BRIEF REPORT

Mechanisms of Age-related Decline in Memory Search Across the Adult Life Span

Thomas T. Hills (t.t.hills@warwick.ac.uk)

Dept. of Psychology, University of Warwick, Coventry, CV4 7AL, UK

Rui Mata (rui.mata@unibas.ch)

Dept. of Psychology, University of Basel, Missionsstrasse 60/62, 4055 Basel, CH

Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195

Berlin, Germany

Andreas Wilke (awilke@clarkson.edu)

Dept. of Psychology, Clarkson University, 8 Clarkson Av, Box 5805, Potsdam, NY 13699, USA

Gregory R. Samanez-Larkin (g.samanezlarkin@vanderbilt.edu)

Psychological Sciences, Vanderbilt University, 111 21st Avenue South, Nashville, TN 37240, USA

Abstract

Three alternative mechanisms for age-related decline in memory search have been proposed as resulting from either reduced processing speed (global slowing hypothesis), over-persistence on categories (cluster-switching hypothesis), or the inability to maintain focus on local cues related to a decline in working memory (cue-maintenance hypothesis). We investigated these three hypotheses by formally modeling the semantic recall patterns of 185 adults aged between 27 to 99 years of age in the animal fluency task (Thurstone, 1938). The results indicate that people switch between global frequency-based retrieval cues and local item-based retrieval cues to navigate their semantic memory. Contrary to the global slowing hypothesis that predicts no qualitative differences in dynamic search processes and the cluster-switching hypothesis that predicts reduced switching between retrieval cues, the results indicate that as people age they tend to switch more often between local and global cues, per item recalled, supporting the cue-maintenance hypothesis. Additional support for the cue-maintenance hypothesis is provided by a negative correlation between switching and digit span, and switching and total items recalled, which suggest that cognitive control may be involved in cue maintenance and the effective search of memory. Overall, the results are consistent with age-related decline in memory search being a consequence of reduced cognitive control, consistent with models suggesting working memory is related to goal perseveration and the ability to inhibit distracting information.

Keywords: search; semantic memory; executive control; animal fluency; age-related cognitive decline.

What are the mechanisms underlying age-related declines in the rate at which information is retrieved during memory search? We tested three alternative hypotheses by investigating age differences in a widely used fluency task ("name all the animals you can think of"; Thurstone, 1938). In this task, numerous studies have demonstrated that older adults consistently recall fewer items relative to younger adults within a fixed time interval (e.g., Kozora & Cullum, 1995; Lanting, Haugrud, & Crossley, 2009; Tomer & Levin, 1993). The three hypotheses proposed to explain this difference each suggest a different underlying cognitive mechanism. The *global slowing hypothesis* holds that retrieval deficits associated with age are the result of age-related cognitive slowing (Mayr, 2002; Mayr & Kliegl, 2000). According to this view, aging is associated with slower (cf. Salthouse, 1996) but otherwise unaffected cognitive processing, and therefore predicts no age differences in the utilization of memory retrieval cues. Two alternative proposals, however, suggest that the impact of aging goes beyond global slowing. In particular, these proposals suggest that aging may impact how memory is queried through the ability to appropriately handle perseveration or abandonment of memory retrieval cues.

One of the proposals, the *cluster-switching hypothesis* (Troyer, Moscovitch, & Winocur, 1997), sees memory retrieval as a dynamic process involving, first, a search for semantic categories (e.g., PETS), and second, a search for and recall of words within a category (e.g., DOG). Troyer et al. (1997) developed a hand-coded categorization of animals based on typical recall patterns, and developed measures of switching (i.e., the number of switches between categories) and clustering (i.e., the mean number of items in a cluster). Several studies have now used these switching and clustering measures to investigate age-related differences in semantic fluency. A common finding is that, along with a recall of fewer items overall, aging is associated with fewer total switches between categories (Lanting et al., 2009; Troyer et al., 1997; Troyer, 2000; but see Rosselli, Tappen, Williams, Salvatierra, & Zoller, 2009). Results regarding the size of each cluster are less clear with two studies finding no differences (Troyer et al., 1997; Troyer, 2000), and others finding increased numbers of items per cluster with

increased age (Lanting et al., 2009), at least for some semantic categories (Rosselli et al., 2009). Based on this empirical observation, Troyer and colleagues (Troyer et al., 2007; Troyer, 2000) proposed that aging is associated with *reduced* switching between retrieval cues (categories), an outcome they proposed must be linked to cognitive control brain networks affected by aging (Bäckman, Nyberg, Lindenberger, Li, & Farde 2006; Braver & Barch, 2002; Buckner, 2004; Hedden & Gabrieli, 2004; Li, Lindenberger, & Bäckman, 2010). However, this prediction lacks a clear development of the cognitive mechanisms that would reduce switching and has come under fire because it uses the absolute number of switches as a dependent measure, which has been shown to be problematic for several reasons (Mayr, 2002; and see below).

The third view of age differences in memory search, the *cue-maintenance hypothesis*, arises out of the literature on working memory. According to this view, memory retrieval can also be seen as a dynamic process by which individuals utilize specific retrieval cues to access memory. Theories of cognitive control see working memory as the ability to maintain focus on one cue while ignoring other potentially distracting ones (e.g., Kane & Engle, 2000; Hills, Todd, & Goldstone, 2010). Consequently, loss of executive capacities leads to a loss of cue focus, and thus more frequent switching between retrieval cues (e.g., Unsworth & Engle, 2007). Aging is associated with lower working memory capacity, as measured by either simple or complex span tasks (Bopp & Verhaghen, 2007), likely due to age differences in the structure and function of prefrontal brain networks (Buckner, 2004; Hedden & Gabrieli, 2004), specifically the frontostriatal dopamine system (Bäckman et al. 2006; Braver & Barch, 2002; Li et al., 2010; Paxton, Barch, Racine, & Braver, 2008). Notably, studies have found that lower working-memory capacity is associated with increased switching between retrieval cues during recall in fluency tasks (Hills & Pachur, 2012; Rosen & Engle, 1997; Unsworth, Spillers, Brewer, 2011). Consequently, the cue-maintenance hypothesis predicts that age-related decline in working memory capacity should lead to *increased* switching between retrieval clues.

Though declines in the rate of retrieval follow directly for global slowing, the other two hypotheses are based on the implicit assumption that optimal retrieval from memory requires switching between categories that is neither too frequent nor too infrequent—in much the same way a bird may forage among patches of berries. Switching too frequently leads to inadequate retrieval of items that are related to one another, and thus leaves clusters of related items before many of these items are recalled. In this case, most of the time is spent moving between clusters. Switching too infrequently leads to perseveration in regions of memory where most of the items have already been recalled, again leading to lower recall rates because most everything in the local region of the memory space has already been retrieved. Indeed, recent research has shown that individuals who do not leave categories either too soon or too late retrieve items faster from memory in a fluency task, whereas those who switch between categories either too frequently or too infrequently retrieve items more slowly (see Hills, Jones, & Todd, 2012).

To summarize, three hypotheses—global slowing, cluster-switching, and cue-maintenance—make different predictions about age-related differences in dynamic memory search.[i] While the *global slowing* hypothesis predicts slowing but otherwise no differences in the nature of cue utilization with increased age, the other two hypotheses posit specific changes in cue utilization with increased age, with *cluster-switching* predicting *reduced* switching (per item recalled) between retrieval cues, and *cue-maintenance* predicting *increased* switching (per item recalled) between retrieval cues. In what follows, we investigate this issue by providing the first test of these competing hypotheses through computational modeling of semantic retrieval patterns in a fluency task.

To do this, we utilized the computational search framework provided by the Search of Associative Memory model (SAM; Raaijmakers & Shiffrin, 1981; see also Hills, Jones, & Todd, 2012) to describe memory retrieval given a well-defined memory representation (Jones & Mewhort, 2007). Crucially, this framework allowed us to build different models that varied concerning the cues used to

sample memory. We were thus able to assess how different assumptions concerning cue utilization could account for age differences in memory search. Specifically, this approach made it possible to assess the link between age and individual differences in cognitive control and specific model parameters (i.e., the amount of switching between retrieval cues). Our goal was to infer whether, contrary to the global slowing hypothesis, we would find age differences in cue utilization, and if so, whether age differences in cue utilization were consistent with prior characterizations of the cluster-switching or cue-maintenance hypotheses. Finally, the participant data we use to address these questions also collected digit-span measures as part of an additional study, which allowed us to further investigate more general predictions of declines in cognitive control.

Methods

Participants and Procedure

Two-hundred and one participants (30% males, 70% females), with ages ranging from 27 to 99 (Median = 68, M = 65.45, SD = 13.53), participated in the study. Sixteen participants were excluded due to low scores on a dementia screener (Mini-Mental State Examination, Folstein, Folstein, & McHugh, 1975). We also excluded individuals with education-adjusted scores below 27 (Crum, Anthony, Bassett, & Folstein, 1993) leaving 185 participants for analysis with ages ranging from 29 to 99 (Median = 69, M = 66.66, SD = 12.74).[ii] Participants had different levels of education ranging from 7 to 20 years of schooling (Median = 16; M = 15.26, SD = 2.54). Participants were recruited through postings at a health center in the San Francisco Bay Area as part of a study on health-related decision making. Participants were tested individually at the health center and first completed questionnaires regarding preferences for treatments within hypothetical health states (e.g., cancer, dementia), followed by a small battery of cognitive tests, which included the animal naming (Thurstone, 1938; Lindenberger, Mayr, & Kliegel, 1993) and digit span (Wechsler, 1997) tasks—used as measures of fluid intellectual ability in the health preferences study—and a vocabulary task (Wechsler, 1997). The animal naming task consisted of asking

participants to name all the animals they could in 1 minute, using the following instructions "When I say start please tell me as many words that you can think of that fit into the category of animal. Start!" The experimenter wrote down all words in the order that participants recalled them and stopped the participant after 1 minute. Both forward and backward span were included in the digit span task[iii]. The scores reported are the sum of trials completed in the forward span (which assesses short-term maintenance) and backward span (which assesses both maintenance and item manipulation) and we use it as a measure of working memory. The vocabulary task requires participants to provide open-ended definitions to each word in a list. Summary statistics of individual difference measures by age decade are provided in Appendix A. Participants were compensated for their participation ($50).

The Representation of Semantic Memory

The first step towards formalizing search in semantic space is to provide an explicit representation of the space being searched. In our modeling, we used the lexical semantic representations of animals computed in prior work (Hills et al., 2012) using the BEAGLE semantic space model (Jones & Mewhort, 2007). BEAGLE learns semantic relations by assigning each word an *initial vector* with vector elements sampled randomly from a Gaussian distribution with $\mu = 0$ and $\sigma = 1$, where D is the vector dimensionality (set to 1000 in these simulations). As the text corpus is processed, a second vector, the word's *memory vector*, is updated each time the word is encountered as the sum of the initial vectors for the other words appearing in context with it. Once the entire corpus has been learned, a word's memory representation is a vector pattern reflecting the word's history of co-occurrence with other words. By this method, words that frequently co-occur will develop similar vector patterns (e.g., bee-honey), as will words that commonly occur in similar contexts, even if they never directly co-occur (e.g., bee-wasp). We used the pairwise similarity metric vector cosine (a normalized dot-product between two word vectors) for our comparisons. BEAGLE was trained on a subset of Wikipedia, composed of approximately 400 million word tokens and 3 million word types. Support for

BEAGLE comes from its success at accounting for a variety of effects, including semantic typicality, categorization, sentence completion (Jones & Mewhort, 2007), priming (Jones, Kintsch, & Mewhort, 2006), and retrieval from long-term memory (Hills et al., 2012).

Search in Semantic Memory

Memory retrieval can be viewed as the result of probing a memory representation with one or more cues to activate a response (e.g., Gronlund & Shiffrin, 1986; Walker & Kintsch, 1985). Under this assumption, the simplest possible model of memory search would assume that search is guided by a single cue. More complex models could assume a combination of cues, or include dynamically switching between cues as a function of retrieval success. To model these various approaches we used a model framework similar to the item-level recall probability equation from SAM (see Raaijmakers & Shiffrin, 1981):

$$P(I_i \mid Q_1, Q_2, ..., Q_M) = \frac{\prod_{k=1}^{M} S(Q_k, I_i)^{w_k}}{\sum_{j=1}^{N} \prod_{k=1}^{M} S(Q_k, I_j)^{w_k}} \tag{1}$$

where $S(Q_k, I_i)$ represents the retrieval strength from cue Q_k to item I_i in memory, and w_k represents the saliency or attention directed at the kth cue. The probability of retrieving a given item, I_i, is given by the ratio of the activation strength of that item and the product of the activation of all other items in memory given those same cues. The saliency parameter, w, provides a measure of the deterministic nature of the activation; higher values of w lead items with higher retrieval strengths for a given cue, Q_k, to gain a larger share of the recall probability, while lower values of w more evenly distribute the probability of recall over all items.

We considered two cues: a *global context* cue and a *local context* cue. The global context cue activates each item in memory in the category of "animals that you know," and we assume that this is best approximated by the frequency of occurrence of each animal name in the Wikipedia corpus (e.g.,

Raaijmakers & Shiffrin, 1981; Hills et al., 2012). The *local context* cue activates each item in memory in relation to its semantic similarity to the cue currently held in primary memory, that is, the previous item recalled. Thus, the most recently recalled item is the cue used to query local memory, and activation is defined as the pairwise semantic similarities produced by BEAGLE (Jones & Mewhort, 2007) with all animals yet to be recalled. Note that for the global context cue, items are activated in proportion to their frequency. For the local context cue, items are activated in proportion to their semantic similarity with the most recently recalled item. Thus, given a set of cues, we can compute the predicted retrieval probability for any sequence of animal names by repeatedly using Equation 1.

As an example, if a participant recalled DOG, which shared a BEAGLE-based semantic similarity of 0.5 with CAT, then S('DOG','CAT') = 0.5. This would be placed in the numerator of Equation 1, while the sum of the semantic similarities over all animals would be placed in the denominator. If CAT were the next item recalled, followed by 'FOX', then S('CAT', 'FOX') would be in the numerator to compute the next probability of recall, with the activation for all remaining animals summed in the denominator. Using this framework to compute probabilities, *w* parameters were then fit to each participant to maximize the observed probabilities of recall and produce a maximum likelihood fit (see Lewandowsky & Farell, 2010, for an overview on model fitting).

We tested five models based on previous work (Hills et al., 2012; Hills & Pachur, 2012) differing in how cues were used to guide the search process. This allowed us to evaluate whether each additional assumption made by a model was warranted, based on whether or not it provided a significantly better fit to participants' recall sequences than a model without the additional assumption. The main distinction between the five models we tested concerns the nature of cue use, namely, whether the cues were used in a static or dynamic fashion, which has not been previously tested in a mature adult cohort.

Static models (Models 1, 2, and 3) used the same cue arrangement over the entire recall interval. Model 1 used a single global cue: frequency. This assumes that individuals recall sequences of animals

that reproduce their natural strength of activation in memory as a consequence of frequency alone. Model 2 used a single local cue: semantic similarity. This assumes that individuals rely only on the previously recalled item as a cue for the next recall, producing a chain of pairwise associated animals. Model 3 represents the simultaneous combination of global and local cues: both frequency and semantic similarity. This assumes a process of recall based on semantic similarity to the previous item that is further informed by the frequency of past experience with those items.

Dynamic models (Models 4, 5) assume that people transition between global and local cues as they search through memory. The logic underlying these models can be best understood by likening the search process in semantic space to foraging in physical space. The dynamic models we tested assume that the representational space is patchy or clustered and that people rely on global cues to travel from one patch to another while they rely on local cues to move within a patch. Search thus involves both traveling between distant patches and searching within a patch. In line with previous work, our models share the assumption that frequency alone guides search between clusters of items while both frequency and similarity guide search within a cluster (Raaijmakers & Shiffrin, 1981; Hills & Pachur, 2012; Hills et al., 2012).

Models 4 and 5 differ in their assumptions concerning when transitions between clusters occur and, consequently, they make different predictions about switches in the composition of the memory probe. Model 4 switches between local to global search when two successively recalled items do not share a category as defined by the cluster-switching hypothesis (Troyer et al. 1997). For example, if DOG were immediately followed by CAT during memory retrieval, since both are in the category 'pets', the model predicts local search. However, if DOG were followed by SHARK, since these do not share a category, the model predicts a transition to global search (frequency only) to recall SHARK. The Troyer et al. (1997; see also Troyer, 2000) categorization schema contains 22 non-exclusive categories, including "African animals", "water animals", and "beasts of burden." The categorization contained 155

unique animal names, but was extended by Hills et al. (2012) to contain an additional 214 animal names found in Wikipedia.

Model 5, which we call the similarity-drop model, implements a different view of transitions between patches. Previous research indicated that similarity drop was a plausible alternative hypothesis to Troyer et al. (1997) categories (Hills et al., 2012). Similarity drop defines switches between memory cues wherever a sequence of four items, A, B, C, D have BEAGLE-defined semantic similarities that follow the pattern $S(A,B) > S(B,C)$, and $S(B,C) < S(C,D)$. In words, the semantic similarity decreases between clusters and then increases again once search resumes locally with a subsequent cluster (for example, a sequence of A, B, C, D items might be DOG, CAT, SHARK, WHALE), with a similarity drop between CAT and SHARK. Previous work identified these similarity drops as places where individuals left a patch of locally semantically similar items to transition globally to another patch of locally semantically similar items (Hills et al., 2012). Note that, unlike model 4, model 5 prohibits two consecutive items to both represent global switches.

For all models, once an item was recalled, that item was removed from the retrieval structure. As in previous work (Hills et al., 2012; Hills & Pachur, 2012), removing or replacing items after retrieval had no affect on our conclusions. Also note that all dynamic models predict switches after they occur— and thus test where the most plausible locations for switches are, given the underlying representation. Predicting switches before they occur in fluency data is still an open question in the literature (see Hills et al., 2012 for further discussion).

Results

Participants recalled on average 16.8 (*SD* = 5.3; range 5–33) animal names. As expected, age was associated with recalling significantly fewer items. For all analyses we included a quadratic term, but this was the only case where the quadratic term was significant (for more information on linear and non-linear aging effects, see Verhaeghen & Salthouse, 1997; Salthouse, 2004). A regression with a

linear and quadratic term for age as the independent variables and total number of animals recalled as the dependent variable resulted in a significant effect for the linear term, $B = 0.41$, $t(182)=2.14$, $p = 0.03$, and quadratic term, $B = -0.004$, $t(183) = -2.80$, $p < .01$, $r = .33$, (see Figure 1). In all subsequent analyses, the quadratic terms were not significant, and we therefore report only the linear terms.

In what follows, we first present results showing that the modeling of semantic retrieval supports dynamic memory search. We then correlate age and digit span with the frequency of switching per item to address the three hypotheses for age-related memory decline.

Do People Switch Dynamically Between Retrieval Cues?

Table 1 presents the Bayesian Information Criterion (BIC) of models 1 to 5. BIC is a commonly used measure of model fit/comparison with penalization for the total number of free parameters to reduce overfitting (Lewandowsky & Farrell, 2011); smaller values of BIC indicate better model fit. The results show that the best single predictor of recoveries is frequency rather than similarity as suggested by the superior fit of Model 1 relative to Model 2. In turn, Model 3 outperformed both Models 1 and 2, showing that combining both frequency and similarity information was an improvement over single-cue models. Adding dynamic transitions between frequency and similarity further improved the model fit (compare Model 3 with Model 5), but only when using similarity drop to predict transitions between retrieval cues. Model 4, which predicted switches based on the Troyer et al. (1997) categorization scheme, did not produce an improvement over Model 3. This suggests that the Troyer et al. (1997) categories could benefit from additional computational evaluations to ascertain how they might be improved to capture patterns of retrieval in the fluency task.

Note that the poor performance of the Troyer et al. (1997) categorization scheme is a conservative estimate, because dynamic models use information from the data about the to-be-retrieved items to identify switching points. Thus, the dynamic models should outperform the static models if transition points are accurately identified. Crucially, while the difference in BIC between Models 4 and

5 is small (as it is based on only those locations where switches are predicted to occur, and thus represents a difference of approximately 5 data points per participant), the dynamic model using similarity drop provides a better explanation of the data for the majority of the participants. Moreover, this pattern is the same regardless of the age group considered: Using a median split on age (Median = 68), the similarity drop model provided the best fit to the data of both younger (56%) and older (53%) participants in our sample, with the remainder spread over the other four models. Overall, our results suggest that the process of recall is dynamic, consistent with both the basic assumptions of the cluster-switching and cue-maintenance hypotheses. Though these dynamics are not well characterized by the Troyer et al. (1997) cluster-switching categorization scheme, it may still be that age-related changes in switches-per-item are consistent with the cluster-switching hypothesis. In what follows, we focus on similarity drop switches (supported by the best fitting Model 5) to compare the three hypotheses—global slowing, cluster-switching, and cue-maintenance—by evaluating differences in switching across adulthood.

Is Age Related to Dynamic Switching Between Retrieval Cues?

Participants produced a mean of 5.09 similarity drop switches (SD = 1.73) over the course of the recall interval. The total number of similarity-drop switches derived from the dynamic model of memory search (Model 5) was negatively correlated with age, $r = -.21$, $t(183) = -2.89$, $p < .001$. However, the total number of similarity drop switches was also strongly correlated with the total number of recalls, $r = .88$, $t(183) = 24.96$, $p < .001$. This produces an obvious confound. Because switches may occur at regular intervals during recall, the reduction in switching may not represent a second search component as proposed by the cluster-switching hypothesis (Troyer et al. 1997), but simply the outcome of a single slower recall rate as suggested by the global slowing hypothesis (see also Mayr, 2002). For example, if there is a switch every four items, fewer total items recalled would lead to fewer total switches, and this would account for age-related differences in switching without the need for an age-related influence on

cue switching processes.

A more meaningful correlation with age is the number of *switches per item*, which corrects for the total number of items recalled. Switches per item are more weakly negatively correlated with total number of recalls, $r = -.16$, $t(183)= -2.16$, $p = .03$. More importantly, switches per item are positively correlated with age, $r = .17$, $t(183) = 2.30$, $p = .02$ (see Figure 2). Controlling for vocabulary, education, and gender in a multiple regression did not alter these results. Further, note that switches based on Model 4 (i.e., Troyer et al., 1997) were not significantly ($p = .10$) correlated with age. In sum, the number of switches between retrieval cues per item recalled correlates *positively* with age. This result is inconsistent with the global slowing hypothesis that predicts *no qualitative differences* in switching as a function of age, as well as the cluster-switching hypothesis that predicts *decreased* switching as a function of age. This result is, however, consistent with the cue-maintenance hypothesis that assumes *increased* switching between retrieval cues as a function of age-related decline in cognitive control.

Is Cognitive Control Related to Dynamic Switching Between Retrieval Cues?

According to the cue-maintenance hypothesis, age-related cognitive control deficits are responsible for the increase in switching between retrieval cues. Therefore, the effect of age should be similar to the effect of reduced cognitive control. We used digit span as a measure of cognitive control. Digit span ranged between 8 and 28, with a mean of 16.52 ($SD = 4.30$), and was not significantly correlated with age, $r = -.12$, $t(183) = -1.62$, $p = .11$. A multiple regression predicting similarity drop switches per item resulted in a significant independent effect of both digit span, $B = -0.002$, $t(183)= -2.10$, $p = .04$, and age, $B = 0.0006$, $t(183) = 2.06$, $p = .04$. Figure 3 shows the correlation between similarity drop switches per item and digit span, $r = -.15$, $t(183) = -2.34$, $p = .02$. Again, controlling for vocabulary, education, and gender in a multiple regression did not alter these results. Individuals with higher digit spans switched less often per item, supporting the cue-maintenance hypothesis, which predicts that reduced cognitive control abilities lead to increased switching between retrieval cues.

Note that a multiple regression predicting total number of recalls is still significant for age, $F(1,180)=12.17$, $p < .001$, after controlling for the effect of both similarity drop switches and digit span (both of which are significant; similarity drop switches, $F(1, 180)=5.58$, $p = .02$, and digit span, $F(1,180)=24.00$, $p < .001$). This indicates that these measures do not fully capture individual differences in cognitive control that contribute to age differences in memory search and that additional factors may also play a role in age-related memory decline. One possibility is that although the digit span measure can tap into individual differences in cognitive control it may not be powerful enough to detect age differences in these processes. Future studies should use more specific measures of cognitive control and inhibitory function (Friedman & Myake, 2004) to examine the link between memory search and cognitive control.

Discussion

We modeled the semantic search processes of adults between 27 to 99 years of age in a fluency task to investigate potential mechanisms of age-related decline in memory search. Our results suggest that people transition between local and global cues when searching long-term memory. Specifically, people appear to switch between global frequency-based and local item-based retrieval cues to navigate their semantic memory (cf., Gruenewald & Lockhead, 1980; Hills et al., 2012; Hills & Pachur, 2012; Raaijmakers & Shiffrin, 1981). Our results further show that the number of switches between global and local cues per item recalled increases with age. This is consistent with the cue-maintenance hypothesis, which suggests that age-related differences in cognitive control cause difficulties in focusing on specific item-level cues and thus lead to increased switching between global and local retrieval cues. This leads to less efficient memory search, because efficient memory search requires an appropriate balance between too little and too much focus on local search cues (Hills et al., 2012). The cue-maintenance hypothesis is supported with evidence that the number of switches between retrieval cues per item recalled was correlated with a measure of cognitive control (digit span)—with more switching correlated

with lower spans and fewer items recalled. This corroborates previous research showing a relationship between cognitive control and fluency tasks (Rosen & Engle, 1997; Unsworth et al., 2011; Hills & Pachur, 2012). It also further supports proposals that working memory plays an important role in the ability to maintain focus on specific cues (Kane & Engle, 2004; Hills et al., 2010) and may be associated with age-related cognitive decline (e.g., Paxton et al., 2008). Nevertheless, future studies should test the ability of cognitive control measures to predict switching relative to others measures of cognitive function, so as to further test the cognitive control hypothesis, that is, the idea that a specific deficit in control processes and not general cognitive decline underlies age differences in switching between memory representations.

Age-related global slowing may also contribute to age-related decline in memory search. However, our results indicate that this cannot be the only explanation, because the global slowing hypothesis proposes that age-related differences in memory retrieval are the result of slower memory processes overall. Thus, global slowing alone predicts no differences in cue utilization as a function of age (Mayr, 2002). Similarly, our results contrast with the cluster-switching hypothesis, which suggests that age-related differences in cognitive control are associated with decreased switching between retrieval cues per item recalled (Troyer et al., 1997). Indeed, our results indicate increased switching when using the best predictor of switches in dynamic memory search (similarity drop switches), and show no change in switches per item when using switches based on the Troyer et al. (1997) categorization. This suggests that the Troyer et al. (1997) categories, which were hand-coded, are potentially insufficiently accurate to assess switching behavior, and could benefit from future research on how categories are structured.

Several limitations of the present research should be addressed in future work. First, the study is cross-sectional so we cannot evaluate the extent to which the observed age differences are due to true age-related cognitive change or cohort differences in cognitive ability more generally. Statistically

controlling for education, vocabulary, and gender did not alter our results—but other cohort-related factors may be at work. Second, younger adults in their twenties and thirties are under-represented in the sample. Control analyses not reported here suggest this did not bias the results (e.g., including an indicator variable in our analyses for age above or below the median did not reveal a significant interaction in any of our analyses—suggesting the slopes are not different between age groups). However, future work is needed to more precisely examine age differences in the earliest stages of adulthood. Third, our models assume that age differences take place at the level of search. However, the underlying semantic representation may degrade with age, causing items in memory to have smaller local activation levels, and this may offer an alternative explanation (not based on changes in cognitive control) for increased switching (e.g., Borge-Holthoefer & Arenas, 2010). There is indeed evidence of age differences in vocabulary and priming tasks that are compatible with semantic representations differing between younger and older adults (Laver & Burke, 1993; Myerson, Ferraro, Hale, & Lima, 1992; Verhaeghen, 2003). However, evidence based on free associations suggests that representations do not degrade with age (Burke & Peters, 1986). With respect to the present study, the correlation between switches-per-item and digit span would not necessarily be predicted by such changes in representation. Future research could further clarify this matter, however, by testing search- versus representation-based explanations against one another by using models that have different representational assumptions. An alternative would be to focus on domains where it is possible to evaluate the representation directly, for example, by using a social fluency task based on participants' social networks (e.g., Hills & Pachur, 2012). Fourth, future research could utilize timing data to corroborate patterns of switching (as in Rosen & Engle, 1997) and provide additional data to be accounted for by cognitive models.

In sum, our work provides an example of how formal modeling can provide insights into the cognitive processes underlying adult age differences in memory search. Specifically, our results suggest

that there may be life span changes regarding cue use in memory search, with age-related deficits in cognitive control being associated with increased switching between memory cues.

References

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: current status and future prospects. *Neuroscience and Biobehavioral Reviews*, *30*, 791–807.

Bopp, K. L., & Verhaeghen, P. (2007). Age-related differences in executive control processes in verbal and visuo-spatial working memory: Storage, transformation, supervision, and coordination. *Journals of Gerontology: Psychological Sciences, 62*, 239–246.

Borge-Holthoefer, J. & Arenas, A. (2010). Semantic networks: Structure and dynamics, *Entropy*, *12*, 1264–1302.

Bowden, S. C., Petrauskas, V. M., Bardenhagen, F. J., Meade, C. E., & Simpson, L. C. (2012). Exploring the dimensionality of digit span. *Assessment*. Advance online publication. doi:10.1177/1073191112457016

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience and Biobehavioral Reviews*, *26*, 809–817.

Buckner, R. L. (2004). Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. *Neuron*, *44*, 195–208.

Burke, D. M., & Peters, L. (1986). Word associations in old age: Evidence for consistency in semantic encoding during adulthood. *Psychology and Aging*, *1*, 283-292.

Crum R.M., Anthony J.C., Bassett S.S., & Folstein M.F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. Journal of the American Medical Association, *269* (18), 2386–91. doi:10.1001/jama.1993.03500180078038

Folstein, M., Folstein, S., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Gronlund, S. D., & Shiffrin, R. M. (1986). Retrieval strategies in recall of natural categories of

categorized lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 550–561.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 225–239.

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the aging mind: A review from cognitive neuroscience. *Nature Reviews Neuroscience*, *5*, 7–96.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*, 431-440.

Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 218–228.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: exploration and exploitation in generalized cognitive search processes. *Journal of Experimental Psychology: General, 139,* 590–609.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Peterson, R. C., Kokmen, E., Kurland, L. T. (1992). Mayo's older Americans normative studies: WAIS-R norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6*, 1-30.

Jones, M. N., Kintsch, W., & Mewhort, D. J. K. (2006). High-dimensional semantic space accounts of priming. *Journal of Memory and Language, 55,* 534–552.

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 104*, 1–37.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, Cognition*, *26*, 333–358.

Kozora, E., & Cullum, C. M. (1995). Generative naming in normal aging: Total output and qualitative

changes using phonemic and semantic constraints. *The Clinical Neuropsychologist, 9,* 313–325.

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15*, 196–204.

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age: A meta-analysis. *Psychology and Aging, 8*, 34–43.

Lewandowsky, S. , & Farell, S. (2010). *Computational Modeling in Cognition: Principles and Practice*. London: Sage Publications

Li, S. -C., Lindenberger, U., & Bäckman, L. (2010). Dopaminergic modulation of cognition across the life span. *Neuroscience and Biobehavioral Reviews*, *34*, 625–630.

Lindenberger, U., Mayr, U., & Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging, 8,* 207–220.

Mayr, U. (2002). On the dissociation between clustering and switching in verbal fluency: Comment on Troyer, Moscovitch, Winocur, Alexander, and Stuss. *Neuropsychologica, 40*, 562–566.

Mayr, U., & Kliegl, R. (2000). Complex semantic processing in old age: does it stay or does it go? *Psychology and Aging, 15*, 29–34.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging, 7*, 257–270.

Paxton, J. L., Barch, D. M., Racine, C. A., & Braver, T. S. (2008). Cognitive control, goal maintenance, and prefrontal function in healthy aging. *Cerebral Cortex*, *18*, 1010–1028.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126,* 211–227.

Rosselli, M., Tappen, R. M., Williams, C. L., Salvatierra, J. Zoller, Y. (2009). Level of education and category fluency task among Spanish speaking elders: Clustering and switching strategies. *Aging, Neuropsychology and Cognition, 2*, 1–24.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Salthouse, T., A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*, 140–144.

Thurstone, L. L., (1938). *Primary mental abilities*. University of Chicago Press, Chicago.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratigied by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*, 167-177.

Tomer, R., & Levin, B. E. (1993). Differential effects of aging on two verbal fluency tasks. *Perceptual and Motor Skills, 76,* 465–466.

Troyer, A. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, *22*, 370–378.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*, 138-146.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *The Quarterly Journal of Experimental Psychology*, *64*, 447–466.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging, 18*, 332–339.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, *122*, 231–249.

Walker, W. H., & Kintsch, W. (1985). Automatic and strategic aspects of knowledge retrieval. *Cognitive Science*, *9*, 261–283.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale: Revised Manual (WAIS-R)*. New York: Psychological Corporation.

Table 1

Median BIC for models of memory recall proposed to explain animal fluency data.

Model	Model Characteristics	Cue(s)	Switches	Median BIC	Median w	Median No. of Switches	Best Fit (%)
1	Single-cue, Static	Frequency	—	84.8	2.22	—	8
2	Single-cue, Static	Similarity	—	88.5	1.33	—	0
3	Multiple-cue, Static	Frequency	—	82.3	1.80	—	17
		Similarity			1.06		
4	Multiple-cue, Dynamic	Frequency	Troyer	81.2	2.02	9	19
		Similarity			1.39		
5	Multiple-cue, Dynamic	Frequency	Similarity Drop	79.4	1.88	5	55
		Similarity			1.28		

Note. Models were fit using the maximum likelihood method to find the optimal w for each Model/Cue combination for each individual.

Figure 1. Number of animals recalled during a one minute interval as a function of age. Each point represents a single individual. The line is the best fit regression, which used the linear and quadratic terms reported in the text.
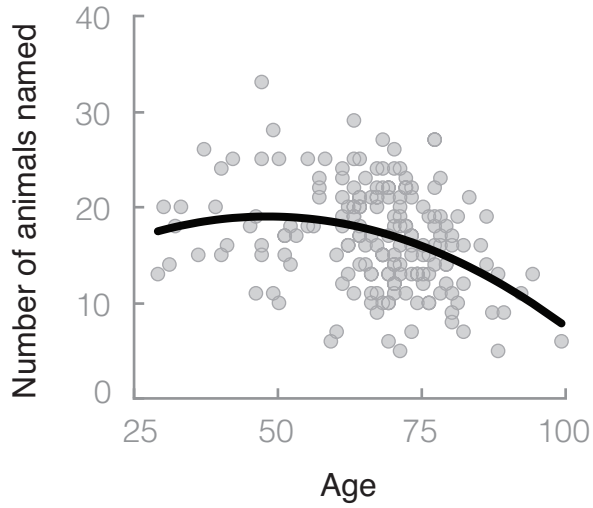
Figure 2. Similarity drop switches per item as a function of age. Similarity drop switches were computed as described in the text, representing places in the production interval where the sequence of recalled animals showed very low and then very high semantic similarity as measured by BEAGLE. The line is the best fit regression line ($B = 0.0007$, $r = .19$).
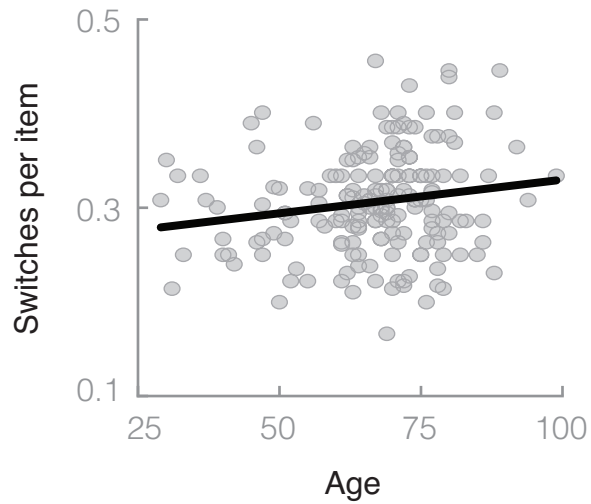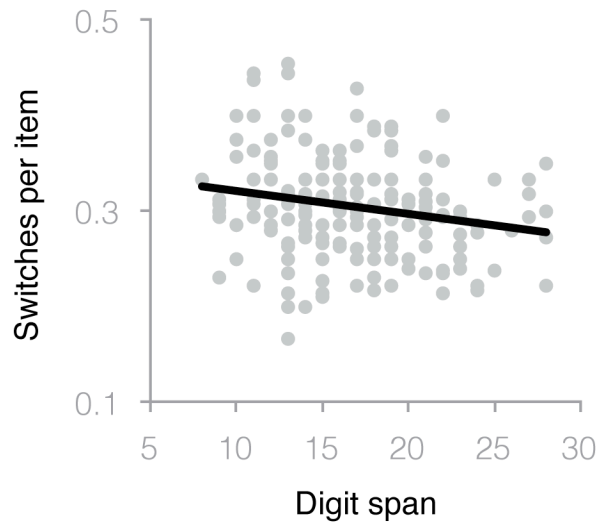
Figure 3. Similarity drop switches per item as a function of digit span. The line is the best fit regression line ($B = -0.002$, $r = -.20$).

Appendix A.

Participant counts, and means (standard deviations) for education, vocabulary, digit span, and animals recalled by age decade.

Age Range	#	Education	Vocabulary	Digit Span	Animals recalled
29–39	8	15.9 (2.9)	41.5 (7.4)	17.5 (3.6)	18.3 (4.2)
40–49	13	14.3 (2.0)	45.0 (14.5)	16.8 (2.7)	19.7 (6.8)
50–59	16	15.7 (2.2)	43.5 (14.7)	16.1 (4.8)	18.2 (5.4)
60–69	62	15.3 (2.3)	44.6 (10.9)	17.1 (4.7)	17.5 (5.1)
70–79	65	15.5 (2.7)	43.4 (12.6)	16.4 (4.2)	16.5 (4.8)
80–89	18	13.8 (2.9)	34.0 (16.8)	15.1 (3.9)	12.8 (4.6)
92–99	3	13.3 (4.2)	40.7 (14.6)	15.3 (3.2)	10.0 (3.6)

Endnotes

[i] Our approach takes the falsifiable position made by Mayr (2002) that slowing is proportional across cognitive faculties and as such "slowing alone predicts no differences in cue utilization as a function of age." Note that a slowing hypothesis that leaves room for age-related changes in switching—without stating how switching will change—is unfalsifiable. Specific interpretations of global slowing that do not assume proportional slowing may be viable, but only those consistent with the increased switching predicted by the cue-maintenance hypothesis are supported by the present data. Indeed, future research should investigate the assumptions necessary for non-proportional global slowing models to predict the observed results

[ii] To be even more conservative, we also tested all of our hypotheses after excluding an additional 6 individuals who performed in the lower 5th percentile of age-normed digit span or fluency (Ivnik et al., 1992; Tombaugh, Kozak, & Rees, 1999). Results lead to the same statistical conclusions as those presented, and indeed do so even if we exclude individuals in the lower 10th percentile. However, it is impossible to exclude all pathological processes with any test, as some may yet still be undefined, and thus our results—like all aging studies—may reflect influences beyond normal aging.

[iii] Forward and backward digit span were combined at the time of data entry. Thus, we cannot determine whether the same results would be obtained if forward and backward span were considered separately. However, recent research in this area suggests there is "no evidence of a discriminant validity pattern" between these two measures (Bowden, Petrauskas, Bardenhagen, Meade, & Simpson, 2012).