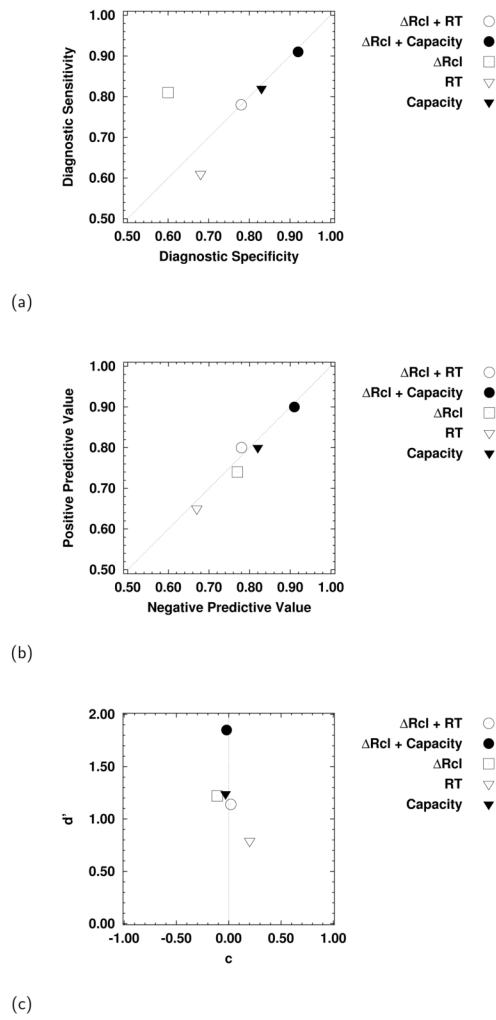


**Figure 8.** Effects of increasing cues on capacity, as estimated using the proportional hazards model, for each combination of group and single-cue condition. Error bars are 95% confidence intervals on the parameter estimates.



**Figure 9.** Relations among five models for predicting MCI status on the basis of three predictor variables. Numbers in parentheses indicate the number of predictor variables in each model, with those variables listed in the lower portion of each box. Single-headed arrows indicate relationships among models differing in the number of predictor variables, double-headed arrows indicate relationships involving the same number of predictor variables





**Figure 10.** Measures of diagnostic sensitivity, accuracy, and bias, based on the outcomes of the discriminant function analyses. Note:  $\Delta Rcl$  = total change in recall, RT = mean cued recall RT.

Cues		Cues	
Item	Category	Item	Instance



mountain earth formation climb



dog domestic animal pet

Cues		Cues	
Item	Category	Item	Instance

reading material chapter

vehicle dashboard

*J Math Psychol.* Author manuscript; available in PMC 2011 February 1.

Item	Cues	
	Category	Instance

*J Math Psychol.* Author manuscript; available in PMC 2011 February 1.



sheets

furniture

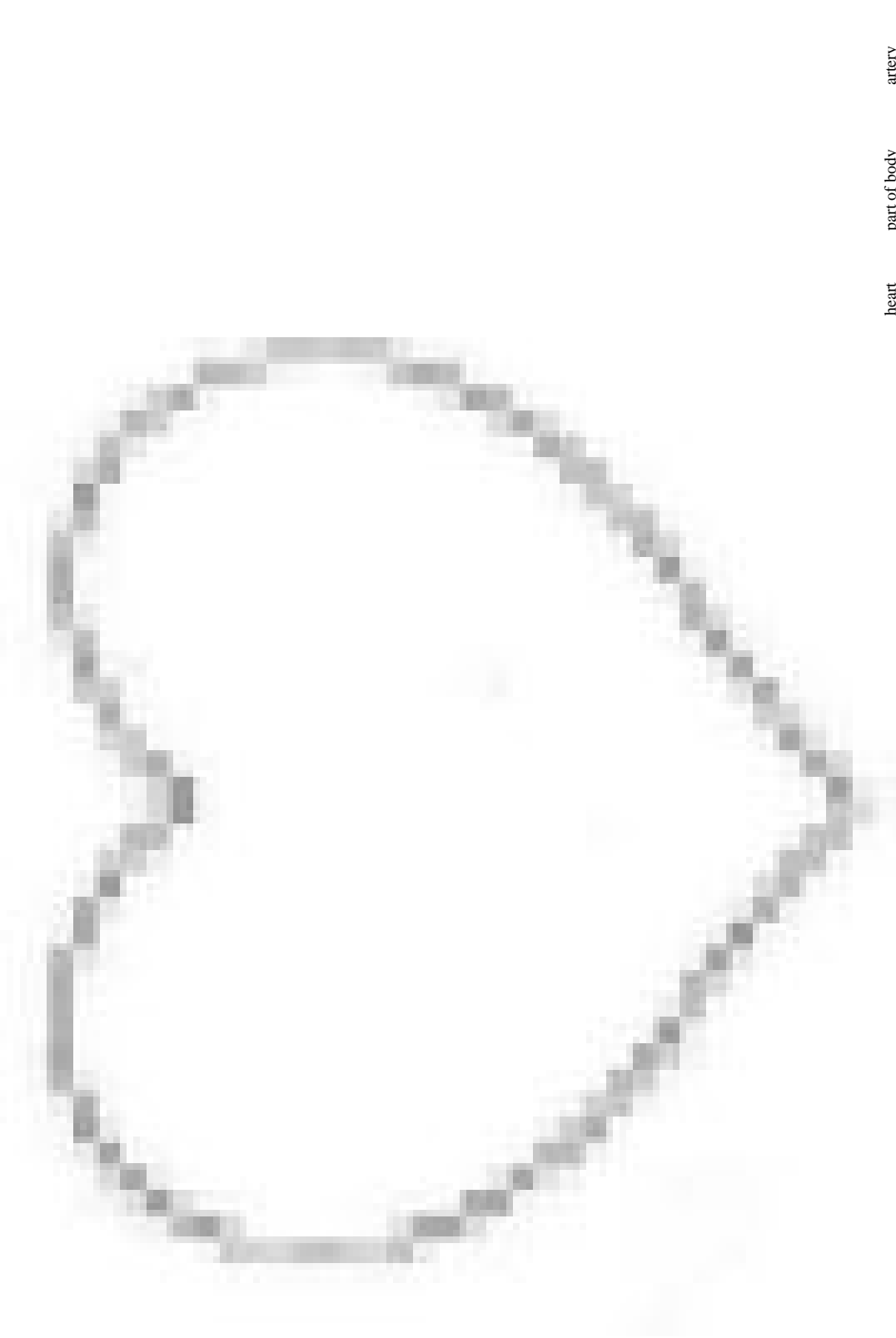
bed

collar

clothing

Cues	
Item	Category
	Instance

Cues	
Item	Category
	Instance



heart

part of body

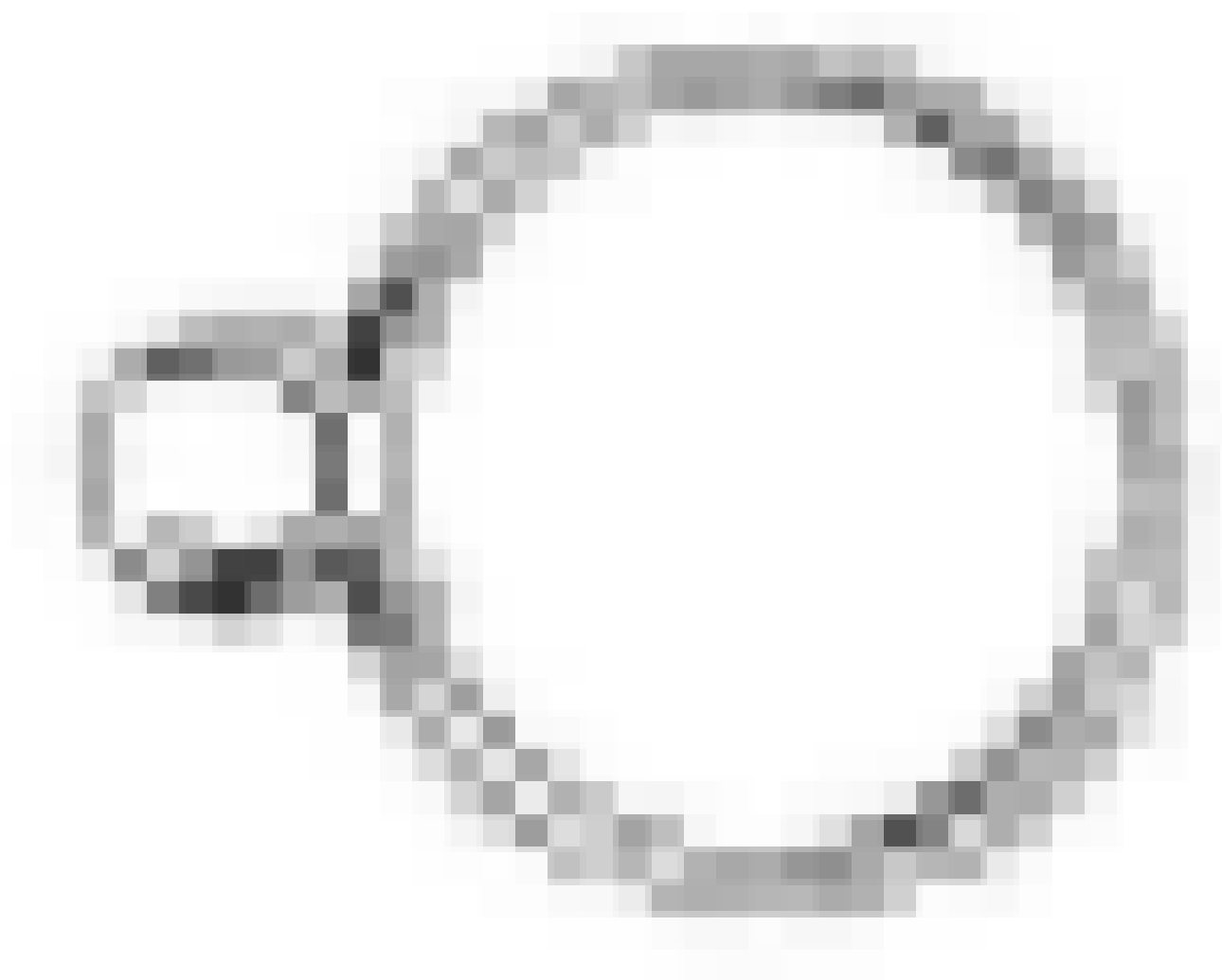
artery

copper

kitchen tool

Item Category Instance

Item Category Instance



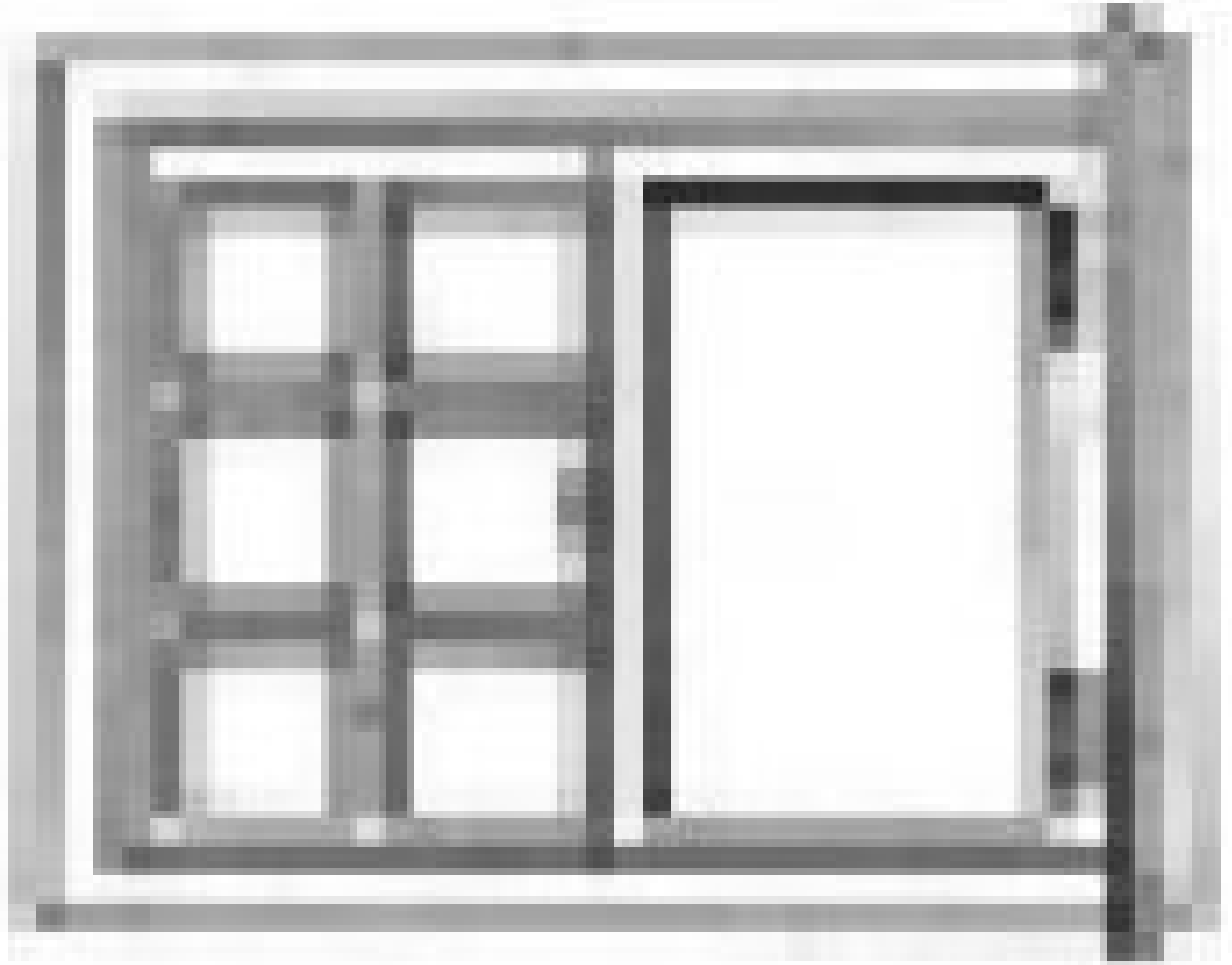
*J Math Psychol.* Author manuscript; available in PMC 2011 February 1.

ring type of jewelry pearl

chapel type of building

Item Category Instance

Item Category Instance



*J Math Psychol.* Author manuscript; available in PMC 2011 February 1.

window part of building shutter

type of toy soccer



Cues		Cues	
Item	Category	Item	Instance



strawberry

fruit

cherry

trigger

weapon

Item Category Instance

Item Category Instance



*J Math Psychol.* Author manuscript; available in PMC 2011 February 1.

stove major appliance range

keyboard instrument organ

Item	Cues	
	Category	Instance



bread

type of food

butter

lace

footwear

Cues	
Item	Category
	Instance

Cues	
Item	Category
	Instance



cow

farm animal

herd

sight

part of the face

Cues	
Item	Category
	Instance

Cues	
Item	Category
	Instance



knife  
eating utensil  
fork

brass instrument  
trombone

Cues		Cues	
Item	Category	Item	Instance
	pencil	writing implement	lead
		insect	spider



Participant characteristics, as reported on the background questionnaire. Note: for the education variable, it was assumed that all participants in the CA group would graduate.

**Table 2**

	CA	MA	HE-SB	MCI-C	MCI
<i>N</i>	100	104	97	48	50
Age: mean	20.3	46.5	72.7	77.3	79.3
Sex: % female	56	63	64	50	37
Race: % reporting white	90	82	100	100	98
Education: mean years	14.0	16.2	14.3	13.4	13.6
Education: % with bachelors or advanced degree	100	62	32	36	29
English: % native speakers	97	97	97	93	98
English: (mean yrs spoken / mean age) × 100	99	97	99	97	98
Employment: % reporting retired	0	5	66	86	87
Head injury: % reporting ever experiencing	13	9	0	4	0
Depression: % reporting ever experiencing	6	16	14	0	0

**Table 3**

Mean RT for cued recall trials, summary of the ANOVA.

<b>Factor</b>	<b><i>df</i></b>	<b><i>F</i></b>	<b><i>MSE</i></b>
Group (G)	4	185.47 ‡	324786
Single-cue condition (SC)	1	0.05	324786
Number of cues (NC)	1	60.00 ‡	104847
G × SC	4	3.90 †	324786
G × NC	4	2.91 *	104847
SC × NC	1	66.91 ‡	104847
G × SC × NC	4	8.47 ‡	104847

Note:

\*  $p < .05$ †  $p < .01$ ‡  $p < .001$ .

**Table 4**

Estimated parameters for the effect of varying the number of retrieval cues on capacity (proportional hazards), for each combination of group type and single-cue condition.

Group	Single-cue Condition	$\beta$	SE	$\chi^2$	% Change
CA	Category	-0.380	0.264	2.08	-32
	Instance	-1.205	0.270	19.90 <sup>‡</sup>	-70
MA	Category	-0.201	0.271	0.55	-18
	Instance	-1.411	0.278	25.78 <sup>‡</sup>	-76
HE-SB	Category	0.114	0.249	0.21	12
	Instance	-0.807	0.253	10.20 <sup>†</sup>	-55
MCI-C	Category	-0.229	0.418	0.30	-20
	Instance	-1.317	0.398	10.93 <sup>‡</sup>	-73
MCI	Category	0.437	0.206	4.04 <sup>*</sup>	55
	Instance	-1.681	0.305	14.96 <sup>‡</sup>	-81

Note:

\*  $p < .05$

<sup>†</sup>  $p < .01$

<sup>‡</sup>  $p < .001$ .

**Table 5**

Evaluation (top section) and comparison (bottom section) of the five models for MCI status, based on the discriminant function analyses.

Model	$\ln(L)$	$\nu$	AIC
1	-37.864	2	79.728
2	-34.587	2	53.174
3	-54.775	1	111.550
4	-57.964	1	117.928
5	-44.499	1	90.998

Restricted Model	General Model	$df$	$\chi^2$
3	1	1	33.822‡
3	2	1	60.376‡
4	1	1	40.200‡
5	2	1	39.824‡

**Table 6**

Measures of diagnostic accuracy, sensitivity, and bias based on the frequencies of hits (H), misses (M), correct rejections (CR), and false alarms (FA), obtained from the discriminant function analyses. Note:  $z(HR)$ ,  $z(FR)$  = inverse normal transformation of the hit and false alarm rates.

Measure	Form
Diagnostic sensitivity (hit rate, HR)	$\frac{H}{H + M}$
Diagnostic specificity	$\frac{CR}{FA + CR} = 1 - FR$
Positive predictive value	$\frac{H}{H + FA}$
Negative predictive value	$\frac{CR}{CR + M}$
2AFC $d'$ (sensitivity)	$\frac{1}{\sqrt{2}}[z(HR) - z(FR)]$
2AFC $c$ (bias)	$\frac{1}{2\sqrt{2}}[z(HR) + z(FR)]$

**Table 7**

Correlations between individual predictors and total hippocampal volume, for MCI-C and MCI participants in each of the two single-cue conditions.

Measure	MCI-C		MCI	
	Category	Instance	Category	Instance
Age	.052	−0.427	−0.008	−0.050
Total change in recall	−.411	−0.020	−0.759*	0.636*
Mean RT, cued recall	.034	−0.042	−0.267	−0.181
Capacity	.194	0.062	0.845†	0.775*

**Table 8**

Parameters, their roles, and values in the numeric simulations of the hippocampal circuit.

Parameter	Role	Values
$dt$	simulation time step	1 ms
$t_+$	time, input on	200 ms
$t_-$	time, input off	100 ms
$N_E$	total excitatory cells, each region	3,000
$N_I$	total inhibitory cells, each region	$\frac{1}{5} \times N_E$
$N_\emptyset$	total cells receiving no input	$\frac{1}{6} \times (N_E + N_I)$
$a$	proportion of cells in each region deleted, atrophy	0.10 to 0.50 by 0.10
$N_S$	max synapses from any cell $i$ to any other cell $j$	35
$\tau_\alpha$	membrane time constant	5 ms
$V_0$	resting value, membrane potential	5 mV
$V_{max}$	maximum value, membrane potential (spike)	25 mV
$V_{min}$	minimum value, membrane potential	1 mV
$\gamma$	spiking threshold	20 mV
$\lambda$	asymptotic inter-spike (waiting) time	100 ms
$\beta$	slope, spike-rate adaptation function	200
$J_{ij}, i, j \in E$	mean synaptic efficiency, excit./excit. connections	0.20
$J_{ij}, i \in E, j \in I$	mean synaptic efficiency, inhib./excit. connections	0.30
$J_{ij}, i \in I, j \in E$	mean synaptic efficiency, excit./inhib. connections	0.60
$J_{ij}, i, j \in I$	mean synaptic efficiency, inhib./inhib. connections	1.00
$\sigma(J)$	standard deviation, synaptic efficiency (all cases)	0.10
$s$	sparseness parameter, synaptic efficiencies	0.25
$\delta_{ij}$	transmission delay between cells	1 ms within a given region 1 ms between EC and DG 2 ms between EC and CA3, CA1
$I_i^{\text{ext}}$	mean external input	20
$\sigma(I_i^{\text{ext}})$	standard deviation, external input	$\frac{1}{4} \times I_i^{\text{ext}}$
$\gamma_L, \gamma_H$	criterion number of output spikes, low and high load	1000, 2000
$N_R$	number of replications at each level	5,000

**Table 9**

Parameters, their roles, and values in the numeric simulations used to assess the bias in estimates of proportionality, as a function of sample size and effect (true level of proportionality).

Parameter	Role	Values
$\alpha$	Shape parameter, Weibull distribution	1.90
$\beta$	Location parameter, Weibull distribution	600
$r$	Proportionality, two samples	0.25 to 1.00 by 0.05
$n$	Sample size	5 to 200 by 5
$N$	Replications at each level	10,000