STUDIES IN HUMAN AUTOBIOGRAPHICAL MEMORY AND RODENT SPATIAL NAVIGATION

by

Robert Gardner
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Psychology

Committee:

_________________________________ Director

____

____

_________________________________ Department Chairperson

_________________________________ Program Director

_________________________________ Dean, College of Humanities and Social Sciences

Date: ___________________________ Fall Semester 2014
George Mason University
Fairfax, VA

Studies in Human Autobiographical Memory and Rodent Spatial Navigation

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Robert Gardner
Bachelor of Science
The Pennsylvania State University, 2003

Director: Giorgio A. Ascoli, Professor
Molecular Neuroscience Department
Department of Psychology

Fall Semester 2014
George Mason University
Fairfax, VA

DEDICATION

I dedicate this dissertation to my family.

ACKNOWLEDGEMENTS

I would like to thank Drs. Giorgio Ascoli and Ted Dumas who have given me invaluable scientific training and support throughout the years. I would also like to thank Dr. Linda Chrosniak for her encouragement and guidance, and Dr. Matteo Mainetti and the host of student researchers who helped advance my dissertation projects. I am grateful to the members of the Center for Neural Informatics, Structures, and Plasticity of the Krasnow Institute for Advanced Study for their support and assistance.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

Appendix:

LIST OF ABBREVIATIONS

Activity-Regulated Cytoskeleton-Associated Protein .... Arc
Activity-Regulated Gene 3.1 .... Arg 3.1
Autobiographical Memory .... AM
Dentate Gyrus .... DG
Dorsal lateral striatum .... DLS
Dorsal medial striatum .... DMS
Froot Loop .... FL
Immediate Early Gene .... IEG
In Situ Hybridization .... ISH
Memories Per Hour .... MPH
Opposing T's .... OpT
Plus Win-Shift .... P-Wsh
Prospective Memory .... PM
Supplemental Information .... SI
Years Old .... yo

ABSTRACT

STUDIES IN HUMAN AUTOBIOGRAPHICAL MEMORY AND RODENT SPATIAL NAVIGATION

Robert Gardner, Ph.D.

George Mason University, 2014

Dissertation Director: Dr. Giorgio A. Ascoli

This dissertation takes a multi-species approach to investigate several types and focuses of memory. The first set of experiments quantified two important, yet relatively unexplored, dimensions of human recollection: its content and frequency. These studies primarily focused on memory of personally-experienced events (autobiographical memory: AM). In particular, participants reported the number of details (among specified content categories, e.g., *People*, *Places*, and *Things*) that comprised word-cued AMs dated to various life periods (see http://cramtest.info). Application of this methodology to subjects from numerous age groups revealed several age-related effects on memory content. Notably, the amount of detail associated with typical AMs increased with the age of the participant, and decreased with the age of the memory. To estimate the occurrence of recollection in natural settings, an experience sampling approach was utilized. In addition to targeting AM, this design probed recollections of to-be-experienced events (termed here, prospective memory: PM). Younger subjects experienced AM and PM

equally often. In contrast, while older adults engaged in AM as often as younger subjects, they experienced PM about twice as frequently. In addition, AM and PM occurrence rates were positively correlated, most strongly among younger individuals. Altogether, these experiments on human memory demonstrate differences between younger and older adults both in how personal events are remembered and in the temporal focus of typical recollection.

The second set of experiments investigated, in rodents, the interplay between different decision-making strategies (which rely upon distinct memory types) during the execution of spatial navigation tasks. A place strategy is attentive, depends on memory of the spatial arrangement of landmarks, and is used to flexibly locate the position of a goal. In contrast, a response strategy is generally thought to be automatic, and engages a fixed motor sequence. Replicating prior work, rats increased their reliance on response navigation with repeated training on a dual-solution choice (i.e., one that can be solved using either strategy), suggesting an experience-dependent shift from attentive to automatic performance. Additionally, during serial navigation, when a subsequent (secondary) choice required spatial working memory, the strategy transition on the dual-solution (primary) choice was blocked. However, once response navigation was established on the primary choice, subsequent initiation of secondary spatial training did not increase the use of primary place strategies. Taken together, these results suggest that strategy reliance is sensitive to the cognitive demand of subsequent behaviors—an influence from which well-practiced actions are protected. In a related experiment, neural activation, based on activity-dependent gene expression, was estimated in structures

implicated in place (e.g., hippocampus) and response (i.e., dorsolateral striatum) navigation among rats assigned either to a relatively brief or to a protracted training schedule on a dual-solution task. Results suggest that patterns of activation across neural structures may reflect the experience-dependent emergence of response navigation. Implications and limitations of all experiments are discussed.

CHAPTER ONE: INTRODUCTION

Memory is not a unitary phenomenon. Rather, memory can refer to several types and focuses of retrieval. For example, among other functions, it confers the ability to relive, imagine, and plan experiences, recall knowledge of the world, and perform complex tasks and behaviors. Characterization of memory according to unique sub-types can be traced throughout history. Nineteenth and twentieth century psychologists (James, 1950; Tolman, 1948; also see Ryle and Dennett, 2000) have suggested a distinction between memories that are recollected consciously and those which are expressed by learned actions. This division has been described as “knowing that” versus “knowing how,” and illustrated using contrasting terms such as memory versus habit, and declarative memory versus procedural memory.

Evidence of dissociated memory systems

From a neurocognitive perspective, the existence of distinct memory systems has been highlighted through observation of individuals who endured specified brain damage. For example, patient H.M. was diagnosed with severe epilepsy. To minimize its impact on his everyday activities, surgery was performed to remove bilaterally portions of his medial temporal lobes, including his hippocampus (Corkin, 1968; Corkin et al., 1997; Milner, Squire and Kandel, 1998; Scoville and Milner, 2000). The magnitude of his seizures was attenuated, but an unexpected effect of the surgery arose. Although H.M.

demonstrated some degree of retrograde amnesia, he could generally recall personal events preceding the surgery; however, he could no longer remember events and experiences transpiring after the surgery, and had trouble acquiring new knowledge of the world. In stark contrast, he was able to acquire new procedural or motor abilities, e.g., as measured through performance of mirror drawing and rotary pursuit tasks (Corkin, 1968; Milner, 1962; Scoville and Milner, 2000). The specificity of H.M.'s amnesia suggested that indeed memory is expressed in multiple forms and that these forms rely on distinct brain regions. In particular, these observations fit the view that a propositional or declarative memory system (reliant on the hippocampus and surrounding areas) enables the formation of memory for facts and events, and a disparate non-declarative memory system (reliant on areas not housed within the medial temporal lobes) permits the acquisition and expression of actions or skills.

The declarative system can be further elaborated based on observation of additional patients who exhibit varied patterns of memory impairment (e.g., Rosenbaum et al., 2005; Vargha-Khadem et al., 1997). For example, patient K.C.'s hippocampus was damaged bilaterally as a result of a motorcycle accident. K.C. could no longer relive the details associated with personally-experienced past episodes, but had relatively little trouble remembering general factual information (Rosenbaum et al., 2005). These findings suggest a distinction in the neurological substrates of richly detailed remembering of experiences (episodic memory) and retrieval of context-free knowledge (semantic memory; Tulving, 1972, 1985). Tulving (1985) described this distinction as remembering versus knowing. Notably, in addition to exhibiting deficits in past-oriented recollection, K.C. similarly had difficulties imagining future events. Consistent findings were reported in other patients who displayed qualities of temporal lobe amnesia (e.g., Tulving, 1985). In light of these observations, taken together with findings among healthy individuals (e.g., Botzung, Denkova and Manning, 2008; Schacter and Addis, 2007), it appears that, among other regions, the hippocampus is integral to both past- and future-oriented context-specific or episodic recollection.

Non-declarative memory describes several forms of learning and retrieval (e.g., procedural memory, priming, and classical conditioning), each of which is proposed to rely on a distinct set of neural circuits (for a full review of long-term memory systems see Squire, 2004). These non-declarative memories are expressed in terms of action or performance, as opposed to recollection, and may be retrieved without conscious awareness. Frequently, the declarative system is contrasted with procedural learning, which includes memory for skills that are incrementally acquired over time (and may become habitual). Mirroring patients with hippocampal disruption who typically show pronounced deficits in declarative (and less so in procedural/habitual) memory, patients with striatal disruption generally demonstrate impairments of procedural/habitual learning (and less so of declarative recall; e.g., Knowlton, Mangels and Squire, 1996; Schmidtke, Manner, Kaufmann and Schmolck, 2002; Squire, 2004; Tranel, Damasio, Damasio and Brandt, 1994). Importantly, these conclusions, drawn from observation of patients with targeted brain damage, have been largely corroborated through neuroimaging studies of healthy individuals (e.g., Poldrack, Prabhakaran, Seger and

Gabrieli, 1999; Poldrack et al., 2001), supporting the multiple systems hypothesis of human memory.

Multiple memory systems across species

This dissociation between brain regions and cognitive abilities noted in humans appears to be remarkably well conserved across species, e.g., as observed during spatial navigation in rodents. In many instances, performance of a spatial navigation task can rely on flexible decision-making that engages context-specific memory of the position of a goal (place strategies), or inflexible behaviors (well-learned fixed routes) that lead to the goal (response strategies). The distinction between these navigational strategies was emphasized in the mid-twentieth century when psychological research was dominated by behaviorism, and hypothesized mechanisms of memory were confined within its theoretical framework, asserting that learning stems from inflexible stimulus-action (response) associations. Tolman (1938, 1948), however, showed that a portion of rats used an alternative and flexible strategy (place) proposed to leverage a mental or cognitive map of the spatial layout of the environment, and to be associated with an expectancy of an outcome.

Subsequent research suggested that both strategies may underlie performance and that their relative contributions to behavior can be modulated by several factors. For example, a multitude of studies has documented an experience-dependent shift in strategy engagement. Across training sessions on a dual-solution task (one that can be solved using either strategy), place strategies underlie performance early in training whereas response strategies become dominant after extensive practice and experience, when habits are thought to emerge. Investigation into the neural correlates of these strategies revealed that place navigation relies on an intact hippocampus, and response navigation relies on an intact striatum (Packard, 1999; Packard, Hirsh and White, 1989; Packard and McGaugh, 1996).

Thus, place navigation appears to rely on the same memory system as episodic memory: they both demand hippocampal-dependent contextual retrieval. Supporting this view, medial temporal lobe patients who show deficits in episodic recollection also display marked impairments in spatial (place) memory tasks (e.g., Aradillas, Libon and Schwartzman, 2011; Holdstock et al., 1999; Vargha-Khadem et al., 1997). In contrast, response learning is synonymous with procedural or habit memory; it is associated with inflexible striatal-dependent expression of incrementally acquired actions. Additionally, these same findings summarized from studies in rodents translate to human spatial navigation (e.g., Bohbot, Iaria and Petrides, 2004; Iaria et al., 2003; Schmitzer-Torbert, 2007). Therefore, investigation into decision-making strategies using animal models holds great potential to further our understanding of human memory and its dissociated systems.

Dissertation outline and organization

These seminal discoveries in human memory and rodent navigation suggest that memory is best described from a multiple systems perspective. Unsurprisingly, they have also spurred extensive research to further understand and characterize dissociated forms of memory, e.g., their characteristics, inter-relationships, and fundamental neural mechanics. This dissertation presents two main efforts which build off previous approaches that probe human episodic recollection and rodent spatial navigation.

The first series of experiments examines episodic recall in humans. Relevant to the current work, relatively recent investigations into episodic memory (Tulving, 1972, 1985) have focused on naturalistic recollections of the unique personally-experienced events of our lives (i.e., autobiographical memories; Berntsen and Rubin, 2012; Rubin, 1988). In particular, this dissertation presents research which quantifies the content (the number of details within specified categories: e.g., *People*, *Places*, and *Things*) associated with autobiographical memories across the life span from various age groups. This approach extends previous studies which predominantly focused on evaluation of phenomenological qualities of memory (e.g., reliving, vividness, uniqueness, importance; see Johnson, Foley, Suengas and Raye, 1988; Rubin and Schulkind, 1997a; Rubin, Boals and Berntsen, 2008; Sutin and Robins, 2007), rather than counts of reported detail.

Additionally, relying on an experience sampling procedure, estimates of the frequency at which autobiographical recollections occur in natural settings are presented for younger and older subjects. The occurrence rates of autobiographical memories are contrasted with occurrence rates of future-oriented thoughts (i.e., thoughts related to potential future experiences and planned actions: termed here, prospective memory; Atance and O'Neill, 2001; McDaniel and Einstein, 2007). Thus, these data shed light on the typical temporality of recollection.

In particular, Chapter Two introduces the methodologies employed to quantify autobiographical memory content and occurrence rates, and presents data collected from college-aged subjects. Chapter Three extends this line of research by presenting work that applies the same methodology used to assess memory content to individuals across the adult life span. Results suggest that while recall is relatively stable across age groups, there are several notable differences in how past experiences are remembered and recounted between younger and older adults. For example, the amount of detail reported for a given recollection is generally increased among older adults and decreased among older memories, showing two opposing influences of age on retrieval. Chapter Four presents data which quantify retrieval frequencies of autobiographical and prospective memory among younger and older subjects. These data highlight a drastic age-related change in the relative time spent engaged in past and future episodic thought. Notably, older subjects appear to engage in thought associated with future events almost twice as often as younger subjects and twice as often as they engage in autobiographical recollection.

The second series of experiments focuses on spatial navigation in rodents in order to understand the interplay between place and response memory systems during task performance. Previous findings on the experience-dependent transition from place to response navigation suggest that these systems may compete to control behavior at a single choice point (e.g., Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999; also see Packard, 1987). However, whether these systems interact during serial choice learning remains an open question. Chapter Five introduces a novel two-choice maze that can be used to assess interactions between place and response navigation during serial spatial navigation tasks. Results from these studies demonstrate several interactions between navigational strategies that are not isolated to a single choice and offer insight into the malleability of the experience-dependent strategy transition. For example, spatial working memory training on the second of two choices appears to prevent the expression of response navigation on the first choice that otherwise occurs with task repetition. However, this effect is only present when the secondary spatial task is administered early in training; once response navigation is established, secondary training does not increase the use of place strategies.

In the Appendix, preliminary studies evaluate the relationship between region-specific neural activation (estimated by activity-dependent immediate early gene expression: Arc/Arg 3.1) and the experience-dependent strategy transition. Results are consistent with straightforward hypotheses on hippocampal and striatal contributions to performance: activation patterns across structures correlate with strategy recruitment and/or indicate the emergence of response navigation that occurs with task repetition.

Each chapter in this dissertation corresponds to a free-standing manuscript, written over the past several years. Two have been published (Chapters Two and Five) and two are in the submission process (Chapters Three and Four). The data presented across chapters are generally analyzed and interpreted in a consistent manner. However, as these manuscripts were written at different times, chapter-to-chapter variability e.g., in how data are analyzed, reflects changes in the authors' perspectives and acquired knowledge during the intervals between manuscript constructions, as well as variability in journal requirements and reviewer criticisms.

CHAPTER TWO: QUANTITATIVE MEASUREMENTS OF AUTOBIOGRAPHICAL MEMORY CONTENT

[Gardner, R. S., Vogel, A. T., Mainetti, M. and Ascoli, G. A. (2012) Quantitative measurements of autobiographical memory content. *PloS One*, *7*(9), e44809. doi:10.1371/journal.pone.0044809] [Used with permission from the publisher]

Abstract

Autobiographical memory (AM), subjective recollection of past experiences, is fundamental in everyday life. Nevertheless, characterization of the spontaneous occurrence of AM, as well as of the number and types of recollected details, remains limited. The CRAM (Cue-Recalled Autobiographical Memory) test (http://cramtest.info) adapts and combines the cue-word method with an assessment that collects counts of details recalled from different life periods. The SPAM (Spontaneous Probability of Autobiographical Memories) protocol samples introspection during everyday activity, recording memory duration and frequency. These measures provide detailed, naturalistic accounts of AM content and frequency, quantifying essential dimensions of recollection. AM content (~20 details/recollection) decreased with the age of the episode, but less drastically than the probability of reporting remote compared to recent memories. AM retrieval was frequent (~20/hour), each memory lasting ~30 seconds. Testable hypotheses of the specific content retrieved in a fixed time from given life periods are presented.

Introduction

Autobiographical memories (AMs) are recollections of the first-person experience of past episodes. They refer to spatially and temporally specific events rather than factual (semantic) knowledge about the world (Manns, Hopkins and Squire, 2003; Tulving, 1972, 1985; Vargha-Khadem et al., 1997). From the neuropsychological perspective, AMs are long-term, i.e., the potential for their retrieval lasts from minutes to the entire life span, and are distinguished in the underlying brain organization from short-term or working memories (Baddeley and Warrington, 1970; Bayley, Hopkins and Squire, 2006; Remondes and Schuman, 2004; Squire, Knowlton and Musen, 1993).

AMs are believed to subserve fundamental thoughts and behaviors (Bluck, Alea, Habermas and Rubin, 2005; Pillemer, 1992) and important dimensions of autobiographical recall have been extensively characterized. Some psychophysical quantities have been measured objectively, such as the time to report a memory in response to a cue (e.g., Barsalou, 1988; Rubin and Schulkind, 1997a, 1997b). Among other aspects, emotional level, importance, and rehearsal have been rated on a Likert scale (Johnson, Foley, Suengas and Raye, 1988; Rubin, Schrauf and Greenberg, 2003). In addition, AM accuracy, intensity, and retrieval efficacy have been probed by systematically documenting daily events over extended periods of time (Linton, 1986; Wagenaar, 1986; White, 1982).

The temporal distribution of AMs has also been studied comprehensively. Galton (1879) initially described a systematic protocol for eliciting his own AMs in order to sample their distribution over his life span. Crovitz and Schiffman (1974) revised this method with the introduction of the cue-word technique: single words are sequentially presented to participants as prompts for generating memories, which are labeled for later recall and subsequently dated to when the recalled event had occurred. An effect of specific cue words on the resulting temporal distribution was soon documented (Robinson, 1976), and the consequent adoption of that same fixed word set in following studies created a *de facto* standard protocol of this cue-word technique (Fitzgerald and Lawrence, 1984; Jansari and Parkin, 1996; Rubin and Schulkind, 1997a, 1997b; Rubin, 1982; Rybash and Monaghan, 1999). The resulting studies yielded a reliable characterization of three components of the temporal distribution of AMs: the retention function, the reminiscence bump, and childhood amnesia. Retention of AMs declines steeply according to a power decay backward from the present day (Rubin, 1982; Rubin, Schulkind and Rahhal, 1999). The reminiscence bump is an increase in the relative recall of episodes that occur between 10 and 30 years of age (Rubin, Wetzler and Nebes, 1986; Rubin, Rahhal and Poon, 1998), and is best observed in adults older than ~45 years. Childhood amnesia is a drastic reduction of episodes recalled from 0 to 4 years of age (Rubin, 2000). Notably, these characteristic changes are considerably robust to aging (Rubin and Schulkind, 1997a, 1997b), pathology (Fromholt and Larsen, 1991; Fromholt et al., 2003), the sensory modality used to elicit recollections (Chu and Downes, 2000; Rubin, Groth and Goldsmith, 1984) and the cueing technique (Fromholt et al., 2003).

Despite the history of cognitive and clinical research, certain dimensions of AM, namely the number and types of details comprising the recollection (i.e., content) and its spontaneous rate of occurrence (i.e., frequency), have received relatively little attention.

Nevertheless, these dimensions of autobiographical recollection play pivotal roles in AM theory. For example, Conway and Pleydell-Pearce (2000) proposed a hierarchical organization of AM termed the *autobiographical knowledge base* (also see Conway, 2005). This theory places abstracted thematic knowledge and summarized events at the top, and specific details of distinct episodes rich with context at the bottom. The dynamic relation between these components would support reminiscence: abstracted knowledge comprising one's life story is constrained by the sensory-perceptual and contextual details of individual events; in turn, storage of episodic details in long-term memory and their retrieval are influenced by abstracted knowledge. Several hypotheses on the relationship between the amount and types of details from individual events and the organization of knowledge structures according to certain features (e.g., person, place, feeling, time; Barsalou, 1988), emerge from this model. For example, the amount of detail for selected features could predict the probability that an AM will be incorporated into certain knowledge sets expounding how episodic detail and abstract knowledge come together during memory storage and retrieval.

AM content also features prominently in theories of *source monitoring*, i.e., the process of retrieving and assigning the information in a memory to the original source (e.g., person A or person B). Under a framework set forth by Johnson and Raye (1981; also see Johnson, Foley, Suengas and Raye, 1988; Johnson, 1997), evaluating the amount and type of recalled detail (e.g., emotion, location, time) is key for accurate source determinations. When the detail of various events (e.g., real and imaged) is uncharacteristic, mistakes may arise. Misattribution of select details can have drastic consequences, e.g., as observed in eyewitness testimony, and pathology (Johnson, 1997).

The importance of both AM content and frequency is also highlighted in cross-sectional research. Aging and major neurodegenerative diseases selectively impair episodic and contextual memory (Spencer and Raz, 1995), as compared to e.g., semantic memory (Vargha-Khadem et al., 1997). This observation also applies to recollection of AM content (Fromholt and Larsen, 1991; Fromholt et al., 2003; Hashtroudi, Johnson and Chrosniak, 1990; Levine, Svoboda, Hay, Winocur and Moscovitch, 2002), including source monitoring (Cohen and Faulkner, 1989; Johnson, O'Connor and Cantor, 1997). Similarly, the number of AMs elicited by the cue-word method, extracted from participant narratives or collected through participant diaries (Schlagman, Kliegel, Schulz and Kvavilashvili, 2009) also declines in aging and pathology (Fromholt and Larsen, 1991; Fromholt et al., 2003). Moreover, specificity of AM content, when preserved in aging, is suggested to be a result of frequent rehearsal (Cohen and Faulkner, 1989; Cohen, 1998). Altogether, theories on AM structure and function, and age-related and pathological impairment, identify content and frequency as significant AM dimensions, urging their comprehensive measurement through the life span.

Several methods have been utilized to characterize aspects of AM content. The autobiographical memory interview was developed to assess the extent of what can be remembered by amnesic patients (Kopelman, Wilson and Baddeley, 1989). The memory characteristics questionnaire (MCQ; Johnson et al., 1988) and the autobiographical memory questionnaire (Rubin, Boals and Berntsen, 2008) used Likert scales to rate, among other features, spatial and temporal specificity, vividness, and sensory detail. Subjective content has also been quantified by counting the number of categorical details (a measure complementary and distinct to rating scales; Dawes, 2008) in written or spoken narratives of recalled autobiographical events (Addis, Musicaro, Pan and Schacter, 2010; Berntsen, 2002; Johnson et al., 1997; Johnson, Kahan and Raye, 1984; Levine et al., 2002) or experimentally created “autobiographical” episodes (Hashtroudi et al., 1990; Johnson et al., 1997).

To the best of our knowledge, the rate of spontaneous AM occurrence has not been measured. A widely used method that holds potential to reveal the frequency of typical AMs provides participants with journals to document and annotate AMs as they occur in daily life. Several such diary-based studies investigated differences between voluntary and involuntary AMs. Voluntary AMs are retrieved deliberately, and involuntary AMs, while consciously recollected, are retrieved without intention. Although research designs included both qualitative (e.g., content ratings) and quantitative (e.g., temporal distributions, counts of detail) measures, a focus on voluntary-involuntary AM comparison precluded frequency assessment. Specifically, these studies limited daily AM documentation, influenced voluntary AM retrieval, or did not collect a precise time window of diary utilization (Berntsen and Hall, 2004; Berntsen and Jacobsen, 2008; Johannessen and Berntsen, 2010; Schlagman and Kvavilashvili, 2008; Schlagman et al., 2009).

Here we report quantitative measurements of content and frequency of everyday AMs from numerous life periods, obtained using two new naturalistic methods: 1) The

Cue-Recalled Autobiographical Memory (CRAM) test, and 2) The Spontaneous Probability of Autobiographical Memories (SPAM) protocol.

The CRAM test elicits AMs using a word-set designed to replicate everyday written and spoken language cues. Participants are then asked to identify the age of each cued AM, and count the number of identifiable details within specified categories, similar to those investigated in the AM literature and highlighted as important components in AM theory (Johnson and Raye, 1981; Johnson et al., 1988; Johnson, 1997; Johnson et al., 1997; Levine et al., 2002; Nadel, Samsonovich, Ryan and Moscovitch, 2000), e.g., temporal and spatial details, persons, objects, and emotions. Such categories have also been shown to be age-sensitive (e.g., Hashtroudi et al., 1990; Levine et al., 2002). CRAM adapts the cue-word method (Crovitz and Schiffman, 1974) and combines it with content assessments, asking the participant to count the occurrence of different features in an AM. This approach offers several advantages over previously used methods. The novel cue-set, designed to mimic natural language cuing experiences, should elicit AMs more closely comparable to those retrieved under real-life conditions. Moreover, using participant-counts rather than time-intensive experimenter-scored narratives (Addis et al., 2010; Hashtroudi et al., 1990; Johnson et al., 1997; Levine et al., 2002) greatly reduces the data collection workload. This permits analyses of more numerous AMs, thus increasing representation and temporal resolution across the life span. Much like seminal studies of AM temporal distribution (Crovitz and Schiffman, 1974) reliably quantified AM retention (Rubin and Wenzel, 1996), counts of feature-specific details of AMs naturalistically sampled across the life span can provide the foundation for quantitatively characterizing feature-specific retention. This methodology has already been adapted to an internet-based application (http://cramtest.info).

The SPAM protocol expands experience sampling techniques to measure the probability and duration of AM recall during everyday life, yielding estimates of the number of AMs experienced in a given time period. SPAM was inspired by and adapted from an original experiment by Brewer (1988) in which participants carried a buzzer that prompted them at random times to annotate their behavioral and mental states and surrounding events for later analysis. This approach has been employed, among other applications, to assess visual activities (Rah, Walline, Lynn Mitchell and Zadnik, 2006) and psychopathology such as mood disorders (Peeters, Nicolson, Berkhof, Delespaul and deVries, 2003). The technique provides several advantages over assessments performed in clinical settings, including real-time monitoring, elimination of retroactive reporting errors, and performance evaluation in the natural environment (Trull and Ebner-Priemer, 2009). In addition, this technique reduces the participant workload as compared to a strictly diary-based research design. To our knowledge, SPAM is the first application of experience sampling to measure AM probability and duration, resulting in the first quantification of spontaneous AM frequency.

Combined data from CRAM and SPAM enable previously inaccessible estimations related to AM recall, e.g., the average number of different features (people, feelings, etc.) recalled from distinct life periods in a given time window. Such a computation assumes that the temporal distributions and content of naturalistically-cued experimental AMs and everyday occurring AMs are similar. While this assumption is untested, the resultant predictions provide a quantitative base for experimentally testable hypotheses about the recall probability of defined subjective content from past life periods. Such a comprehensive characterization is necessary to inform theoretical and computational models of AM (Rubin and Wenzel, 1996) and provides the groundwork to test how AM content and frequency change with age, pathology, and differing physiological conditions.

Methods

The CRAM (Cue-Recalled Autobiographical Memory) test

CRAM is comprised of four parts delivered using computerized interactive forms: (Part 1) Non-identifiable information, i.e., month and year of birth, gender, and whether English is a native language, is collected from the participant; (Part 2) Thirty word-cued AMs are uniquely labeled by the subject with brief text descriptions; (Part 3) The text descriptions of each memory are re-presented one by one for the participant to date the recalled event; (Part 4) The text descriptions of a subset of the dated memories are re-presented once again, one at a time, for the subject to score each AM for content by counting the number of items in the recollection in each of eight specified categories.

The scope and details of how CRAM cues and scores AMs are substantially different from previous methods (Addis et al., 2010; Crovitz and Schiffman, 1974; Fromholt et al., 2003; Hashtroudi et al., 1990; Johnson et al., 1997; Levine et al., 2002; Robinson, 1976). Thus, the full written instructions of these sections of CRAM (parts 2 and 4) are provided as Supplemental Information (SI: sections A and B-C, respectively). These scripts, which remained fixed throughout the study, were progressively developed with the aid of debriefing interviews during extensive preliminary experiments (not reported here).

Memory definition, cueing, and word sampling. In part 2, participants were provided with the following definition:

"Autobiographical memories are recollections of past episodes directly experienced by the subject. These memories should be of a brief, self-consistent episode of your life. An episode can be as short as a single snapshot and up to a few seconds long."

Limiting the duration of the recalled episode to a few seconds allows for segmentation of multiple recalled events (Ezzyat and Davachi, 2011), which facilitates scoring of just one episode rather than a combination of many recollections. Subjects were instructed to read through a set of 7 word cues, to identify the first AM that came to mind, and to label it with either a unique word or phrase. Any single (or group of) cue-word(s) from the set could be used to trigger an AM. If no memory was elicited, participants were able to call up a new set of words, until 30 AMs were successively generated and labeled for later recall. Two major differences distinguish CRAM from the commonly adopted standard cue-word method (Robinson, 1976). First, each cue consisted of a list of 7 words rather than individual words, a number determined by trial and error in early pilot experiments that enabled relatively quick and probable autobiographical recall. Second, the words were not selected from a small, fixed sample as in many previous studies (e.g., Crovitz and Schiffman, 1974; Jansari and Parkin, 1996;

Robinson, 1976; Rubin, 1982; Rybash and Monaghan, 1999). Rather, they were chosen randomly from the 100,000,000-word British National Corpus (see SI section D for processing details), a compilation of written and spoken works. Therefore, word-set sampling (see e.g., Fig. 1A) was based on natural usage frequency, and differed dynamically from subject to subject. The rationale for these choices was to achieve a sampling of AMs as close as possible to those occurring under normal circumstances, as elicited, for example, by everyday conversations, readings, or one's internal dialogue.

Memory dating. In part 3, the participant's age was used to divide his or her life span (e.g., Howes and Katz, 1992; McCormack, 1979) into 10 equal temporal periods or bins, numbered 0 through 9. We use the term youth in reference to the age of the subject at the time of the recalled episode relative to their age at the time of study participation. Three equivalent yet complementary methods were made available to participants to help them allocate each memory to one of the youth bins (Fig. 1B): subject's age (e.g., from 15 years and 1 month old to 18 years old), date of event (e.g., from May 1997 to April 2000), and time lapsed (e.g., from 15 years and 1 month ago to 12 years and 1 month ago). Participants were instructed to use (and encouraged to switch between) the method(s) that best helped them accurately date each memory. Given that accurate dating of select memories may be difficult, subjects were able to assign AMs to multiple bins, if needed. This option also addresses the possibility that temporal bins based on the subject's age might create cutoffs intersecting a temporal range associated with a particular episode. Multiple bins were selected for 2.8% of all dated AMs. During subsequent analyses, when applicable, AMs were weighted according to the number of bins to which they were assigned. The first 2 AMs were considered practice with the procedure and excluded from further processing.

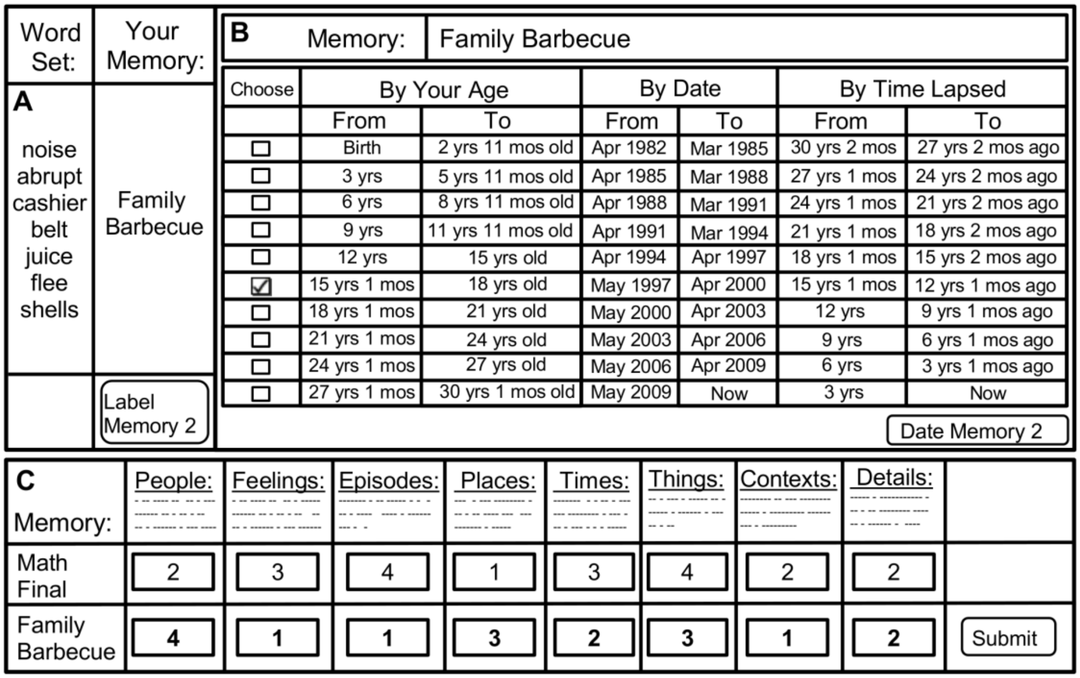


Figure 1. ***Graphical appearance of the Cue-Recalled Autobiographical Memory (CRAM) test user interface.***
Each panel represents a separate part of the CRAM test. (A) Subjects first label 30 memories each recalled upon presentation of 7 words stochastically sampled from their natural language usage frequency. (B) Each memory is then dated into one of 10 temporal bins, based on the subject's age at the time of the event, the date of the event, and/or the time lapsed from the event. (C) Finally, participants score the content of 10 memories by counting the number of elements recalled from the event for each of eight distinct features (People … Details). Every feature is accompanied by a brief definition, schematically illustrated here by a few dotted lines underneath (see section B of the Supplemental Information for their full text).

Count of elements within specified memory features. In part 4, a subset of the twenty-eight dated memories was pseudo-randomly selected for content scoring. Participants using an earlier version of the interface (implemented in Microsoft Excel) scored one AM from each reported bin. Participants using the later version (running in regular internet browsers) scored 10 AMs: one memory from each reported bin, plus if applicable (i.e., if not all bins were represented by a subject) additional AMs were selected starting in order from the least to most represented bins in the entire data base for all memories across subjects at that point in time. This latter procedure maximizes coverage of scored AMs across bins, while relaxing the constraint of uniform bin coverage for each subject. The labels of the sampled memories were re-presented one by one in random order. Participants were asked to count as many details as they could recall for each memory with the following instructions:

"In this part you will revisit your recorded memory. For this memory, your task is to count how many elements you remember. There are 8 categories of elements, each with a short description and example - click on the category's name to see the example. Once you have counted the elements of a given category, enter that numerical value in the proper box, and proceed to the next category. After completing all categories for a memory, press 'submit' to display the next memory. Click here for additional guidance on what constitutes 'an element'."

In this paper, we refer to details as elements, their categories as features, and the summed element counts for all features as total content. The eight specified features were *Contexts, Episodes*, *Feelings, People*, *Places*, *Things, Times*, and (other) *Details*. These CRAM features are similar to those investigated in the AM literature. For example, a meta-analysis of 84 articles on episodic memory (Spencer and Raz, 1995) compiled a categorical list of commonly characterized variables. These included descriptors of temporal sequences (*Episodes*), events (*Contexts*), temporal specificity (*Times*), perceptual features (*Details*), objects (*Things*), self (*Feelings*), persons (*People*), and spatial features (*Places*).

Although all features were displayed together, their order was randomized for each participant (and kept fixed for the ~10 scored AMs). The description of each feature was in the form of a question that remained visible throughout the scoring section. However, the subject had the option to collapse or expand back the descriptions at any point after the first scored memory. An example of how elements should be counted for each feature was also offered through a clickable link. The descriptions and clickable examples are provided for each of the features in section B of the SI. Participants were offered the option of additional guidance on whether an element should be counted within a particular feature by means of a hyperlink (which remained available throughout the scoring section) to a detailed explanation, reported in section C of the SI. According to the test logs, individual feature-counting examples and general additional guidance were invoked on average 1.10 and 0.12 times per participant, respectively. All instructions and additional guidance were written to minimize biases; in particular, the content of the examples focused on the definition of the respective feature, avoiding references to specific life periods (except for the *Times* feature), and traumatic or important episodes.

Graphical user interface, implementation, and availability of CRAM. All components of the CRAM test, including the word sampling algorithm, graphical user interface (GUI), and the response-driven transitions within and between the four parts, were developed and deployed in two separate formats and environments. One was based on Microsoft Excel and implemented in Visual Basic, while the other was based on standard internet browser protocols (HTML) and implemented in PHP/Java script (Fig. 1). Each of these two versions was complete, independent, and fully functional. Although the "touch and feel" of the two GUIs was different, the exact wording and sequential order of the functions were identical.

All results described in this paper are derived from a procedure in which subjects took the test on a local computer in the lab with one of the investigators present in the room. A version of the CRAM test for internet browser was later adapted, and is currently available for online use (http://cramtest.info). A version of the Excel implementation has also been developed in Italian, with faithful translation, based on an established 500,000 word spoken Italian corpus (http://badip.uni-graz.at). All versions of CRAM are available from the corresponding author upon request.

Participants, data screening, and analysis. The subject pool consisted of George Mason University undergraduate students recruited through the Psychology Department's enrollment web site. Students received course credit for successful study completion. This research was approved by the George Mason University Human Subject Review Board in accordance with Federal regulations and Mason policies for the protection of human subjects. Written consent for participation was obtained prior to data collection. All reported data are from subjects 18 through 36 years of age (mean ± standard deviation: 21.17 ± 3.77; median: 20). Memory dating involved 111 participants (83 females, 28 males; 72% native speakers) using the Excel format and an additional 83 participants (63 females, 20 males; 74% native speakers) using the web browser format. Out of the 5,432 dated memories, 1,424 memories were scored for content by 103 participants (77 females, 26 males; 75% native speakers) using the Excel format plus 79 participants (61 females, 18 males; 72% native speakers) using the web browser format.

All data were stored in a relational database (MySQL 5) and queried for quantitative measurements in SQL language. The extracted parameters were imported into R (Dalgaard, 2008), SPSS, and Excel for statistical analysis and graphical output. Multiple tests of probability were corrected using the Bonferroni method. Data were initially inspected by the investigators to ensure the reasonable authenticity of the responses and to minimize the impact of intentional hoax, lazy entries and honest typos. Representative screening examples are reported in section E of the SI. This process resulted in the exclusion of data from 5 subjects plus 54 individual memories.

The SPAM (Spontaneous Probability of Autobiographical Memories) protocol

The SPAM protocol was devised to estimate the “Spontaneous Probability of Autobiographical Memories” by measuring the probability and duration of naturally occurring AMs during the course of everyday life using experiencing sampling (Brewer, 1988). The SPAM procedure randomly prompts participants at specific instants during the day to note whether in those very moments they are recalling an AM. When they are in fact experiencing an AM, they are asked to estimate the length of time of their reminiscence up to the point of the prompt. From these data, the fraction (f) of random prompts that correspond to an AM event is obtained by dividing the AM-associated prompts by the total number. Moreover, the duration (d) of an AM is computed by doubling the time estimate of the reminiscence, because on average the prompt interrupts the middle of the AM. The number (N_t) of memories that are spontaneously recalled in a given period (t) can be estimated as $N_t = f \cdot t / d$. In this formula, f represents the probability of experiencing an AM at any one moment in time, and t / d corresponds to the total number of possible AMs experienced in a given period of time. Their product ($f \cdot t / d$) captures the number of memories in a temporal window, given a participant-specific probability of occurrence and average duration.

The protocol design capitalizes on the widespread technology of mobile telephony. An auto-dialer program was custom written for a computer modem to randomly call participants within variable constraints. Participants were given a choice of the number of daily calls they would receive and the hours to exclude from calls for sleep or other reasons. As the goal of SPAM is to measure the frequency of typically occurring AMs independent of retrieval mechanism, no distinction was made between voluntary and involuntary memory.

Upon initial briefing, SPAM participants were given a packet containing the informed consent form, a concise description of the protocol, a log booklet, and a form to record general biographical information and calling parameters (e.g.,, number of calls allowed per day). The packet also included the definition of AM as well as examples of mental states that should or should not count as AMs. The text of these examples is reported in full in section F of the SI. One of the investigators verbally reviewed the contents of the packet with all participants. When subjects received a call, they were instructed to perform a mental check on whether they had been experiencing an AM at that very moment and to write on the log book their best estimate of the memory duration up to the point of the phone call. Otherwise, a dashed line was used to indicate the absence of an AM event at the time of call. Subjects were encouraged to program a specific ring-tone for the number used by the auto-dialer to allow for a more instant reaction to the prompt; alternatively, SPAM calls were identified by caller ID.

Participants, data screening, and analysis. The subject pool consisted of George Mason University undergraduate students recruited through the Psychology Department's enrollment web site. The pool of subjects was the same as used for CRAM; however, the participant samples were entirely non-overlapping. Students received course credit for successful study completion. The protocol was approved by the George Mason University Human Subject Review Board in accordance with Federal regulations and Mason policies for the protection of human subjects. Written consent for participation was obtained prior to data collection. A total of 53 subjects underwent testing, and 16,801 phone calls were made altogether. On average, subjects received 17 calls per day (range 8-22) and selected daily calling windows of 11 (range 5-18) hours. The mean number of calls received by each participant over the course of the entire experiment was 317, with a standard deviation of 58 (range 184-480). The exact numbers depended on individual choices of parameter settings and an additional random factor due to the stochastic nature of the calling algorithm. For a typical subject, the SPAM experiment lasted an average of

19 (range 14-34) days. All reported data are from subjects (29 females, 24 males; 77% native speakers) between 18 and 37 years of age (mean ± standard deviation: 22.25 ± 3.85; median: 21). Data were collected, entered in Excel for analyses, and excluded if values were more than 3 standard deviations from the mean, resulting in the exclusion of one data point pertaining to memory rate per hour (N_{hour}).

Results

Memory content is more resilient to temporal decay than retrieval probability

The temporal distribution of memories collected with CRAM is shown in Figure 2. More than 50% of AMs referred to episodes that occurred in the most recent 20% of the subject's life. This result quantitatively and qualitatively reproduces previous seminal findings using similar protocols (Fig. 2A). The reminiscence bump in the data of Jansari and Parkin (1996), and its absence in ours and those of Rubin and Schulkind (1997a), are consistent with the participant age pools. Specifically, to discriminate the bump from the retention function, AMs must be sampled from time periods between these two phenomena. As the age of our sample falls within the constraints of the reminiscence bump, it is expected to be occluded by the retention function (Jansari and Parkin, 1996; Janssen, Rubin and St Jacques, 2011).

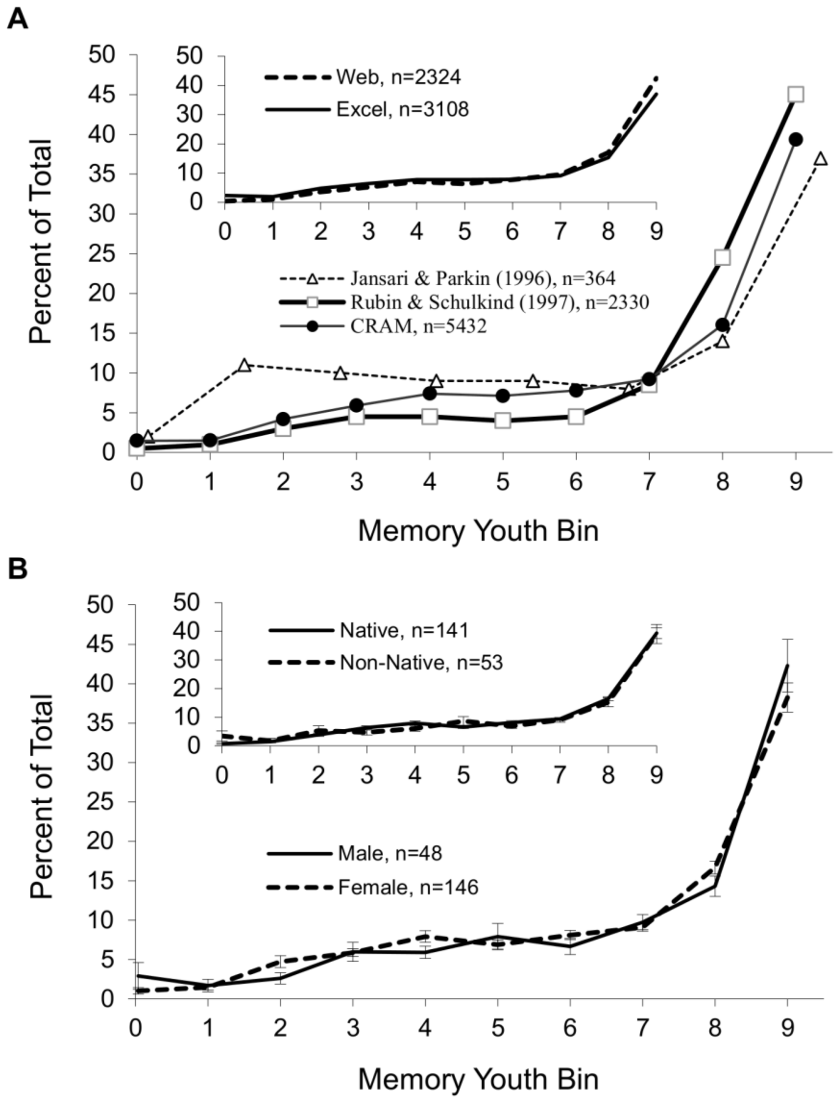


Figure 2. ***Temporal distributions of memories collected with the CRAM test.***
(A) Consistent with previous findings, the CRAM temporal distribution of autobiographical memories shows a retention effect in the most recent time bins and a power decay toward remote bins. Data re-plotted (with permission) from previous studies, i.e., Jansari and Parkin (1996), and Rubin and Schulkind (1997a), are adapted by converting memory age to youth. The term youth reflects the age of the subject at the time of the recalled episode relative to their age at the time of study participation, and is grouped into 10 bins from the most remote (0) to the most recent (9) episodes. Temporal distributions are also equivalent between the Excel and web-based CRAM test formats (inset). (B) Temporal distributions, plotted as mean ± standard error across individual subjects to enable statistical comparison, are not significantly different between genders and between native and non-native English speakers (inset).

Temporal distributions between the two CRAM interfaces (Excel- and web browser-based) were nearly identical (Fig. 2A inset). Also in agreement with earlier literature, the same distribution was observed for males and females (Fig. 2B), and for native and non-native English speakers (Fig. 2B inset). Overall, our analysis confirms the robustness of the temporal distribution of AMs to different experimental procedures.

The average total content of AMs was the sum of elements from every feature computed as a weighed mean over all bins. The weight of each memory was its retrieval probability according to the temporal distribution of AMs, thus correcting for the sampling procedure that selected AMs for scoring equally between bins. For example, more recent AMs had greater weights based on the relatively large proportion of cued AMs occurring in more recent life periods. Given this formulation, a typical AM contained ~20 elements. Although total content varied across individual memories (coefficient of variation ~0.5), the average values were extremely similar between genders, graphical interfaces, and native/non-native English speakers (Table 1).

Table 1. ***Total content from all 1424 scored memories reported as the sum of elements from all features of each memory.***
No differences in total content are found between test formats, genders or native vs. non-native English speakers.

	Total	Test Format		Gender		English	
		Web	Excel	Female	Male	Native	Non-Native
n	1424	756	668	1095	329	1059	365
Mean	20.07	20.53	19.68	20.18	19.68	20.41	18.94
SD	9.82	10.17	9.40	9.84	9.76	9.76	9.94

The total content of AMs varied as a function of the age of the memory, with AMs recollecting recent episodes typically containing more elements than those retrieved from the more remote past (Fig. 3). However, the reduction of content with time appeared to be less compared to the reduction in the probability of memory recall. For example, a memory from the most recent past (bin 9) had only 40% more elements (~23 vs. ~17) than one from a middle period (e.g., bin 4). In contrast, the retrieval probability from the same examples (Fig. 2) was more than 400% greater (38.2% vs. 7.5%). Similarly, the ratio of total content from bins 8 and 1 (which for a 20 year old subject corresponds to episodes that occurred at ages 17 and 3, respectively) was 1.25, compared to more than a 10-fold factor in the retrieval ratio.

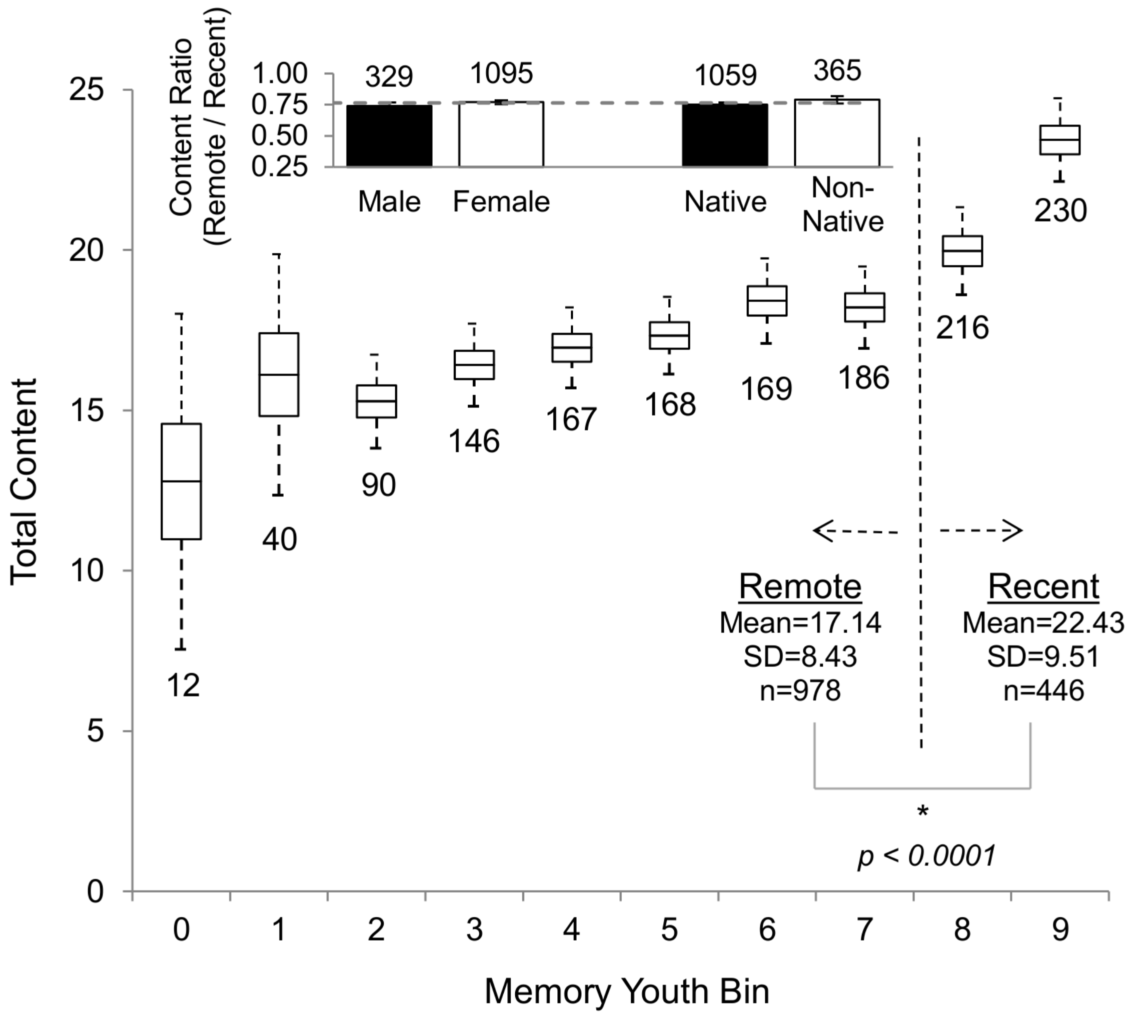


Figure 3. ***Total content across youth.***
Total content increases for memories retrieved from the first to the last tenth of life. Boxes represent the mean, 25th, and 75th percentiles, while whiskers indicate 95% confidence intervals. Scored memories are divided into remote (bins 0-7) and recent (bin 8 and 9) time intervals, which account for 45.4%, and 54.6% of all dated memories, respectively. The total content of recent memories is 31% greater than that of remote memories ($p < 0.0001$). The ratio between the two, i.e., the memory content lost to temporal degradation, is shown in the inset as the average (dotted line) and for two subject partitions as mean ± standard error.

AMs were operationally divided into "recent" and "remote" halves, corresponding to the last two and first eight bins, respectively. This approximates a median-split

(Bäuml, Hanslmayr, Pastötter and Klimesch, 2008; Gibbs and Rude, 2004) with recent and remote AMs accounting for ~55% and ~45% of all dated recollections, respectively. Recent memories contained an average of 5 more elements than remote AMs, $t(1422) = 10.54$, $p < 0.0001$, $r = 0.27$. This temporal effect was consistent across males, females, native, and non-native English speakers (Fig. 3 inset). To assure that this effect was not due to an arbitrary division of temporal periods, analyses were performed using various separations, e.g., between bins 8 and 9 or between bins 6 and 7, and provided equivalent results: $t(1422) = 8.88$, $p < 0.0001$, $r = 0.23$; and $t(1422) = 7.83$, $p < 0.0001$, $r = 0.20$, respectively.

Some features are more memorable than others in remote and recent memories alike

The overall content of AMs was analyzed in terms of the number of elements in each individual feature. This breakdown reveals two features that are particularly memorable, *Places* and *People*, each with more than 3 elements counted in the average memory. In contrast, *Contexts* and *Episodes* have fewer than 2 elements per memory each. The other four features have a number of elements per memory that remains close to the overall average of 2.5 (Fig. 4A). Interestingly, although the total content decays with time, the relative feature composition of AMs remains considerably stable from the most remote to the most recent memories (Fig. 4B); an equivalent pattern of feature composition emerges upon individual bin analysis. In particular, the relative ranking of the eight features remains largely unaltered, with *Places* being the most and *Episodes* the least represented features in both recent and remote memories. Together, *Places*, *People*, and *Things* amount to approximately half of the counted elements in recent and remote memories alike. In general, the ratio of the number of elements between remote and recent AMs for each and every feature remained close to that of overall content.

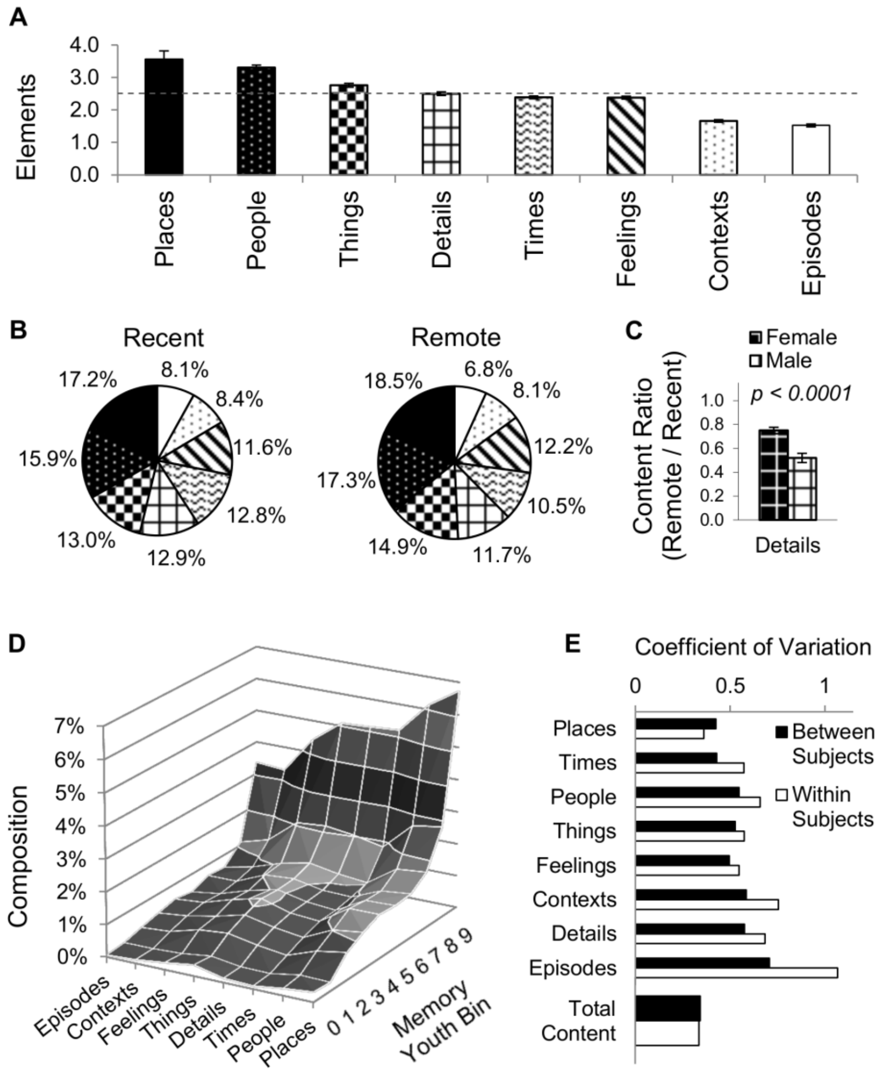


Figure 4. ***Feature composition of memory content.***
(A) The count of elements in each of the eight features is reported as mean ± standard error over all scored memories. More elements pertaining to Places and People are recalled than those pertaining to Contexts and Episodes. The dashed line represents the average number of elements (2.5). (B) The relative feature composition of recent and remote memories is similar (the slices are ordered by the recent rank). Although Feelings shifts from 4th to 6th rank from the remote to the recent distribution, the actual value change is very modest. (C) The content ratio (remote/recent) across all features was similar between genders with the only exception of Details, for which the decay was significantly greater in males (light hatching) than in females (dark hatching). (D) Recall composition of elements and youth in a typical AM. (E) For all features except one, the variability in the number of elements within subjects is greater than that between subjects. However, variability of total content is similar within and between subjects.

The overall feature composition of AMs was found to be remarkably similar between males and females. Although females remembered slightly more *Feelings* (2.5±1.7 vs. 2.1±1.5) and fewer *Details* (2.4±2.1 vs. 2.7±2.6) than males, these differences were not statistically significant, and the number of elements for all other features differed by less than 10% between genders. Similarly, the temporal decay, reflected by the remote/recent ratio was similar between female and male subjects across all features with the one noticeable exception of *Details* (Fig. 4C). Specifically, relative to the overall content decay of 0.76 (Fig. 3 inset) females tended to retain significantly more *Details* than males (ratios of 0.75±0.90 vs. 0.52±0.72, $t(1422) = 4.25$, $p < 0.0001$, $r = 0.07$).

Combining the temporal and feature distributions, it is possible to compute the typical composition of recalled memories (Fig. 4D). Elements of *Places* from the most recent tenth of one's life are over 100 times more represented in AMs than elements of *Episodes* from the most remote tenth (6.7% vs. 0.06%). *Feelings* from bin 7 are approximately as likely to be recalled as *Things* from bin 4 (~1%). Interestingly, for most features, content varied more among different memories of individual subjects, than across subjects. In contrast, the variability of total content was essentially identical within and between subjects (Fig. 4E). Altogether, these findings suggest that individual memories can vary substantially in their feature composition (e.g., one retrieved memory might have richer information on *Times* than *Contexts*, and another just the opposite), yet these effects tend to average out when considering all combined content and/or a large pool of memories.

A certain amount of correlation among the number of elements in the various features is expected, as richer memories are likely to have more elements in several features. However, some features may be more "independent" than others. In order to identify these more "fundamental" features, we computed the cross-correlation among features, as well as the correlation of each feature with all other content (Table 2). Interestingly, *People* and *Places*, the most memorable features, are also the least correlated with the rest of AM content, but *Episodes* (the least memorable) ties for second in this ranking, while *Feelings*, *Details*, and *Times* (the three features with average memorability) are last. These correlation values were essentially identical for males and females (data not shown).

Table 2. *Feature cross-correlation.*

The numbers of elements in each feature are all positively correlated with each other across memories, as quantified by Pearson's correlation coefficients. [a]The last row reports the correlation of each feature with the sum of the elements from all other features combined. Elements pertaining to People, Places, and Episodes are more independent, as indicated by lower correlation coefficients, while elements pertaining to Times, Details, and Feelings are more interdependent. The mean Pearson Coefficient across features with All Other Features is 0.46.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1. People	-							
2. Places	0.17	-						
3. Episodes	0.22	0.16	-					
4. Things	0.21	0.37	0.21	-				
5. Contexts	0.18	0.24	0.31	0.26	-			
6. Feelings	0.24	0.29	0.29	0.39	0.39	-		
7. Details	0.26	0.26	0.38	0.30	0.39	0.36	-	
8. Times	0.20	0.44	0.38	0.28	0.38	0.31	0.34	-
All Other Features[a]	**0.33**	**0.43**	**0.43**	**0.46**	**0.48**	**0.52**	**0.52**	**0.52**

Reminiscence of AMs occupies a substantial fraction of cognitive time

The SPAM protocol measured the rate at which AMs are recalled under normal conditions. On average, one in seven subjects reported reminiscing an AM at the time of a phone call. This corresponds to a mean fraction of time spent reminiscing of ~15% (median: 14%, mode: 15%). However, this sampling probability varied considerably among individuals, ranging from less than 2% of the calls for some subjects to more than 40% for others (Fig. 5A).

The subjective estimation of the duration of AMs varied considerably both within and between subjects. In particular, the length of memory recall, averaged in each individual over the "positive" cases in which a phone call interrupted reminiscence, ranged from less than 5 to more than 60 seconds, with a grand mean around half a minute (median: 28s, mode: 28s). The standard deviation of this mean across individuals was 13.5 seconds. There was no correlation between the mean and the coefficient of variation of memory duration from subject to subject (Fig. 5B).

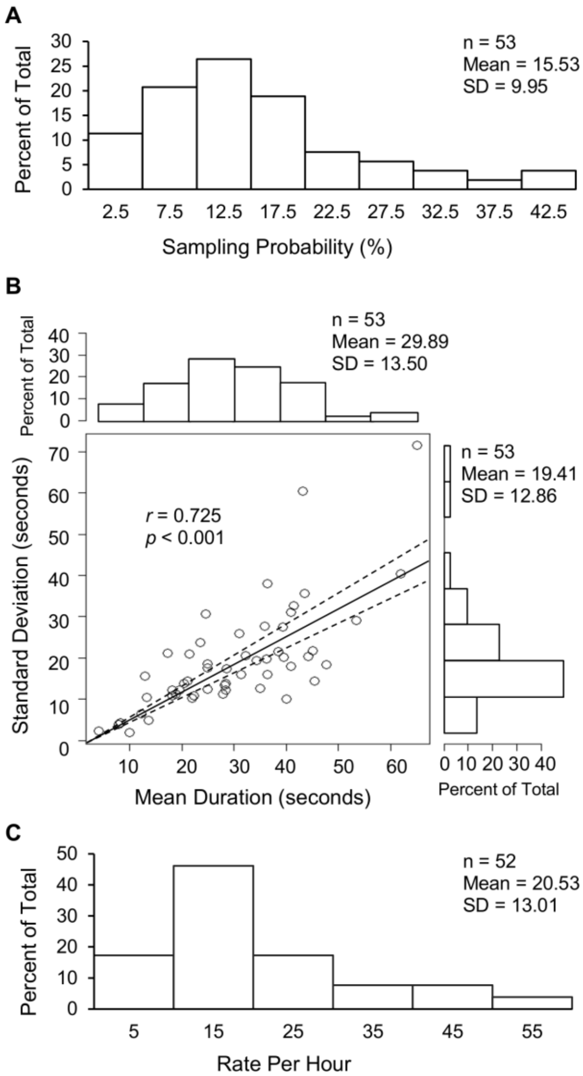


Figure 5. ***Frequency and duration of autobiographical memories.***
(A) The Spontaneous Probability of Autobiographical Memories (SPAM) was sampled on a participant-by-participant basis by dividing the number of memories reported by the total phone call prompts. (B) The mean and standard deviation of AM duration was computed for each subject over an average number of memories per subject of 48.5 (the standard deviation of this number was 29.3, range 3-133). Mean durations are directly and significantly correlated with the standard deviation. The slope of the best fitting line indicates a coefficient of variation of 0.65 with a 95% confidence interval of ±0.07 (dashed lines). (C) The number of AMs experienced per hour was calculated for every subject by multiplying his/her sampling probability by 3600 and dividing the result by the same individual's mean duration in seconds.

The probability of recall and memory duration data were used to compute the number of AMs retrieved in an hour. Such estimation yielded a right-tail skewed distribution, with an average of 20.5 AMs recalled per hour (median: 17, mode: 11). Except for one outlier, the range across individuals was 2 to 54 memories per hour. None of these metrics (i.e., probability, duration, and rate per hour) varied significantly between males and females or native and non-native English speakers (Table 3). Moreover, no recall differences were found within (i.e., early and late) or across (i.e., initial and subsequent, weekday and weekend) days of sampling (data not shown).

Table 3. ***Quantification of the spontaneous occurrence of autobiographical memories.*** *The probability of occurrence, duration, and hourly rate are not significantly different between males and females or between native and non-native English speakers.*

		Gender		English	
		Female	**Male**	**Native**	**Non-Native**
	n	**29**	**24**	**41**	**12**
Sampling Probability (%)	Mean	15.35	15.74	14.03	20.65
	SD	10.22	9.83	8.70	12.48
Memory Duration (s)	Mean	28.34	31.75	28.34	35.15
	SD	12.89	14.25	13.97	10.61
Rate Per Hour	Mean	21.10	19.88	20.12	21.92
	SD	14.19	11.76	13.47	11.77

Quantitative estimates of AM content retrieval by feature and youth

Data from CRAM and SPAM were combined to estimate the quantitative profile of subjective content in naturally-occurring AMs. Such analysis assumes that the temporal distributions and feature content assessed with CRAM on word-cued memories is sufficiently similar to that expected of AMs recalled in everyday life. While this vital assumption remains to be tested, such integration yields a useful baseline of quantitative hypotheses about probability of feature recollections from distinct life periods. In particular, the measured temporal distribution of cued memories, together with the spontaneous retrieval rate, allows the computation of the average period that elapses between recalls of AMs from a given bin (Fig. 6A). According to this analysis, a typical subject recalls a memory from the first fifth of one's life every ~3 waking hours or ~5 times a day. In contrast, only minutes separate consecutive retrievals of AMs from the most recent tenth of one's life.

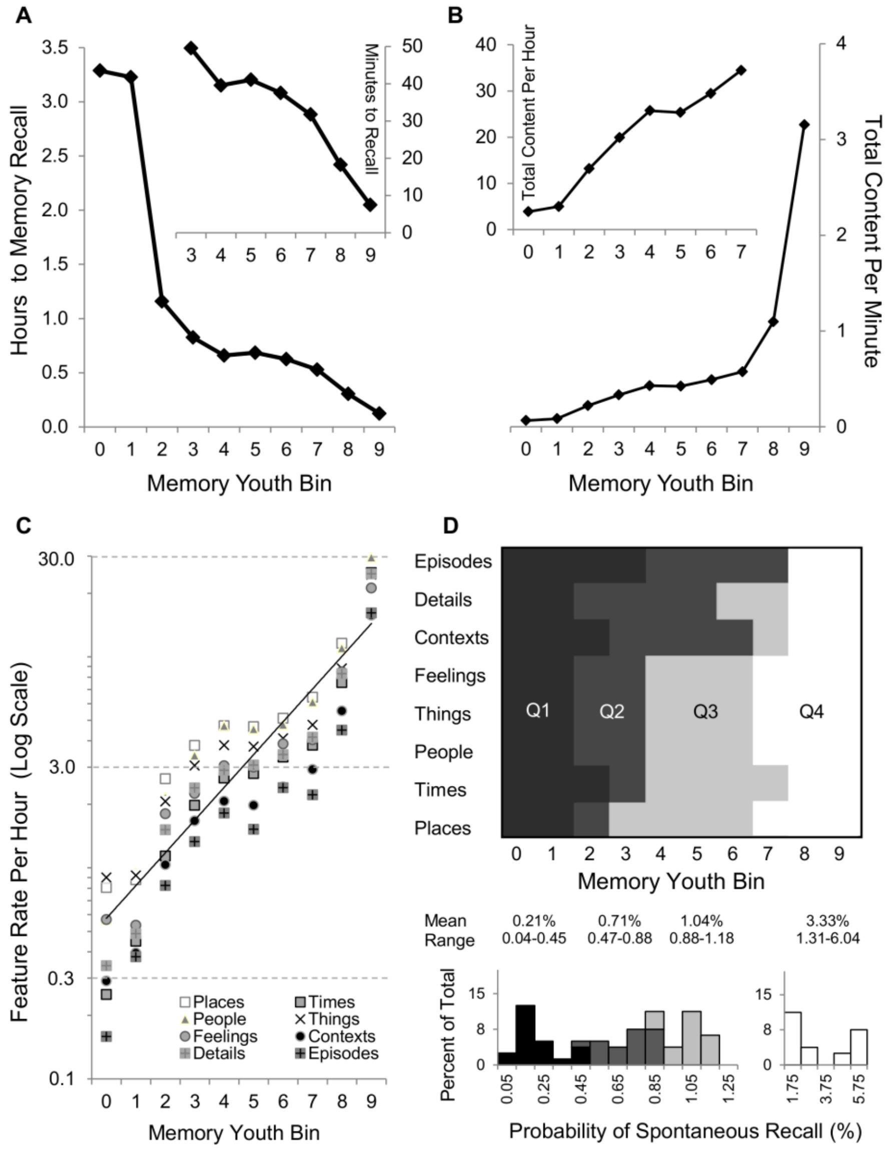


Figure 6. ***Retrieval probability of memories and features across youth.***
Integration of CRAM and SPAM data allows a number of quantitative estimates. (A) The average time elapsing between two memory recalls (e.g.,, more than 3 hours for the most remote AMs, but less than 10 minutes, see inset, for the most recent AMs). (B) The average recall rate of total content (e.g.,, more than 3 total elements per minute from the most recent tenth of life, but less than 3 per hour from the most remote tenth of life, see inset). (C) The number of elements for each feature recalled in one hour from each life period; the linear fit in the log scale represents the average across features. (D) Momentary probability of recalling at least one element of a particular feature from a distinct bin, grouped in quartiles.

In the same vein, total content retrieval can be computed as the number of elements recalled from various bins per unit of time. A total of 412 elements are recalled per hour (7 per minute, or one every 9 seconds on average); 62% of this amount is from the recent past, i.e., from when a 20 year old subject was 17 or older (Fig. 6B). This analysis can be further broken down by individual features (Fig. 6C). For example, while 3 contextual elements are recalled in an average hour from bin 7, the same rate of 3 features per hour holds for the number of recalled places from the more remote bin 2. Moreover, one may estimate the probability of recall, at any given time, of at least one element of a particular feature from a specific period of his/her life (Fig. 6D). For example, these results indicate that in any given moment, approximately one in 167 awake people (0.60%) are experiencing an AM containing an object (*Things*) from their third tenth of life, while ten times as many individuals (6%) are recalling a place from their most recent tenth of life. More generally, these data may be summarized by dividing all 80 combinations of 8 features and 10 bins in quartiles based on the relative probability that at least an element of the corresponding pair would be recalled at any one time. The first (i.e., least probable) 3 quartiles are uniformly distributed in decreasing order of the age of the memory, as indicated by the respective probability means and ranges (Fig. 6D, bottom histograms). In contrast, the fourth (most probable) quartile introduces a discontinuity (from the 0-1% to the 2-5% range) displaying a bimodal distribution, clearly reflecting the retention effect.

Discussion

This work begins to answer open yet integral questions in the scientific characterization of AM: For typical everyday AMs, what features are recalled and how many? From when are the recalled events, and how often are they remembered? In order to tackle these questions, we introduced two novel tools, the CRAM (Cue-Recalled Autobiographical Memory) test and SPAM (Spontaneous Probability of Autobiographical Memories) protocol. Compared to commonly used methods, these procedures permit relatively efficient data collection and comprehensive analysis of AM content and occurrence. The results reported here complement and expand existing knowledge on the temporal distribution, subjective content, and frequency of AMs. This is accomplished by quantifying the number of elements retrieved in specified features from naturalistically-cued AMs of numerous life periods in addition to the spontaneous retrieval probability and overall occurrence of AMs during everyday life.

The seminal cue-word research design of Crovitz and Schiffman (1974) opened a path to quantify the effect of AM age on its relative frequency. Several studies adopted this method in the popular variations of Robinson (1976) and Rubin (1982), which fixed the word set to maximize reproducibility. Specified qualities of AMs were also measured with rating scales (e.g., MCQ; Johnson et al., 1988), and in relatively few cases those aspects of content were adopted to count the number of details recalled (e.g., Hashtroudi et al., 1990; Levine et al., 2002).

CRAM combines an adaptation of the cue-word technique with a variation on feature-counting assessments by asking subjects to date retrieved episodes and to count the identifiable elements of their own AMs in each of several features. In order to emulate everyday cuing instances, we chose to sample cue words stochastically based on their usage frequency in natural language. Moreover, to further mimic naturalistic conditions of memory retrieval, we prompted the subject with a list of 7 words instead of just one. Consistent with previous reports (e.g., Jansari and Parkin, 1996; Rubin and Schulkind, 1997a, 1997b) , we observed a retention effect for the most recent AMs, a power decay for intermediate AMs, and childhood amnesia for the most remote AMs. This temporal distribution of retrieved episodes was also robust with respect to subject gender, native language, and details of the computer interface.

To the best of our knowledge, this is the first report of the absolute number of elements retrieved in AMs using cues based on their natural usage frequencies. Both the mean value (~20) and the coefficient of variation (~0.5) were equivalent between genders, native language, and computer interface. Moreover, total content values collected with CRAM are similar to those reported when cuing memories with typical life events and scoring event narratives (Levine et al., 2002). Those same previous studies showed that experimenter probing of specified feature categories may aid retrieval of content. Nevertheless, the amount of total content collected through CRAM closely resembles that obtained without such specific probes (Levine et al., 2002), suggesting the counted details are more akin to those recalled spontaneously. This apparent conflict may be reconciled by highlighting methodological differences. CRAM provides description and examples of what constitutes an element in general feature categories, which differs from overt requests using an extensive number of item (or element) categories (as done in earlier approaches).

Adding to previous studies, the combination of a naturalistic cue-word technique with content counts enabled analysis of how the content of everyday occurring AMs varies by the age of the memory. Total content was found to decay temporally, but not as prominently as observed in the temporal distribution of AMs. For example, AMs from the most recent two tenths of one's life are nearly five times more likely to be retrieved than AMs from the most remote eight tenths, but their total content is only 30% greater. Further breakdown of AM content composition revealed certain features (*Places* and *People*) that are generally more memorable than all others in both remote and recent memories. *People* and *Places* were also determined to be the most "independent" features, i.e., those least correlated with other features. Moreover, these two features had a relatively high ratio of elements recalled in remote vs. recent AMs, indicating resilience to temporal degradation. These results complement previous observations that memory cues pertaining to the "what" of the event produce the greatest amount of recalled details (Catal and Fitzgerald, 2004). If AMs are supported by an underlying skeleton of core features, these data suggest *Places* and *People* as likely candidates for such a core. Barsalou (1988) reported that cues for people and locations elicit AMs more quickly than other cues (e.g., time), possibly due to the underlying organization of AM. Similarly, our findings may relate to the notion that the number of details recalled in distinct features corresponds to how an AM is organized for retrieval (e.g., by location, person).

The variability of total content from memory to memory within a participant was similar to that of the content average over memories from subject to subject. In contrast, the variability of the content of individual features was greater among memories of a single subject than among all subjects. This finding reflects the intuitive expectation that different memories could have different composition, emphasis, and themes, while maintaining a consistent amount of retrievable information. Thus, one feature (e.g., *People*) might be richly represented in some episodes of one's autobiographic life yet absent in others. This expectation is consistent with other empirical reports. For example, in assessing the effects of emotional arousal on AM composition, Berntsen (2002) found a greater proportion of central details recalled from shocking compared to happy AMs, but an equivalent total number of details.

Similar to the temporal distribution, total content, feature composition, overall content decay, and feature cross-correlation were all similar for males and females (as well as for native vs. non-native English speakers and between computer interfaces). The only aspect that discriminated between genders was the difference in temporal decay of the number of other *Details* between recent and remote memories, which was more acute in males than in females.

As the CRAM techniques differ in certain aspects from previous studies, it is appropriate to discuss those design choices and their implications. Numerous definitions of AM have been employed in previous research. While temporal and spatial specificity are commonly used criteria (e.g., Levine et al., 2002; Rubin and Schulkind, 1997a, 1997b), the rules used to segment recalled episodes vary. To facilitate counts of AM content, the definition of AM provided here constrained the recalled episode to a few seconds. As the temporal distributions and content measures replicate previous research with different segmenting rules (e.g., Levine et al., 2002), on average this constraint does not seem to bias the recollections elicited along those dimensions.

CRAM utilizes sets of word cues that replicate natural language frequency to elicit recollections more comparable to those occurring in real life. As word cues were sampled randomly from the British National Corpus, the use of an international subject pool dictated monitoring of differences in AM recall between subjects of varying native languages. Given the absence of differences between native English and non-native English speakers and the lack of complaints or questions about particular cue words by any participant throughout the study, we surmise that CRAM in its current form is suitable for international users. As non-linguistic cues (olfactory, auditory, kinesthetic, pictorial, etc.) may also evoke AMs, the precise degree to which these word-sets mimic natural cuing experiences remains an open question. Since AMs elicited in the lab by varying sensory experiences have equivalent temporal distributions and vividness ratings (Rubin et al., 1984), stricter use of naturalistic cues might not affect measures of AM content across the life span. Nonetheless, further investigation into the proportion of AMs elicited by various sensory experiences and the resulting qualities and quantities of recollection is warranted.

The CRAM protocol does not reveal which word or collection of words elicited retrieval. This drawback impedes cross-sectional comparisons within and between participants as in experimentally-controlled fixed-cue recollections. This limitation reflects the unavoidable tradeoff between lab conditions and naturalistic approaches. In principle, cross-sectional comparisons focusing on naturally-occurring AMs are possible by statistical analysis of very large sample sets.

CRAM collects counts of details within feature categories instead of extracting details from participant narratives. While it is unknown how participants score content in CRAM (e.g., by enumeration or approximation), a smooth count distribution was found for each feature (data not shown), suggesting that the results are not significantly distorted by rounding bias (Huttenlocher, Hedges and Bradburn, 1990). Furthermore, during pilot experiments to determine the effectiveness of instructions, test subjects were asked post-test to provide a description of the details counted for each scored AM. In each case, this detail-by-detail event description corroborated the number of details provided. While such examples do not explicitly identify a scoring strategy, they are consistent with the idea that counts collected through CRAM are representative of the subjective details comprising AMs.

SPAM revealed that subjects spent a substantial fraction of their day reminiscing AMs. In particular, when unexpectedly asked whether they were experiencing an AM, on average participants had a 15% probability of “being caught in the act.” Combined with the assessed duration of ~30 seconds for a typical recollection, this result leads to the estimate of a mean recall rate of ~20 AMs per hour. Additional investigations will be necessary to determine what proportion of the considerable inter-subject variability of these values (range: 2-54 AMs/hour) reflects genuine cognitive diversity and how much is due to experimental error or systematic bias. In particular, each subject had the prerogative to select the temporal windows to receive calls. This was necessary to avoid interruption of sleep, privacy, or professional activities such as class attendance. If the times people are willing to entertain unexpected phone calls are also well suited for reminiscence, this protocol would tend to overestimate the occurrence of AMs. Moreover, although participant instruction was delivered systematically, differential interpretation could underlie the resulting disparity among subjects.

The integration of CRAM and SPAM data enabled interesting estimates of the number of elements retrieved in a fixed time for each feature from a specific life period. Although such detailed inferences demonstrate the potential of these novel research approaches, future studies will have to verify the underlying assumptions.

We stress that this research design focuses on the **subjective** aspect of AMs. In particular, the exact meanings of the eight features, as well as that of autobiographical memory, are taken to consist of the subject's interpretation of the corresponding definitions and accompanying instructions. Thus, by construction of the research protocol, these empirical measurements reflect what a subject considers to be a *feeling* or a *context*, given the definitions and examples communicated in the briefing sessions. In particular, our data do not discriminate between "true" or "false" memories, because the analysis targets mental representation of autobiographic episodes rather than their historical occurrence in the material world (e.g., Johnson and Raye, 1981). Moreover, we are measuring the subjective recalled content, independent of the total content that could possibly be recalled from those past episodes. Similarly, reminiscence duration in SPAM consists of subjective time estimates (Morillon, Kell and Giraud, 2009). At the same time, both the selection of features and the specific wording adopted in the explanations and interactions with subjects were chosen throughout the course of extensive pilot studies on the basis of spontaneous suggestions from participants and debriefing interviews. In this sense, the analyzed features should in fact correspond to observable aspects of subjective experience.

This research utilized a college-aged subject pool, a common practice in AM research (Berntsen and Hall, 2004; Bluck et al., 2005; Crovitz and Schiffman, 1974; Ezzyat and Davachi, 2011; Johnson et al., 1988; Robinson, 1976; Rubin, 1982; Rubin et al., 1984; Rubin et al., 2003), particularly when exploring novel measures. As such, these data help provide the basis for comprehensive characterization of AM and quantitative testing of AM models (Rubin and Wenzel, 1996). With this aim, these methods are currently being employed with a larger and more diverse pool of subjects of many ages. The high correspondence in data collected between computer interfaces outlined here suggests that CRAM produces reproducible results. Ongoing collection of additional data will help determine the extent of variation due to different testing conditions (e.g., administered in person or online) and the applicability to participants of increasing age. The CRAM dating procedure assigns AMs to bins which vary according to the subject's age (also see Howes and Katz, 1992; McCormack, 1979). While this normalizes the difficulty associated with dating AMs of increasing age (Robinson, 1976), it creates discrepancies in the temporal ranges of two equivalent bins from younger and older subjects. To circumvent these discrepancies, when age is a central variable, analysis will require AM comparison both in terms of relative time periods (e.g., defined by participant-specific bins) and absolute time periods (e.g., defined by the age of the participant and event). Nevertheless, the precision associated with the absolute date of an AM will decrease with increasing participant age. Moreover, certain comparisons remain restricted, e.g., AM content from the most recent year of life from younger and older participants (Levine et al., 2002).

Given the benefits of CRAM and SPAM to quantify efficiently the content and frequency of naturalistically sampled AM, these methods may be of value to cognitive and clinical psychology. Normative data may be useful in assessing the effects of particular conditions (fatigue, stress, psychotropic substances, etc.) or genetic variants on AM recall, as well as monitoring the progression of memory disorders (e.g.,, Alzheimer's disease, Korsakoff's syndrome, and anterograde/retrograde amnesia) in individual patients. Attempting cross-sectional analyses in cognitively impaired populations, however, demands consideration of key factors, including task difficulty, interpretation of instruction, and comfort level with technology (Mitzner et al., 2010; Schaie, 1977).

Further application of CRAM and SPAM could prove useful in clarifying several open questions about human recollection. It has been reported that ratings of vividness and reliving are not related to a higher retrieval probability observed in the reminiscence bump (Janssen et al., 2011). As expected from the age range of the subject pool, our temporal distributions do not show a reminiscence bump. However, applying CRAM to older populations may help elucidate possible relationships between feature counts and retrieval probabilities. In addition, AM has been theorized to have specific functions (e.g., directive, self, and social; Pillemer, 1992), which were supported by empirical findings (Bluck et al., 2005). SPAM, by its current design, does not establish why sampled AMs are recalled. However, the protocol may be adapted to isolate the frequencies of functionally-distinct AMs, and explore how their everyday usage frequencies change with age. SPAM also holds promise to clarify the relationship between voluntary and involuntary recollection. As past voluntary recollection may prime involuntary recollection (Mace, 2005), a variation of SPAM could be devised to obtain the frequency correlation between these two forms of retrieval on a participant-by-participant basis. A similar variation could directly compare the frequencies and durations of recollecting past events and future intentions (retrospective and prospective memory, respectively).

The data reported here provide measures of the content and frequency of AMs over the life span of young adults. Moreover, with the tools introduced in this work, collection and storage of age-specific population statistics in a large-scale informatics database (http://cramtest.info) is underway using Internet sampling. Altogether, these advances constitute a further step towards the inclusion of subjective mental content in the realm of quantitative, reproducible science (Ascoli and Samsonovich, 2008; Brewer, 1986), enabling deeper and broader queries of human memory content.

Acknowledgements

We are grateful to Dr. Praveena Kudrimothi for technical assistance in the pilot projects, Drs. Pamela Greenwood and Linda Chrosniak for critical discussions and suggestions, and Dr. Patrick McKnight for constructive feedback on statistical analysis and an earlier version of this manuscript. We thank the Air Force Office of Scientific Research (AFOSR) Award No: FA9550-10-1-0385.

Supplemental Information

A. Instructions for CRAM part 2 (word-cuing of AMs)

In this part you will be prompted with a set of words. Your task is to recall the first autobiographical memory that the words bring to your mind. This means just the first memory that you think of, not the earliest memory. The episode may have occurred any time from birth to now. Your memory does not need to be associated in any way to the given set of words. It is also ok if the memory comes to your mind before you have finished reading all of the words in the set. This memory should be of a brief, self-consistent episode of your life. An episode can be as short as a single snapshot and up to a few seconds long.

After you experience a memory, write down either a word or a simple phrase in the text area provided below that will allow you to recall the specific memory later in the test, and press “Label Memory” to move on with the test. The notes you write are for your own use only; at the end of this test they will be deleted and not recorded anywhere. You should feel free to write any personal information. The purpose of these notes is just for you to recall this same memory later in the test. They should be as brief and informal as you like, so long as they positively identify the specific memory in your mind.

If the memory you think of refers to a typical and repeated episode that happened regularly or multiple times in your life, you can use it only if you can fixate on a specific individual event. If you can only recall the generic (repeated) event, look for another memory.

If after a few seconds you cannot retrieve any autobiographical memory after reading the set of words, leave the text space blank and click the “No memory comes to mind” button to produce a different set of words.

B. Instructions for CRAM part 4 (feature descriptions and examples)

People: How many uniquely identifiable persons (excluding yourself) do you remember in the episode? Example: You were at some party. Your best friend John was there, and so was his second wife, whose name you don’t recall. The host, Marc, was there, and some of his relatives, but you cannot remember which. *Count 3 elements* (John, his wife, and Marc).

Feelings: How many distinct subjective feelings (tastes, odors, temperature, emotions, etc.) do you recall in the episode? Example: It was the last day of school. You had a stomachache, the room smelled like fish and it was too warm. Still, you felt very happy. *Count 4 elements.*

Episodes: How many other episodes that immediately precede or follow this one can you recall? Example: John remembers the first homerun he hit. He remembers the instance the ball hit his bat and a thunderous crack rang out. Recalling the episodes that led up to and following this moment, John recalls taking a warm-up swing before entering the batter’s box. He also recalls focusing on the pitch right before the ball was thrown. After he hit the ball he recalls running around the bases after which his memory fades but he recalls later that his team went out for a celebratory pizza. Though going out for pizza is a memory, it is not counted because it is not sequential - there is a gap in time. *Count 3 elements.*

Places: How many spatial features do you remember of this episode: town, house or road, room or vehicle, your exact position, etc. Example: You recall chatting with a friend in her apartment in New York, but not whether in the living room or bedroom, nor whether sitting or standing. *Count 2 elements* (for the apartment and the town, even if knowing the apartment "automatically" specifies the town).

Things: How many uniquely identifiable objects do you remember (must have at least one detail such as texture, material, size, color, or else be out of context)? Example: If you were inside a bedroom, the window doesn't count as an object (since almost all bedrooms have one), unless you remember that it was open, or that it had pink curtains… Same with a bed, a closet, etc. If, on the other hand, you remember there were skates on the floor, an apple on the table, or something not usually found in the standard bedroom, then *you should count those objects*.

Times: How many temporal features do you remember of this episode: the exact year, month or season, day of the week, time of the day, etc. Example: You remember getting a speeding ticket while driving to church. You can't remember the exact year, nor time of the day, but you recall it was summer and Sunday. *Count 2 elements.*

Contexts: How many other explicit contextual details (weather, situations, events, etc.) do you remember? Example: You remember that on that same day, the Lakers' won the league, your grandma was at the hospital, and it was freezing cold outside... *Count 3 elements*.

Details: How many other particular details do you recall (words uttered or heard, facial expressions, actions, clothing...)? Example: It was your first date. When you arrived, she said: “late at your first date!?”, and you smiled. She had already ordered a drink. *Count 3 elements.*

C. Optional (hyperlinked) guidance on what constitutes “an element” for CRAM part 4 (feature count of AMs)

A detail could be practically defined as the minimum element of information you would include in a very extensive and exhaustive account of this episode in a hypothetical personal diary. As with a personal diary, you would not describe over and over objects or people you are very familiar with.

Suppose that the episode consisted of an argument you had with an old friend in your kitchen. Even if you can probably visualize in your recollection many details of the kitchen, such as the position of the refrigerator, the color of the walls, and whether you had a gas or electric stove, these are not really part of the specific episode. You would not describe them in your diary, because they would be implied by the fact that the episode occurred in your kitchen. Thus, you should not count these details in the test. Similarly, you should not count the fact that your friend had blond hair and blue eyes. However, if the argument degenerated and the friend broke a dish on your head, you should probably count that dish as an object even if you had seen it many times before in the kitchen.

If, on the other hand, you are describing a hotel room you spent one night at, then every uniquely identifiable detail you can remember should count. You can’t, however, consider “the room had a door, a bed, and a lamp” as three valid details, unless you remember something specific about them.

If in another recalled episode you changed a flat tire, and remember that you had to unscrew as many as 16 bolts, should you count 16 details? Not unless you remember something specific for each and every bolt. In your diary you would probably write that there were 16 bolts, and this single element of information should be counted as one detail. Similarly, if you recall a dinner with 12 people, but only specifically remember 3 of them, you should count three details under “people” and one under “contexts” (corresponding to the fact that there were 12 people). If you remember that one person at dinner was a lawyer, but you don’t remember his face nor any other detail about him, should he count as a person? You can count him in, or alternatively you could count the fact that one person in the group was a lawyer as an “other” detail (it would in any case count as one detail overall).

In general, there is no objectively “right” or “wrong” way to exactly count details in a remembered event. What matters most is what you consider a detail in your memory, and as such you are the ultimate decision maker. No need to agonize over the specific category of the element. The distinctions between various categories are often ‘soft’, and you can decide just based on your intuitive preference.

D. Processing of the British National Corpus to sample cue-words

The initial corpus (http://www.natcorp.ox.ac.uk) consists of adjectives, cardinal and ordinal numbers, proper and common nouns in all their forms, and non-auxiliary verbs from the "demographic" (i.e., conversational) file. Articles, conjunctions, and prepositions, such as "and", "the", and "of" were excluded from this master list. A set of 187 obscene or otherwise questionable terms was identified with a freeware word filter tool (http://www.discusware.com). Potentially offensive terms were removed from the word pool including (but not limited to) the original and all derivatives of the seven terms that the U.S. Supreme Court ruled cannot be used on television (FCC v. Pacifica Foundation, 438 U.S. 726, Decided July 3, 1978). Finally, words with a usage count (as reported in the British National Corpus) of 4 or less out of 10^8 were also excluded to avoid arcane terms. The remaining list consisted of 13,241 distinct terms with an average usage frequency of 90 times for every 10^8 words. Cue-words were sampled with a weight proportional to their usage frequency, with the additional constraints that repeated words, the plural and singular forms of the same regular nouns, and the first- and third-person forms of the same regular verbs could not be re-sampled within the same test. An example of a randomly sampled set of 7 words was: "noise, abrupt, cashier, belt, juice, flee, shells." The number of times that these words appear in the British National Corpus is 215, 11, 7, 109, 136, 6, and 10 times (out of 10^8), respectively.

E. Representative examples of data screening

One suspicious pattern was detected in the AM temporal distributions from one participant, who dated all memories into the first (most remote) time bin. These data were excluded from analysis. Unusual entries were noted in the analysis of memory content, and screened based on the variability within and between subjects. In particular, outliers were identified at the level of individual participants and of individual memories. As a real data example, one memory was found to have a feature count of 100 elements in the *People* category, far beyond what could be explained by the variance observed in the remaining memories scored by that participant, who recalled a median of 2 people per AM. The entry was marked as a possible mouse slip and the individual memory was excluded from analysis. In a second case, a participant systematically displayed repeated counts among the eight features in the majority of AMs (e.g., one memory had 3 elements in each and every feature, while another had zero elements throughout). All memories from this subject were excluded from analysis.

F. Examples of autobiographical memories included in the SPAM instruction packet

An autobiographical memory (AM) refers to an episode of your personal past; a memory of something that you have personally experienced in your lifetime. This memory could be of an event that occurred from the very moment you were born to the last second you just lived. The event in the memory is typically specific to a place and a time.

NO *(Not An) AM Example:* You remember a 3 hour long trip to your grandmother's house and thinking how much longer it felt with an annoying little sister to share the time with.

Why Not? This memory would not to be considered an AM because the memory is of an event that was 3 hours long.

Yes *(An) AM Example:* You are remembering that once on a 3 hour ride to your grandmother's house, you saw a cow for the first time. You can see the cow again through your mind's eyes and you remember what you felt when you saw the cow (curious or afraid for example) and the smell of manure, mmmh.

Why? This is an AM because the event that you are remembering is specific to a moment in your life. That is, this event only lasted a few minutes, and did not stretch out over 3 hours, days, weeks, etc. Also, re-experiencing the feeling and smell of that event is a sign of mental time travel, a hallmark of AMs.

Yes *(An) AM Example:* You are remembering the moment your dog ran through the door of your home after being lost for two days. You are re-living the feeling of happiness and relief that Murdock (your dog) is safe and licking your face like a giant lollipop. You can actually feel his wet tongue on your face, that's how much you're into this memory.

Why? This is an AM because you are obviously re-living some portions of this event (i.e., emotional and physical feelings). Also, this memory is of an event that happened personally to you.

NO *(Not An) AM Example:* You remember a story about a person who for two days hopelessly looked for her dog, when a neighbor finally brought it over to her. After about a week of giving her dog extra love and extra long walks to make up for lost time, she realized that this was not her dog.

Why Not? This is not an AM because it is not an event that happened to you, this is a memory of an event from someone else's life.

NO *(Not An) AM Example:* You remember that for the past week, every time you turn your head fast you get a shooting pain down your spine. You can still feel what this pain was like.

Why Not? This is not an AM because the event is repetitive; you have experienced this event several times in your life.

Yes *(An) AM Example:* You remember that for the past week, every time you turn your head fast you get a shooting pain down your spine. Specifically, you remember the first time that it happened, when someone called out your name and you jerked your head around to see who it was.

Why? Though this is a memory of an event that has happened several times over your life, you are remembering one specific instance that happened.

Your Turn

YES or NO? You remember that earlier today, when you were heading out for lunch you saw a deer that seemed so calm that you could have pet it. You feel nervous just remembering this situation.

YES or NO? You remember when you were a child, every 4th of July you went to your grandparents' house and they always had your favorite dish prepared for you. You remember the smell and the taste of it, maybe your mouth is watering.

YES or NO? You remember last year your car broke down and you swore to your car that you were going to sell it. You remember how frustrated you felt. You remember that you popped the hood of your car to find out that it was only a loose battery cable. You re-experience the relief you had when you found this out. You remember getting into your car and apologizing for yelling at it.

YES or NO? You remember that you went to Jefferson High school and that your locker combination was 23-44-02.

YES or NO? You remember a story that your parents always tell about you whenever you bring a date home. The story is about when you were 10 years old and got so scared that you wetted your pants, how embarrassing.

CHAPTER THREE: OLDER ADULTS REPORT MODERATELY MORE DETAILED AUTOBIOGRAPHICAL MEMORIES

Abstract

Autobiographical memory (AM) is an essential component of the human mind. Although the amount and types of detail (content) that comprise AMs constitute important dimensions of recall, age-related changes in memory content are not well characterized. Previously, we introduced the Cue-Recalled Autobiographical Memory test (CRAM; see http://cramtest.info), an instrument that quantifies AM content, and applied it to college-aged subjects. CRAM elicits AMs using naturalistic word-cues. Subsequently, subjects date each cued AM to a life period and count the number of remembered details from specified categories (features), e.g., temporal detail, spatial detail, persons, objects, and emotions. The current work applies CRAM to a broad range of individuals (18-78 years old) to quantify the effects of age on AM content. AM content was positively correlated with subject age: older compared to younger adults showed a ~16% increase in the number of reported details (~25 vs. ~21 in typical AMs). This age-related increase in memory content was similarly observed for remote and recent AMs, although content declined with the age of the event among all subjects. In general, the distribution of details across features was largely consistent among younger and older adults. However, certain types of details, i.e., those related to objects and sequences of events, contributed more to the age effect on content. Altogether, this work identifies a moderate age-related

feature-specific alteration in the way life events are remembered, among an otherwise stable retrieval profile. By combining AM life span retrieval probabilities with measures of content, we present previously inaccessible estimates of the relative likelihood that a specified amount of retrievable content is associated with a particular age range and life period.

Introduction

Autobiographical memory (AM) refers to the recollection of personally-experienced episodes specified in time. Despite its critical value among adults of all ages (Bluck et al., 2005; Bluck and Alea, 2011; Pillemer, 1992; Waters, 2014), quantitative characterization of the content that comprises AMs (i.e., the types and amounts of associated detail) is lacking, most notably across a range of ages representative of the adult population, and across the life span of a given individual.

Several assessments collect subjective ratings of AM content generally using ordinal scales. For example, the Memory Characteristics Questionnaire (Johnson et al., 1988), Autobiographical Memory Questionnaire (Rubin et al., 2008), and Memory Experiences Questionnaire (Sutin and Robins, 2007) rate the amount (or clarity) of sensory (e.g., visual, auditory), spatial, temporal, and emotional detail associated with a particular memory. Likewise, event specificity has also been rated using numerous approaches (e.g., Kopelman et al., 1989; Piolino et al., 2002; Sutin and Robins, 2007). Relatively few studies, however, have reported absolute counts of the number of details retrieved in AMs (Addis et al., 2008; Addis et al., 2010; Berntsen, 2002; Hashtroudi et al., 1990; Levine et al., 2002; St. Jacques and Levine, 2007). These studies typically use standard sets of cues (e.g., event-cues) to elicit memories and subsequently collect written or spoken narratives of recalled experiences. Experimenters process each narrative (e.g., segment unique AMs), and score each memory for content across several categories of detail. The procedure, being relatively time-intensive, introduces a barrier to extensive AM content analysis; as such, data collection has generally been confined to memories from few and restricted life periods and age ranges.

We introduced the Cue-Recalled Autobiographical Memory test (CRAM; see Gardner et al., 2012) to address these limitations. CRAM elicits AMs using a modification of the word-cue technique (Crovitz and Schiffman, 1974). In contrast to traditional methods, word-cues are generated based on their usage frequency in spoken and written language in order to emulate naturalistic cues. Therefore, elicited AMs should be more closely matched to those recalled in everyday situations. Participants subsequently identify the age of each AM, and count the number of details recalled from specified features (e.g., temporal detail, spatial detail, persons, objects, emotions, temporally linked events, and other contextual elements) similar to those used in previous designs (e.g., Hashtroudi et al., 1990; Johnson et al., 1988; Levine et al., 2002). Given CRAM's reliance on participants to specify what constitutes a detail within a feature category, this technique permits efficient data collection thus enabling collection of larger data sets.

Despite these methodological differences, CRAM reliably reproduces several results of prior studies. For example, AMs cued by CRAM produce temporal distributions which completely replicate characteristics of those produced by traditional techniques, e.g., the retention interval and childhood amnesia (Rubin, 1982; Rubin, 2000; also see Janssen et al., 2011; Rubin et al., 1986; Rubin and Schulkind, 1997a, 1997b). Moreover, AMs scored by CRAM show a temporal decay in content, a typical component of AM retrieval (e.g., Levine et al., 2002; Janssen et al., 2011; Piolino et al., 2002).

The current work builds on this research, which focused on college-aged subjects, by applying CRAM to individuals of various ages across the adult life span (18-78 years old: yo). We utilized both in-person and Internet-based testing (http://cramtest.info) to further enhance data collection. The resulting data provide numerical counts of AM content from a diverse subject pool that should expand our understanding of the relationship between aging and recollection.

This research was conceived to describe age-related changes in AM content. Similar to studies that quantitatively described age-dependent modulation of the temporal distribution of AM retrieval (Rubin et al., 1986; Rubin and Wenzel, 1996; Rubin and Schulkind, 1997a, 1997b; Rubin, 2000) this work aims to quantitatively characterize age-dependent modulation of feature-specific recollection. Nonetheless, application of CRAM to older subjects may also contribute to AM theory. The reminiscence bump is an increase in retrieval of AMs that recall episodes from adolescence to early adulthood and is most clearly observed in older adults (see Jansari and Parkin, 1996; Janssen et al., 2005; Janssen at al., 2011; Rubin et al., 1986; Rubin et al., 1998). While previous studies show that AMs from the bump, as collected using the word-cue technique, are not associated with enhanced phenomenological characteristics of recollection, e.g., vividness or re-living (Janssen et al., 2011; Rubin and Schulkind, 1997a), whether content counts of these memories correlate with their retrieval probabilities remains an open question. For example, it is possible that memories rich with detail have relatively high association probabilities with a given memory cue, causing these AMs to be frequently accessed. This approach and resulting data may be useful to inform theories of memory, e.g., multiple trace theory (Nadel et al., 2000).

Methods

The Cue-recalled Autobiographical Memory test (CRAM)

CRAM is a computerized interactive test presented in web-browser format. It collects counts of the number of details (elements) within categories (features) that comprise naturalistically, word-cued AMs dated to specific life periods. Complete details of the test and instruction provided to participants have been previously reported (Gardner et al., 2012; see http://cramtest.info). Here, each section of the test is briefly described.

Prior to eliciting AMs, CRAM collects demographic information for each participant. Subjects are then presented the following definition of AM and subsequent instruction:

"Autobiographical memories are recollections of past episodes directly experienced by the subject. These memories should be of a brief, self-consistent episode of your life. An episode can be as short as a single snapshot and up to a few seconds long. … If the memory you think of refers to a typical and repeated episode that happened regularly or multiple times in your life, you can use it only if you can fixate on a specific individual event. If you can only recall the generic (repeated) event, look for another memory."

Naturalistic word-cues are then presented to elicit memories. The participant reads through a list of seven words and labels the first recollection retrieved, for subsequent identification. Subjects are further instructed that the cued AM does not necessarily have to relate to any one word or to the entire list of words, but rather is the first AM that comes to mind. The list of word-cues is randomly selected from the British National Corpus (http://www.natcorp.ox.ac.uk), a compilation of one-hundred million written and spoken words. Thus, this procedure provides cues that are presented proportionally to word frequencies observed in everyday settings.

Once AMs are cued and labeled, participants are presented with their AM labels one by one to date each memory. Specifically, the participant places each memory into one of ten temporal bins, which segment his or her life span into ten equal intervals (termed Youths: 0 – 9). Youth refers to the age of a recalled episode with a higher Youth indicating an AM of a more recent event. If necessary, up to three Youth bins could be assigned to a single AM. To increase dating accuracy, the temporal range associated with each bin is presented in terms of time from the present, age of the participant, and month and year.

The age of the participant at the time of a recalled event is estimated as the mid-point of its assigned Youth(s). The retrieval probability of an AM falling within each

Youth or age range is computed by dividing the number of AMs dated to each temporal period by the total number of dated AMs. Collectively, these measures are used to construct the temporal (life span) distribution of AMs. In this work, Recent AMs are defined as memories of events that occurred within the most recent ten years of life; Remote AMs are defined as memories of events that occurred more than ten years from the present moment. Overall, ~4% of AMs were dated into more than one Youth. This proportion mildly increased for Remote AMs (Remote: 5.0%; Recent: 3.7%, $p < 0.001$), reflecting a reduction of a given subject's confidence in dating these older episodes. In addition, younger subjects dated AMs to multiple Youths slightly more frequently (4.0%) than older subjects (3.0%; $p < 0.05$), potentially due to the age-dependent nature of a Youth's temporal interval size (the temporal interval associated with a Youth from a twenty year old is half the size of a Youth from a forty year old).

After AMs are dated, participants are once again presented with their AM labels to score the content associated with a selected memory. In particular, participants are instructed to count the number of details remembered within each of eight categories, i.e., *Things* (objects), *Feelings* (emotional details), *People* (unique individuals), *Places* (spatial details), *Times* (temporal details), *Episodes* (temporally linked events), *Contexts* (other contextual details), and *Details* (all remaining details, including actions). The order in which each category is presented is randomized for each person but fixed across all AMs for a given individual. The exact definitions of these eight categories were previously reported (Gardner et al., 2012) and are available at http://cramtest.info. In addition, CRAM provides, through clickable links, additional examples of what constitutes a detail within a given category and general scoring guidance. Each detail category is called a “feature,” each reported detail associated with a given feature is referred to as an “element,” and the summed number of elements across all features for a given memory is called “total content.”

In-person and Internet testing

CRAM was completed locally under experimenter supervision or remotely over the Internet (http://cramtest.info). From each subject who completed testing in person, thirty AMs were cued and dated, of which a subset of ten was scored for content (see Gardner et al., 2012). AMs were selected for scoring to maximize life span coverage in the entire dataset. Specifically, a single AM was scored from each Youth represented by a participant. From the participant’s remaining AMs, memories were selected in order from the least to most represented Youth (in the entire dataset across all participants) until ten memories were scored in total.

In contrast to in-person testing, CRAM’s online protocol offers subjects the choice of several test options which differ according to the number of AMs cued, dated, and scored. These options are included to promote test completion by suiting a wide range of subjects who may vary in their commitment and eagerness to participate. The Atomic test cues one AM which is dated and scored for content. At the end of the Atomic test, subjects are asked if they would like to complete the Mini or Full test. The Mini test cues five AMs, which are each dated and scored. At the end of the Mini test, subjects are invited to extend the test. If they agree, an additional fifteen AMs are cued and dated, of which, five are scored (Extended test). The Full test cues twenty AMs, each of which is dated. Subsequently, a subset of ten is scored for content. Memory selection for scoring in the Full and Extended tests follows the same rules as those for in-person testing. Given these selection rules, the proportion of scored AMs in a particular Youth may differ from that typically retrieved. Thus, when presenting aggregate measures of content (i.e., those within a given subject grouping) not restricted to a particular life period, content values across Youths are weighted according to the applicable AM temporal distribution.

Participants are encouraged to complete the Full test (or Extended test, if opting initially for the Mini test). This is accomplished by pre- and post-test advertisement for the opportunity to explore an interactive summary report of one's results with the ability to make direct comparisons to results from specified age ranges, solely after completion of the Full (or Extended) test. We stress that although test types were varied (e.g., in duration), all types provided subjects with the same instruction on AM classification, cueing, dating, and scoring (which were also identical to in-person testing). Unless indicated otherwise, data were collapsed across testing conditions and test types.

Participants

As CRAM is freely accessible online (http://cramtest.info) and indexed by popular search engines, data are continuously collected from Internet-browsing individuals. To supplement these unsolicited data, additional individuals were actively recruited from the undergraduate population of George Mason University (GMU), from GMU staff and faculty, and from the local community, obtaining in all cases informed consent. With the exception of undergraduates, recruited subjects were given the choice to complete testing locally at GMU with a researcher present or remotely over the

Internet. Recruited undergraduates (ages 18-36 yo) invariably completed the study for course credit and took the test under experimenter supervision; these data from recruited students have been reported previously (Gardner et al., 2012) and included here to best estimate recalled content in relatively young subjects. However, the amount of data collected from this age range was substantially augmented by the current approach (exclusively through online testing) almost tripling the previous sample of scored AMs (when pooled together) from these younger subjects (previously reported AMs scored by subjects 18-36 yo: *n*=1424, Gardner et al., 2012; currently reported: *n*=4027). No identifiable personal data were stored. All recruitment and testing procedures were approved by the GMU institutional review board.

In total, 17,482 AMs were dated from 2,561 unique test IDs/subjects (*Mean Subject Age* = 34 yo, *SD* = 14, range: 18-78 yo; 67% female; 81% native English speakers). Thirty-two percent of these AMs were collected in-person (predominantly from college-aged subjects: *Mean Age* = 22 yo, *SD* = 7; 76% female; 74% native English speakers) and 68% online (*Mean Age* = 36 yo, *SD* = 14; 67% female; 81% native English speakers). Of these AMs, 6,492 were scored for content (76% scored online). Most Internet subjects opted for the Full test (from which 52% of Internet-scored AMs were collected); the Atomic (20%), Mini (16%), and Extended (12%) tests were less likely to be completed. Test choice was not associated with subject age (i.e., each test type showed roughly identical distributions of ages to that found overall).

Data screening

Data were inspected to identify data entry errors or otherwise lazy and inauthentic reporting (e.g., see Gardner et al., 2012). Positive cases were removed from analysis. For example, AMs were excluded if a subject reported an identical number of elements for each of the eight features. In addition, all AMs were removed from seventeen participants whose scoring across the majority of their AMs reflected this pattern (201 AMs in total). Memories were also excluded from two participants who reported either one or eleven elements in each feature category across all scored AMs (14 AMs), and from two subjects who reported unique scores for each feature but identically scored all memories (15 AMs). Altogether, these exclusions totaled 232 scored AMs (2.5% of total).

Subsequently, extreme total content values were identified as those greater than three times the Inter-Quartile-Range (IQR) above the 75th percentile, or less than three times the IQR below the 25th percentile within a given age range. Data meeting either one of these criterion were considered outliers and excluded from analysis. This procedure was performed separately for AMs collected from each of the following age ranges: 18-25; 26-35; 36-45; 46-55; 56-65; 66-78 years of age. This resulted in the removal of 223 AMs. Specifically, 78 AMs were excluded from 18-25 yo subjects (3.0% of the total within this age range), 52 AMs (3.5%) from 26-35 yo subjects, 39 AMs (4.0%) from 36-45 yo subjects, 33 AMs (5.1%) from 46-55 yo subjects, 16 AMs (4.5%) from 56-65 yo subjects, and 5 AMs (2.1%) from 66-78 yo subjects.

Statistics

Binary logistic regression was run to assess the effects of participant groups on AM retrieval probabilities across life periods (i.e., Recent and Remote). Bivariate regression was performed to evaluate correlation between measures of AM recall and participant age, and inter-feature relationships. ANOVA was performed to evaluate changes in AM content across participant groups and life periods. Results were corroborated using a generalized estimating equation approach (Davis, 2002). Where applicable, for robustness analyses (i.e., analysis across genders, native languages and testing conditions), subject age, and/or cuing and scoring order were assigned as covariates to control for variation in the outcome variable explained by these sources. Chi-square analysis with Yates correction was run to assess group differences in the proportion of AMs assigned to multiple Youths. Cohen's *d* was calculated for each comparison to estimate effect size. Statistical significance was interpreted using the criterion of $p < 0.05$. False discovery rate correction was applied to multiple comparisons (Benjamini and Hochberg, 1995). Statistical analyses were performed using SPSS (IBM), Excel (Microsoft), and R (Dalgaard, 2008).

Results

AM retrieval probabilities are modulated by subject age and life period

Figure 1 displays AM retrieval probabilities for each of the ten temporal bins among various age groupings (plotted according to participant age at the time of the event: Fig.1A, and Youth: Fig. 1B; see Methods). We observed a large proportion of AMs recalling recent events, which declined steeply with time from the present moment (retention interval). We also found a relative increase in the number of AMs dated to adolescence through young adulthood (the bump), and a relative absence of AMs from early childhood (childhood amnesia).

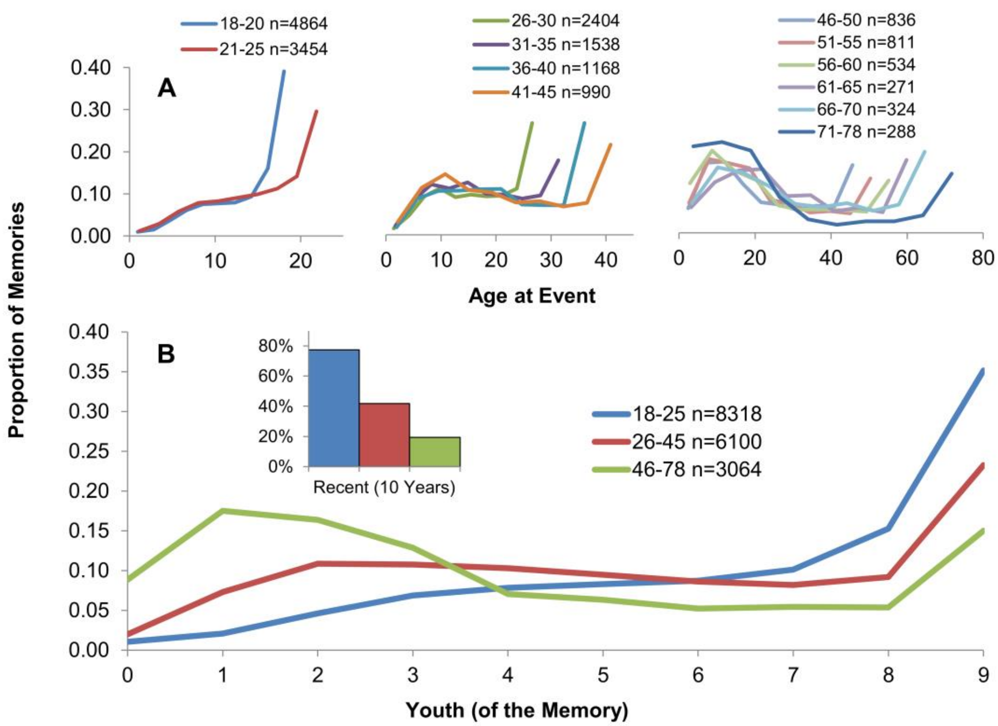


Figure 1. ***Temporal distributions of AMs.***
AMs cued by CRAM were dated into one of ten temporal bins (Youth) that divided a subject's life into ten equal intervals. Participant age at the time of an event for a specified age group was computed as the midpoint of a given Youth (see Methods). The proportions of AMs retrieved across these temporal bins are plotted according to (A) the age at the recalled event, and (B) the Youth of the memory for various age groups (n: number of dated AMs). (A-B) Younger subjects displayed the greatest retention of recent AMs, which gradually decreased with increasing age. (B Inset) The magnitude of this age-effect is illustrated by the proportion of AMs pertaining to events from the most recent ten years across age groups. The reminiscence bump (an increase in AMs recalled from adolescence to early adulthood) emerged in subjects in their mid-to-late twenties and was most prominent in subjects older than 45 yo. Childhood amnesia (a paucity of AMs recalled from the first few years of life) was clearly observed in younger subjects (18-45 yo). Its absence in older subjects (46-78 yo) is likely due to our dating procedure,

as the first tenth of life in this age group is longer than the typical interval of childhood amnesia.

Moreover, these characteristics of the temporal distributions of AMs produced by CRAM changed with participant age. Retention of recent AMs was strongest in younger subjects (18-25 yo) and decreased with increasing participant age ($p < 0.001$). For example, AMs dated to the most recent 10 years of life comprised ~77% of all AMs cued from 18-25 yo subjects, ~42% of those cued from 26-45 yo subjects, and ~19% of those cued from subjects older than 45 years (Fig. 1B Inset; Table 1). Furthermore, while absent in the two youngest age groups (Fig. 1A), the reminiscence bump emerged in subjects in their mid-to-late twenties, and was most evident in subjects older than 45 yo (Fig. 1A-B). The peak of the bump corresponded to the years between ages 8 and 22, depending on the age group (Fig. 1A). Age ranges presented in Fig. 1B reflect the changing clarity of the reminiscence bump quantified as the ratio of the peak retrieval probability across Youths (excluding those associated with the retention interval) to the subsequent minimal retrieval probability. On average, this ratio is undefined in 18-25 year olds (due to the lack of a minimum, reflecting the absence of a bump), greater than one in subjects 26-45 yo, and greater than two in adults older than forty-five. Childhood amnesia was observed in younger subjects (18-45 yo) as demonstrated by a notable drop in AMs dated to the first tenth of the life span (less than 2%; Fig. 1B). Its apparent absence in older subjects (i.e., 46-78 yo; Fig. 1) is likely an artifact of our methodology. Specifically, the first tenth of life of an older individual extends beyond the relatively narrow temporal period associated with childhood amnesia, and thus limits our ability to isolate and analyze memory for these very early life events.

Reported AM content is moderately increased in older adults

A total of 6,037 scored AMs were analyzed for content (see Methods for data screening procedures). On average, subjects reported ~22 elements (*SD*=14) per AM. We found a mild yet significant positive relationship between total reported content and participant age ($r = 0.11$, $p < 0.001$). To further investigate this finding, memories were separated into six age divisions (see Fig. 2A). We observed a moderate increase in the amount of AM content among older (46-55 yo: *Mean Total Content* = 24.8, *SD* = 16.3, *d* = 0.24; 56-65 yo: *M* = 25.6, *SD* = 16.1, *d* = 0.30; 66-78 yo: *M* = 24.3, *SD* = 16.8, *d* = 0.20) compared to younger (18-25 yo: *M* = 21.3, *SD* = 12.1) subjects; whereas reports of content from age divisions younger than 46 yo were found to be relatively equivalent to those from the youngest group (26-35 yo: *M* = 20.7, *SD* = 14.0, *d* = 0.05; 36-45 yo: *M* = 22.6, *SD* = 16.4, *d* = 0.09).

Thus, the age-related increase in the amount of detail reported in typical autobiographical recollection emerges most noticeably in subjects in their mid-to-late forties and persists into old age (i.e., late seventies; Fig. 2A). Given this pattern, to simply describe the magnitude of these effects, AMs were divided into those collected from younger (*Mean Age* = 27, *SD* = 8, range: 18-45 yo; 72% female, 80% native English speakers) and those from older (*Mean Age* = 57, *SD* = 8, range: 46-78 yo; 65% female, 89% native English speakers) participants (Fig. 2B). This grouping revealed a ~16%

increase in the number of reported details for a given AM in older compared to younger subjects (~25 vs. ~21; $p < 0.001$, $d = 0.23$; Fig. 2B; Table 1).

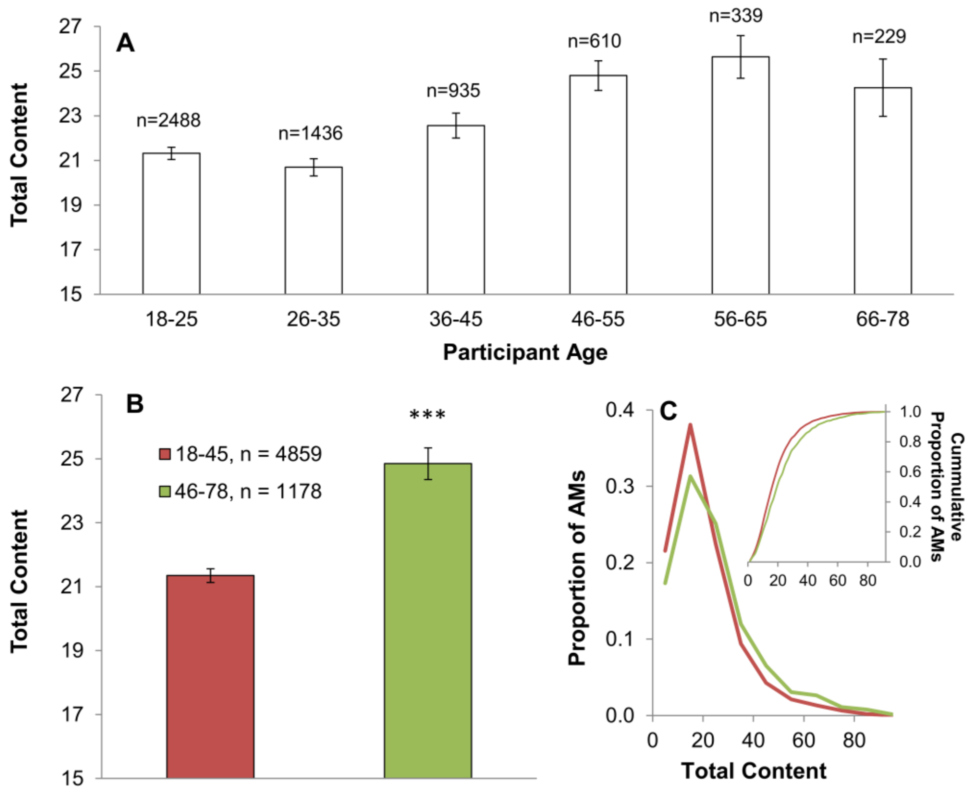


Figure 2. ***Moderate increase of total reported content in older individuals.*** *The total number of unique details reported for a given AM is displayed for various age groups (n: number of scored AMs; error bars: standard error of the mean; *** p < 0.001). (A) Total content moderately increased with the age of the participant, most noticeably in subjects older than 45 yo, an effect that persisted into old age. (B) The magnitude of these effects is illustrated by pooling AMs from those subjects older than 45 yo, and those 45 yo or younger. This grouping revealed a ~16% age-related increase (~4 elements) in total reported content. (C) Histograms of these two age groups show a mild right tailed distribution.*

While reported content was quite variable from memory to memory, the coefficient of variation was similar across all ages (~0.6-0.7; Table 1). Histograms of total content are presented in Fig. 2C. Half of all scored memories in younger subjects were comprised of ~16 or fewer elements; half of all scored AMs in older subjects contained more than ~20 elements.

Total content decays with the age of the memory among younger and older subjects alike

Older subjects reported significantly more content than younger subjects for memories of all life periods, both when comparing equivalent decades of life (Fig. 3), and when comparing relative life periods (e.g., Recent: the most recent 10 years, and Remote: >10 years from the present; see Methods; Fig. 3 Inset; Table 1). The numerical results of content analysis of Recent and Remote AMs only marginally fluctuated depending on how these temporal intervals were defined (e.g., restricting Recent AMs to the most recent five years; restricting Remote AMs to the first decade of life; defining these life periods by Youth). In addition, all reported conclusions remained unchanged and did not depend on the particular grouping applied (data not shown).

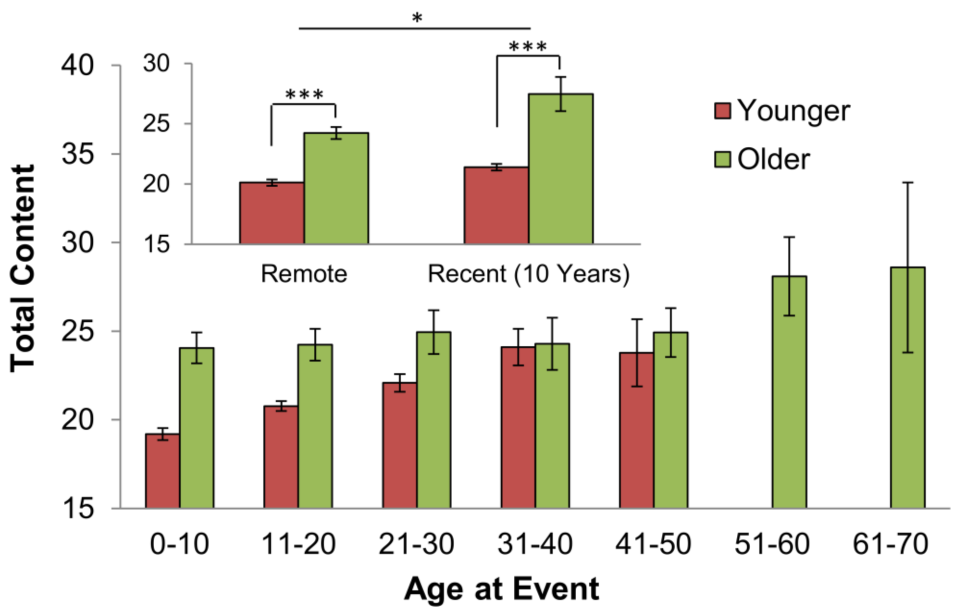


Figure 3. ***Reported details from Remote and Recent AMs across age groups.*** *Older individuals reported a greater number of total details from each decade of life (Younger subjects: 18-45 yo; older subjects: 46-78 yo; error bars: standard error of the mean). Inset: Recent AMs are those dated to within ten years of the present; all remaining AMs are Remote (* p < 0.05; *** p < 0.001; see Table 1 for sample sizes within Remote and Recent intervals). Content declined with the age of the episode among all subjects.*

Older subjects reported ~24 elements from AMs dated to the first two decades of life compared to ~20 elements in younger subjects ($p < 0.001$; $d = 0.27$). Likewise, Recent AMs from older subjects were comprised of ~27 elements compared to ~21 elements in those from younger individuals (Fig. 3 Inset; $p < 0.001$; $d = 0.38$). Total content declined with the age of the episode among all age groups. For example, Remote AMs were comprised of significantly fewer details compared to Recent AMs ($p < 0.01$;

Table 1; Fig. 3). These results confirm those previously reported in college-aged subjects (Gardner et al., 2012) and extend these findings to older adults.

Table 1. ***AM content and retrieval across life periods and age groups.***
Recent AMs are those dated to the most recent ten years; all remaining AMs are Remote (SD: standard deviation; CV: coefficient of variation; IQR: inter-quartile range; *n*: number of AMs within each subgroup).

		Remote			Recent (10 Years)		
		18-25 yo	26-45 yo	46-78 yo	18-25 yo	26-45 yo	46-78 yo
Total Content	Mean	19.05	20.68	24.23	21.04	22.17	27.46
	SD	11.70	14.83	16.07	12.23	15.30	18.02
	CV	0.61	0.72	0.66	0.58	0.69	0.66
	Median	16	17	21	19	19	23
	IQR	13	17	18	15	16	23
AMs Dated	*n* (%)	1890 (23%)	3549 (58%)	2469 (81%)	6428 (77%)	2551 (42%)	595 (19%)
AMs Scored	*n*	826	1592	1016	1662	779	162

Features selectively contribute to the age-related increase in total content

A relatively high proportion of AM content (~46%; ~10.1 elements) was associated with *Places, Things*, and *People*. In contrast, *Times*, *Contexts*, and *Episodes* were less represented, together comprising just ~29% (~6.4 elements; *Details* and *Feelings* were close to the average: ~13% and ~12%, respectively; Fig. 4A). Largely in line with these distributions, the proportion of AMs containing at least one element of a particular feature was highest for *Places* (98%) followed by *People* and *Feelings* (92% each), *Things* (89%), *Details* (84%), *Times* (83%), *Contexts* (78%), *and Episodes* (63%). Feature variation was higher than that observed for total content, but equivalent across age groups (*Mean Feature CV* = 1.0). All these findings uphold those previously reported in younger subjects (Gardner et al., 2012).

Adding to this research, we found that older adults reported a greater number of elements among all features. *Episodes* and *Things* showed the most prominent age-related content increase (~30% and 27%, respectively; $p < 0.001$), while content associated with *People* and *Times* remained close to that observed in young subjects ($p > 0.10$; Fig. 4A-B). These data collectively reflect the finding that the age-related increase in total content is not uniformly distributed across features (in terms of either absolute or relative value). In particular, *Things* (24%), *Episodes* (16%) and *Details* (15%) contributed more substantially to the age-related increase observed in total content (Fig. 4C), whereas contributions from *People* (4%) and *Times* (3%) were definitively smaller. The proportion of AMs containing at least one element from a given feature was significantly higher in older compared to younger subjects for *Feelings, Things, Details, Contexts,* and *Episodes* (~5% higher on average; $p < 0.001$) but not *People, Places* and *Times* ($p > 0.10$). These feature distributions were largely consistent between Remote and Recent AMs in both younger and older subjects (Fig. S1); notably, however, the feature *Times* appeared to be least resilient to temporal decay of content among all subjects.

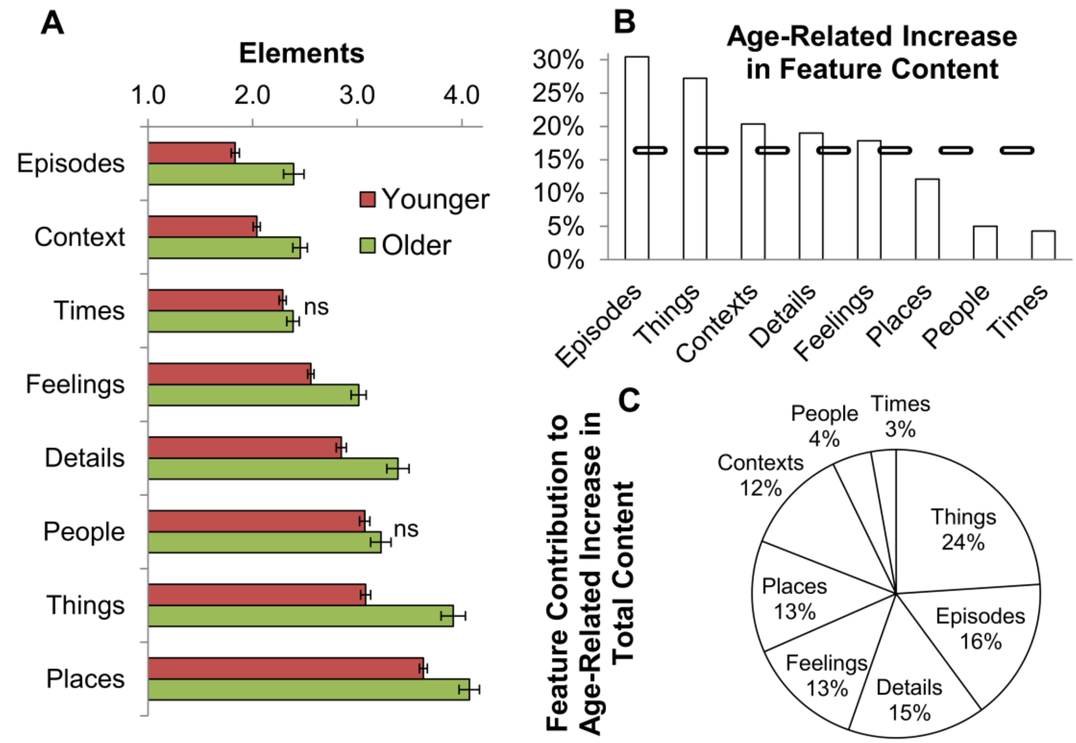


Figure 4. ***Selective feature contribution to the age-related increase in total content.*** *(A) Subjects of all ages reported more AM elements related to Places, Things, and People, and fewer related to Episodes, Contexts, and Times (Younger subjects: 18-45 yo; older subjects: 46-78 yo). Older adults reported more elements within all features (the age-related increase is statistically significant unless marked as ns; error bars: standard error of the mean). (B) Episodes and Things showed the most prominent age-related increase, as People and Times stayed close to that observed in younger adults. The dashed line indicates the overall age-related increase in total content (~16%). (C) The observed increase in total reported content from younger to older subjects (~4 elements; see Fig. 2) was largely composed of Things, Episodes, and Details.*

People is a relatively independent feature of recall among all ages

Correlation analysis showed positive relationships among all features (*mean Pearson* $r = 0.37$). Moreover, these relationships were similar across age groups (younger: $r = 0.36$; older: $r = 0.40$) as well as between Remote and Recent intervals

(Remote: $r = 0.39$; Recent: $r = 0.37$). To further evaluate feature dependence, the correlation between each feature and all other content was computed. The average of these values across all features was equivalent between age groups (younger: $r = 0.53$; older: $r = 0.57$; Fig. S2A) and life periods (not shown). However, *People* exhibited a comparatively mild relationship to the remaining AM content ($r = 0.37$; 32% less than the average: Fig. S2B) across all conditions, suggesting that it is a relatively independent feature of recall. The inter-feature relationships found here confirm those previously reported among younger subjects (Gardner et al., 2012) and extend these findings to individuals across the life span.

Estimates of retrieved content across age groups and life periods

This work provides numerical description of AM retrieval probabilities and reported content associated with distinct life periods. Combining these two measurements, we can estimate the relative distribution of retrieved content across temporal intervals and age groups (Figure 5). Specifically, given the number of elements typically reported for a single retrieved AM, for each of various age groups (as outlined in Fig. 2A and plotted in Fig. 5 by mean age), we computed a probabilistic content distribution among life periods. For example, when experiencing an AM, a typical 70 yo on average reports a similar content amount (~4.5 elements) from his or her middle teenage years and from the last few years of his or her life. In contrast, content from events dated to the most recent two years of the life of a 21 yo is ~4 times more represented in memory than that associated with events from his or her middle teenage years. Assuming that the frequency of AM recollection is stable across age groups, we can further estimate the relative probability that a particular recalled element is associated with a certain age group and life period. For example, the likelihood that content is retrieved by a 30 yo from events dated to his or her early twenties (~1.3%) is equivalent to the likelihood that it is retrieved by a 60 yo from events dated to his or her early thirties.

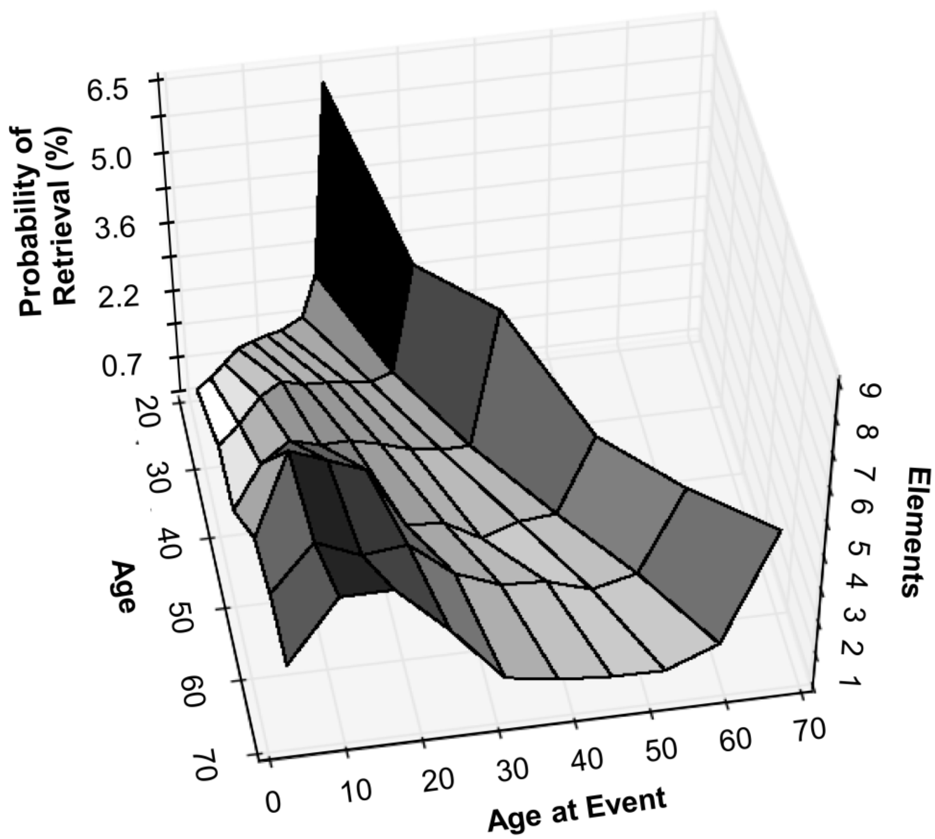


Figure 5. ***Estimated distribution of retrieved content across age groups and life periods.*** *Content measures within each temporal bin for each age group were normalized to their applicable retrieval probability. In particular, the number of elements per life period was computed for each of the six age ranges outlined in Fig. 2A (and plotted here by mean age within each group). These relative values provide estimates of the likelihood that any*

given amount of retrieved content is associated with a particular age and life period (shown as retrieval probabilities).

The bump is unrelated to changes in total content and feature content.

AMs found within the reminiscence bump may have distinct recall characteristics. For example, AMs that comprise the bump may be comparatively rich with recalled detail. Such a finding would explain, at least in part, why this life period plays a particularly prominent role in subjective experience. We evaluated content from AMs within and beyond life periods associated with the bump in older subjects (46-78 yo) for whom this phenomenon was most pronounced. In this age range, we observed a maximal AM retrieval probability (excluding Youths associated with the retention interval) from Youth 1 (~17%) and subsequent minimal probability from Youth 6 (~5%; Fig. 1). However, measures of total content from these Youths did not significantly differ (M = 23.99 vs. M = 22.39; $p > 0.10$, d = 0.10). Likewise, measures of content from ages 11-20 compared to 31-40 were equivalent (M = 24.90 vs. M = 25.28; $p > 0.10$, d = 0.02; Fig. 3). The composition of memories dated to these life periods (both based on Youth and life decade) were also quite stable ($p > 0.10$), with any given feature showing a deviation of ~2% or less (*Mean deviation*: ~1%).

All the main findings of this work are robust to gender, native language, and changes in experimental procedures (see Tables S1-S2). However, several minor distinctions in AM recall were observed between males and females, between native and non-native English speakers, and among testing conditions. A full review of these results is included as Supplemental Information.

Discussion

This work provides quantitative measures of autobiographical memory content reported from individuals that represent a substantial segment of the adult population (18-78 yo). This was accomplished using CRAM, an instrument designed to collect counts of details that fall within specified features from naturalistically elicited AMs dated to particular life periods. Relying on participant counts of memory content (rather than experimenter-scored participant narratives) facilitates data collection and thus enables fine-scale analysis of AMs (e.g., as shown in Fig. 5).

We previously demonstrated that CRAM replicates several noted observations of AM recall (e.g., on the retention of recent AMs and their associated content; Janssen et al., 2011; Levine et al., 2002; Piolino et al., 2002; Rubin and Schulkind, 1997a, 1997b). The present study demonstrates that CRAM is also a reliable tool; the current results confirmed our previous findings from college-aged subjects. In particular, the current estimate of AM total content (~21 elements) among younger subjects is almost identical to that reported (~20 elements) by Gardner et al. (2012). In addition, the current work supports previous findings on the temporal decay of content (Remote AMs contain fewer elements), AM content variability (for any given AM, total content remains relatively stable compared to feature content), AM composition (*Places, Things, and People* are prominent features of recall), and inter-feature correlations (*People* is a relatively independent recall characteristic).

Extending CRAM to older adults, we replicated prior reports of age-modulation of the temporal distribution of AMs. Specifically, Recent AMs were considerably less likely to be recalled in older subjects (e.g., Janssen et al., 2011), and the reminiscence bump emerged in subjects in their mid-to-late twenties. While some studies have found that the bump is not apparent until ~40 years of age, using relatively small temporal bins, Janssen et al. (2011) reported a similar onset to that found here. In addition, the temporal interval associated with the bump using CRAM (i.e., 8-20 years old) is consistent with these studies.

Total content was positively correlated with subject age. Older adults reported ~25 details for a given AM, ~4 elements more than the number reported from younger subjects. Moreover, this age-related increase in content was observed for Recent and Remote memories, and was most drastic for the features *Episodes* (sequences of event) and *Things* (objects) while negligible for *People* (unique individuals) and *Times* (temporal detail). Altogether, these data quantitatively describe an age-associated shift in how events are remembered.

As older adults reported more content than younger adults from memories that originated in the same decade of life (i.e., those AMs that have similar ages of encoding), these findings appear to highlight the constructive nature of AM (Bartlett, 1932; Conway and Pleydell-Pearce, 2000; Hupbach et al., 2007). This interpretation assumes that the initial number of encoded event details and encoding depth are similar between age groups. However, as with all cross-sectional aging research, any inter-group differences may reflect generational differences rather than (or in addition to) changes that occur among individuals across their life span.

While it also remains possible that older and younger adults used divergent strategies to establish feature counts, the finding that the age-effects on content were feature-specific (see Fig. 4) argues against a general change in interpretation of CRAM's instruction, and/or adjustment in content evaluation (e.g., a pervasive tendency for older individuals to report higher scores). Further work, including the use of longitudinal designs, is required to clarify the mechanisms underlying the current findings, and the degree to which reported AM content among distinct age groups (and life periods) contains constructed or created details.

Older individuals are proposed to have altered narrative attention and to tell more interesting life stories (James et al., 1998). Although more detail is not always better, inclusion of information about sequences of happenings (the feature that was most strongly augmented from younger to older subjects) within a life narrative enables placement of an episodic snapshot into a broader context of surrounding events (and may enhance storytelling). It would be interesting to establish how our findings on feature-specific age-related modulation of reported AM content compare to the types and amount of detail shared through social communication of event memories and during the narration of life stories.

Despite the moderate change in total reported content, several properties of recollection were stable across age groups. In particular, AMs from younger and older subjects, and those from Remote and Recent life periods, showed similar feature distributions. In addition, among all ages, fewer details were reported from Remote than from Recent AMs. Thus, these data are indicative of two independent age effects on AM content: a positive correlation with the age of the individual and a negative correlation with the age of the event. These findings confirm and extend prior studies (Janssen et al., 2011; Gardner et al., 2012).

Additionally, independent of age, almost all AMs had at least one detail related to location, and nine out of ten memories included some information about people, objects, and feelings, suggesting that these features are quite memorable and/or at the core of AM. In contrast, less than two-thirds of memories included sequential events. The feature *People* was also relatively independent from other features, further demonstrating its unique role in retrieval: remembrance of the individuals associated with specific life experiences is essential to form and maintain social relationships (a proposed function of AM; see Bluck et al., 2005; Bluck and Alea, 2011; Pillemer, 1992; Waters, 2014).

Combining the observed temporal distributions of AM retrieval with measure of AM content permits computation of detailed probability estimates to report a detail from a given life period at a particular age (Fig. 5). For instance, this approach quantifies how likely it is for a detail retrieved by a sixty year old to be associated with an episode from his or her mid-to-late twenties (~9%). Moreover, we can address questions on how these probability distributions change with subject age. For instance, how does the previously computed probability compare to the likelihood that the same amount of content retrieved by a twenty year old stems from a relatively recent event?

Several factors are proposed to account for the high accessibility of memories that comprise the reminiscence bump, e.g., neurocognitive development and cultural influence (Berntsen and Rubin, 2004; Bohn and Berntsen, 2010; Rathbone et al., 2008; Rubin et al., 1998; Schrauf and Rubin, 2000). We add to an understanding of the bump by showing that neither counts of detail across all features nor counts within individual features explain or are explained by the relatively high probability of recollection associated with this life period. These data are in line with previous studies that have demonstrated that AMs within the bump do not have higher ratings of certain characteristics of recollection (e.g., vividness, rehearsal, reliving, novelty, emotionality; Janssen et al., 2011; Rubin and Schulkind, 1997a).

Past approaches reporting counts of AM content have predominately focused on the distinction between episodic and semantic retrieval (Levine et al, 2002; Piolino et al., 2002; Addis et al, 2008; Addis et al., 2010). Episodic memory recounts a unique personally-experienced event, with some form of contextual information (e.g., spatiotemporal detail). Semantic memory recalls abstracted knowledge of the world or of oneself (generally acquired from repeated experiences) that does not describe or call to mind a unique episode (see Tulving, 1972, 1985). This distinction is highlighted by case reports of neurocognitive deficits following targeted brain damage (Rosenbaum et al., 2005) and is present in numerous theories of AM (e.g., Conway and Pleydell-Pearce, 2000). Using a narrative scoring technique of event-cued AMs specified to life periods, Levine et al. (2002) found that, compared to younger subjects, older adults report fewer episodic but a similar or greater amount of semantic details from typical memories. This age-effect on episodic detail appeared to be feature-dependent as it was absent (but never reversed) for some features (e.g., *Times*). Contrasting accounts, however, have been reported. Hashtroudi et al. (1990) found that older adults reported greater content for feelings and thoughts (albeit these same individuals showed a reduction in sensory-perceptual recollection). In addition, Janssen et al. (2011) found that ratings of AM vividness and re-living, proposed indices of episodic remembering, were higher in older subjects (also see Rubin and Berntsen, 2009; Rubin and Schulkind, 1997a). Direct comparison of memory content scores obtained using a variety of quantitative and qualitative approaches will be a useful endeavor to clarify the apparent discrepancies in findings in AM recollection across age groups.

We emphasize that CRAM was broadly designed to measure the details that a participant considers part of an AM, i.e., the subjective content associated with a retrieved life event. As such, CRAM does not classify reported elements as episodic or semantic. However, each feature definition was worded to collect the details that comprise a memory for a temporally-specific event; likewise, the general guidance provided emphasized reporting of detail unique to the specified episode (see Methods; Gardner et al., 2012). As CRAM collects reports of retrieved subjective detail, this approach also contrasts with those that aim to collect "true" or verifiable detail or to collect all potentially retrievable details associated with an event (e.g., Levine et al., 2002: specific probing; Mello and Fisher, 1996). As CRAM's instruction was identical between all subjects, however, relative measures between and within age groups should reflect genuine changes in subjective recollection.

Altogether, the data presented here provide previously inaccessible fine-scale quantitative characterizations of the subjective content of AMs as a function of the age of an individual and the age of a memory. These characterizations point to a moderate but significant age-associated feature-specific shift in how one's life story is remembered and recounted.

Acknowledgements

This research was supported in part by the Air Force Office of Scientific Research (Award No: FA9550-10-1-0385) and the Office of Naval Research (Award No: 000141010198).

Supplemental Information

AM retrieval and reported content are largely robust to participant subgroups and testing conditions

All the main findings of this work are robust to gender, native language, and testing condition. Similar to the overall findings, each participant group (i.e., males, females, native English speakers, and non-native English speakers) showed age-related modulation of life span retrieval probabilities and age-related augmentation of total content (Tables S1-S2). The distribution of features that composed a typical AM was stable across all of these groupings (plus or minus less than 2% for any given feature), as were feature correlations (data not shown). In addition, Internet-collected AMs from younger and older groups were equally distributed among test types (see Methods), and showed similar measures of recall to those collected in-person.

Nonetheless, several minor differences in AM recall were observed across participant groups and testing conditions. By and large AM retrieval and reported content were equivalent across genders and native languages (Tables S1-S2). However, among subjects aged 26-45 years old (yo), we observed an increase in the number of AMs dated to Remote life periods from native English speakers (61%; compared to non-native English speakers: 47%; $p < 0.001$). In addition, the number of reported elements was increased among females (~22.8; compared to males: ~18.3; $p < 0.001$, $d = 0.31$) and native English speakers (~22.2; compared to non-native English speakers: ~18.3; $p < 0.001$, $d = 0.28$) in this same age range. The general stability of feature distributions across all comparative conditions suggests that these effects are not a result of differential reporting of particular features. Nevertheless, these findings are restricted to a narrow age range (Tables S1-S2) with correspondingly small samples of scored AMs (males: n=738; non-native English speakers: n=460). Thus, proper interpretation of these differences will require further study of the influence of native language and gender on measures of AM retrieval and content.

Table S1. ***AM retrieval probabilities and content across genders.***
Measures of AM content (mean; standard deviation: SD; coefficient of variation: CV; median; inter-quartile range: IQR) showed an age-related increase among females and males alike. Both genders similarly showed an age-dependent decrease in the proportion of AMs dated to a Recent temporal interval (within ten years from the present). In subjects 26-45 yo, however, AMs from females (compared to males) were associated with a greater number of reported details. The number of AMs (n) within each subgroup is presented.

		Female			Male		
		18-25 yo	26-45 yo	46-78 yo	18-25 yo	26-45 yo	46-78 yo
Total Content	Mean	21.48	22.82	24.43	20.78	18.34	25.60
	SD	12.01	15.22	16.37	12.36	14.03	16.36
	CV	0.56	0.67	0.67	0.59	0.76	0.64
	Median	18	19	21	17	14	22
	IQR	14	17	17	15	15	21
AMs Dated	*n*	6174	3929	2036	2144	2171	1028
	(% Recent)	(77%)	(40%)	(20%)	(77%)	(46%)	(18%)
AMs Scored	*n*	1895	1633	786	593	738	392

Table S2. ***AM retrieval probabilities and content across native and non-native English speakers.***

Measures of AM content (mean; standard deviation: SD; coefficient of variation: CV; median; inter-quartile range: IQR) showed an age-related increase among native and non-native English speakers alike. Both groups similarly showed an age-dependent decrease in the proportion of AMs dated to a Recent temporal interval (within ten years from the present). In subjects 26-45 yo, however, AMs from native English speakers (compared to non-native English speakers) were associated with a greater number of reported details and were more likely dated to Remote life periods. The number of AMs (n) within each subgroup is presented.

		English			Non-English		
		18-25 yo	26-45 yo	46-78 yo	18-25 yo	26-45 yo	46-78 yo
Total Content	Mean	21.07	22.19	24.98	22.11	18.27	23.83
	SD	11.83	15.60	16.58	12.98	11.67	14.62
	CV	0.56	0.70	0.66	0.59	0.64	0.61
	Median	17	18	21	19	15	21
	IQR	13	17	19	15	13	18
AMs Dated	*n*	6439	4683	2738	1879	1417	326
	(% Recent)	(77%)	(39%)	(19%)	(78%)	(53%)	(21%)
AMs Scored	*n*	1959	1911	1044	529	460	134

All test variations produced AM temporal distributions with similar shapes and expected age-dependent modulation. Testing completed in-person, however, typically cued a greater number of AMs from Recent life periods compared to those cued over the Internet ($p < 0.001$); comparison between these conditions was restricted to young subjects (18-36 yo) as they accounted for more than 95% of the AMs collected in-person (see Methods). In addition, AM cue order was inversely associated with the age of the episode (i.e., years from the present; $r = -0.22$, $p < 0.001$). Thus, tests that collected more AMs (e.g., Full compared to Mini; Mini compared to Atomic; see Methods) typically elicited a higher proportion of Recent AMs. After controlling for cuing order, however, retrieval was equivalent across all test types ($p > 0.10$).

Total content was equivalent between AMs scored in-person and online (~21.1 and ~21.5, respectively; $p > 0.10$). In addition, feature contributions to total content and inter-feature correlations were generally robust to test location, although there was an increase in reporting of *People* during in-person testing (~13% of content was comprised of *People* in testing done online compared to ~17% when completed in-person). Among Internet-scored AMs, the Full test was associated with content scores lower than those produced by all other test types (Full: $M = 20.1$, $SD = 15.4$; All other Tests: $M = 24.6$, $SD = 14.4$, $d = 0.30$), an effect that appears to be driven by younger subjects. Fatigue potentially associated with longer tests did not underlie this finding, as scoring order did not affect reported content ($p > 0.10$). Since the test options were self-selected, it is plausible that those individuals who chose to complete the Full test vary systematically in the number of elements they typically retrieve, and/or how they interpreted CRAM's instruction. Alternatively, subjects who are more likely to randomly score AM content may be more likely to select specific test types. The finding that AM feature composition and inter-feature correlations were largely equivalent between all test types (data not shown), however, suggests otherwise. Altogether, variations in retrieval probabilities and AM content reported here add to the AM research literature showing that mild alterations in testing conditions can have statistically significant effects on outcome measures (e.g., as shown by demand characteristics and their effect on the temporal distribution of AM; Rubin and Schulkind, 1997a).

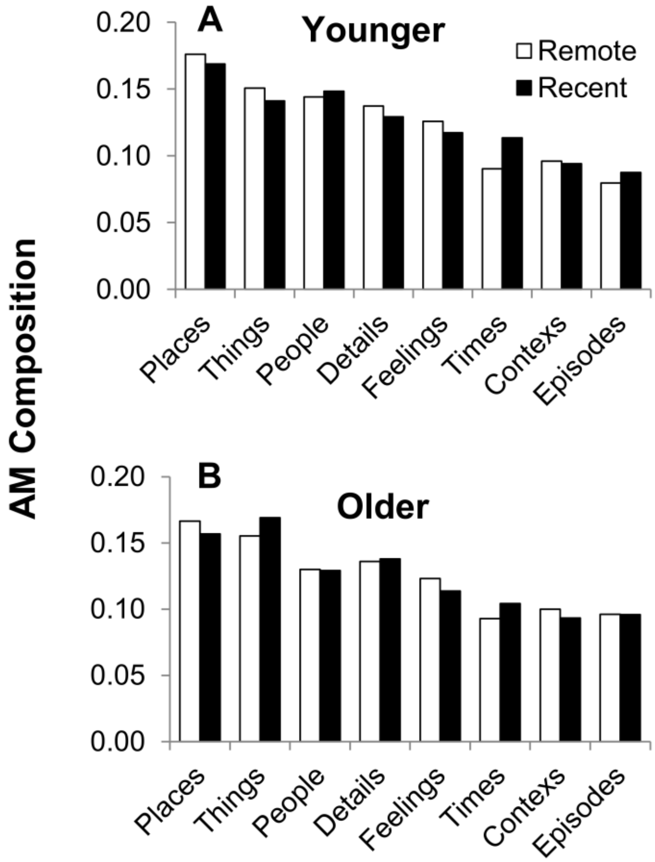


Figure S1. ***AM composition is largely equivalent across Remote and Recent periods.*** *Remote and Recent AMs were similarly composed of individual features among (A) younger (18-45 yo) and (B) older (46-78 yo) subjects; shown as proportion of total content. In addition, feature distributions of AM content among Remote and Recent intervals were quite similar to those observed overall (compare to Fig. 4 in main text).*

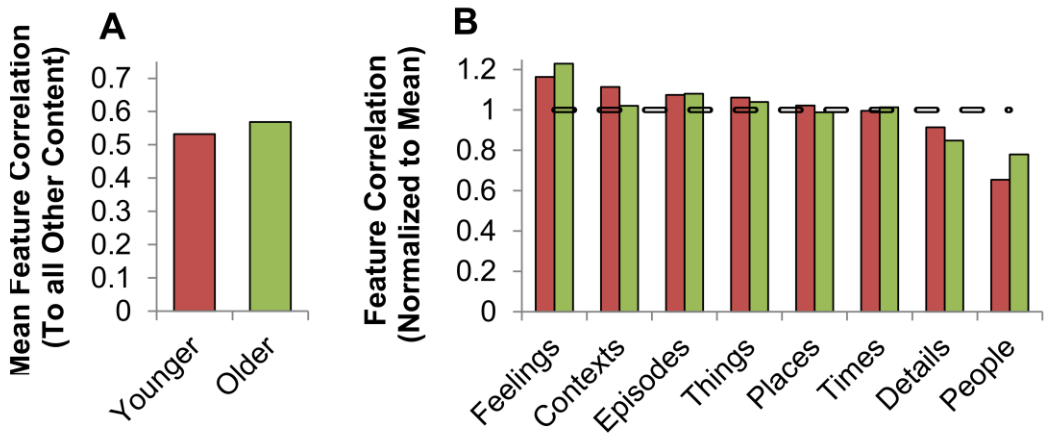


Figure S2. ***People is a relatively independent feature of recollection.***
To assess feature independence, correlation between content from a given feature to content from all remaining features was computed. (A) The mean of these values across all features was equivalent among younger (r = 0.53) and older (r = 0.57) subjects. (B) Moreover, among all ages, People appeared to be a relatively independent feature of recall (32% less than the mean); values shown for a given feature were normalized to the mean (dashed line) across features for the applicable age group. Values represent Pearson correlation coefficients. All correlations are statistically significant (p < 0.001).

CHAPTER FOUR: THE NATURAL FREQUENCY OF HUMAN PROSPECTIVE MEMORY INCREASES WITH AGE

Abstract

Autobiographical memory (AM), the recollection of past experiences, and prospective memory (PM), the prospection of future events, are prominent components of subjective life, yet data on the frequencies of their occurrence are scarce. Utilizing experience sampling, we quantified the incidence of AM and PM in natural settings among various age groups. Individuals of all ages reported engaging in AM ~10% of the time. In contrast, while younger subjects recalled PMs as often as AMs, older subjects experienced PM twice as frequently. AM occurrence was positively correlated with PM occurrence, most strongly among younger individuals. AM and PM durations were also positively correlated and remarkably stable across ages. Together, these data identify an age-associated shift in the temporal orientation of recollection, and quantify the relationship between AM and PM. More broadly, this approach provides a quantitative foundation of the occurrence of AM and PM, a crucial yet largely unexplored dimension of recollection.

Introduction

Adults readily engage in thoughts associated with the past as well as thoughts associated with the future. Recollection of personally-experienced episodes (autobiographical memory: AM) and prospection of to-be-experienced events (prospective memory: PM), together, enable an individual to cohesively recount his or her life story, and effectively plan and deliberate. Despite a recent surge of research into these temporally-distinct cognitive phenomena (e.g., see Berntsen and Rubin, 2012), empirical data of their naturalistic occurrence are scarce. Nonetheless, this knowledge is necessary to address basic yet long-standing questions about the temporality of episodic thought: Among typical adults, are some individuals more likely to ruminate past events while others are more likely to imagine future experiences? Do older individuals spend more time reminiscing their past at the expense of prospecting their future? Is there an age-associated shift in the temporal orientation of recollection?

The importance of these data is further emphasized by several studies which identify retrieval frequency as a meaningful dimension of memory. Self-reported measures of the occurrence of AM retrieval positively correlate with measures of retrieved content (Ritchie, Skowronski, Walker and Wood, 2006; Walker, Skowronski, Gibbons, Vogl and Ritchie, 2009). Likewise, experimental manipulations to enhance AM rehearsal increase a memory's perceived vividness and its associated detail (Nadel, Campbell and Ryan, 2007; Svoboda and Levine, 2009). Frequent reminiscence may also moderate an age-related loss of memory content (Cohen, 1998), an age-effect that can result in serious consequences (e.g., Cohen and Faulkner, 1989; Johnson, 1997). The frequency of PM recall is thought essential for effective planning. For example, without the aid of external reminders, the recall occurrence of a future intention is directly associated with remembering its execution (Kvavilashvili and Fisher, 2007). Moreover, imagining episodic detail associated with a future task can enhance resulting performance (Altgassen et al., 2014).

Naturalistic measurement of the occurrence of prospective recall may also clarify the prospective memory and aging paradox (e.g., Rendell and Craik, 2000). The paradox refers to a robust finding of age differences in completion rates of prospective memory tasks (i.e., remembering to perform an action at a specified temporal interval) depending on the context in which a given task is implemented; performance among younger adults is typically superior to that among older adults for tasks carried out in laboratory settings, whereas performance is superior among older adults for tasks carried out in naturalistic settings. Although questions of the mechanisms underlying this phenomenon remain open, certain explanatory factors have been proposed, including the frequency of thought related to the selected intention.

Previous studies measuring the incidence of subjective thoughts and experiences have typically employed diary-based (Linton, 1986; Wagenaar, 1986) or experience sampling techniques (Brewer, 1988; Csikszentmihalyi and Larson, 1987; Csikszentmihalyi, Larson and Prescott, 1977; also see Kahneman, Krueger, Schkade, Schwarz and Stone, 2004). Both methods ask participants to introspect, record, and often times annotate the occurrence of targeted experiences, and are well suited for use in natural settings. Diary entries are prompted by participants, demanding ongoing introspection, whereas experience sampling randomly prompts participants to evaluate the presence of a targeted experience solely at those prompted moments.

The current study, adapting prior experience sampling designs (Gardner, Vogel, Mainetti and Ascoli, 2012), relied on mobile telephony for prompting participants of ages representative of a large segment of the adult population (18-75 years old: yo; n=106) at random moments in their natural settings to document upon introspection if they were experiencing an AM or PM. If a call interrupted a memory, participants recorded its estimated duration. The proportion of sampled moments coinciding with a memory provides a direct estimation of the amount of time engaged in AM and PM. Together, data of recall probability and duration permit calculation of recall frequency.

This approach differs in several ways from previous studies that quantified the naturalistic occurrence of mental states associated with past and future temporal periods. We adopted the classical definition of AM as recollection of temporally-specific, personally-experienced past episodes. While PM, as defined here (the recollection of to-be-experienced future episodes), overlaps considerably with the recall or rehearsal of planned tasks (Kvavilashvili and Fisher, 2007; McDaniel and Einstein, 2007), it broadly encompasses the concept of future-oriented episodic thought (e.g., Addis, Wong and Schacter, 2008; Addis, Musicaro, Pan and Schacter, 2010; Atance and O'Neill, 2001; see Methods). We further emphasize that the current work measures the occurrence of these PM thoughts, which differs from studies that measured execution rates of specified intentions (or the number of distinct intentions); here, a thought of a particular future action or event was counted as a PM independent of the eventual occurrence of the episode.

Moreover, rather than probing general thoughts (Cameron, 1972; Cameron, Desai, Bahador and Dremel, 1977; D'Argembeau, Renaud and Van der Linden, 2011), including factual information (or recalled events not experienced first-hand), the current work sampled event-based, personally-relevant recollections, which appear to be affected differently by aging (e.g., Addis et al., 2008, Addis et al., 2010; Levine, Svoboda, Hay, Winocur and Moscovitch, 2002). We also emphasized to participants to solely introspect and tally the specific mental state at the moment of a given prompt. This differs from previous studies asking participants to report the most recent stream of thoughts (e.g., Klinger and Cox, 1987), which may be comprised of many divergent mental states associated with more than one temporal period. Thus, our data should estimate more selectively the proportion of time engaged in recall of past and future experiences. Inclusion of estimates of memory durations allows us to compute recall frequencies, enabling quantitative analysis of two distinct characteristics of memory occurrence: recall probability and rate.

Additionally, previous studies have typically focused on college-aged subjects or experiences of a single temporal direction (D'Argembeau et al., 2011; Gardner et al., 2012; Klinger and Cox, 1987; Kvavilashvili and Fisher, 2007; Mace, 2004; Rasmussen and Berntsen, 2011; Schlagman and Kvavilashvili, 2008; Schlagman, Kliegel, Schulz and Kvavilashvili, 2009). Important and novel aspects of the current work are the inclusion of subjects that represent a substantial portion of the adult life span and the focus on both

AM and PM as targeted cognitive phenomena within the same study. This design enables direct comparison of AM and PM occurrence within and across young and old subjects, a requisite methodology to determine how the frequencies of these memory types co-vary. Altogether, the current work characterizes the occurrence of AM and PM across the life span, providing a quantitative base of previously uncharted dimensions of human recollection.

Methods

Experience Sampling

One hundred and six participants (18-75 yo, *M*=34.7, *SD*=18.4; 75% female) were enrolled in the experience sampling experiment from the George Mason University student body (58%), faculty and staff (25%), and from the Northern Virginia community (17%). Within the community, recruitment advertisements were posted in local newspapers as well as at lifelong learning institutes. Undergraduates completed the study for course credit. All subjects reported to be in good health without memory impairment, to live independently, and to own and regularly use a mobile telephone.

This experiment relied on mobile telephony for prompting participants at random moments during their normal daily activities to document upon introspection if they were experiencing an AM or PM at the time of a given call (Gardner et al., 2012). During an initial meeting with a researcher, participants were instructed verbally on the meaning of AM and PM with the aid of script hand-outs (included as SI). AM was defined as the recollection of an episode from the personal past specific to a particular time and place, e.g., remembering your first job interview. The events recalled were emphasized to be contained in duration (less than three hours) to ensure their temporal specificity. PM was defined as the recollection of a task or event that is to occur in the personal future, e.g., bringing to mind an intention to stop at the grocery store on your way home from work. More generally, PM could include first-person perspective thinking of future actions or events, e.g., imagining the route you are going to take to arrive at the store. Before ending the initial meeting, participants were asked to apply these general classifications to various exemplar scenarios to ensure an understanding of the material. For PM classification, we did not place importance on the distinction between thoughts related to the formation of an intention, those related to the recollection of an intention, and musings of possible future events (whether it was the first time the event was called to mind or it was a repeatedly retrieved episode).

Each participant provided demographic information and his or her mobile phone number to receive random prompt calls across days of participation. A custom automatic-dialing program, utilizing a stochastic calling algorithm, was designed for use with a computer modem to initiate the calls. The total number and temporal window of daily calls were predetermined by each participant to optimize their reception and minimize disruption of expected sleep patterns or other inappropriate time periods (e.g., classes). To avoid excessive customization, participants could only select continuous calling windows, which could only differ between weekdays and weekends. The precise timing of each call was unknown to the subjects. Participants were encouraged to assign the incoming number a distinct ringtone to identify each call more quickly; alternatively, caller ID was used.

Altogether, subjects received 219 ± 96 calls (mean ± *SD*; range: 54 - 439) over the course of 19 ± 7 days (mean ± *SD*; range: 7 - 46) during their typical daily activities (between 6am and 12am). Each time prompted, participants were instructed to evaluate the concurrence of that call with an AM or PM. It was emphasized that participants evaluate only the ongoing mental state at that precise moment (that is, the thought actually interrupted by the call), as opposed to thoughts preceding, elicited or otherwise subsequent to the prompt. If participants detected the occurrence of AM or PM, they documented the specific type (AM/PM) and its estimated duration in a pocket-sized data booklet provided at the initial meeting. Alternatively, a mark was recorded to indicate the absence of a memory at the time of that particular call. To facilitate classification decisions of sampled moments, the definitions and selected examples of AM and PM provided during the initial meeting were appended to the back of each data booklet. Duration estimations were subjective and instructed to be between one and sixty seconds. This restriction was applied to facilitate estimation of a unique episodic thought as opposed to a series of thoughts. On the occasion that a memory was perceived to be longer than sixty seconds, participants were asked to estimate the duration of the most recent unique event that comprised the thought. All protocols were approved by the George Mason University Human Subject Review Board. Informed consent was obtained prior to participation.

The probability of recollection was calculated separately for each participant using Jeffreys' point estimate of probability (Jeffreys, 1961). This procedure was implemented to better estimate extreme outcomes while providing relatively small adjustments to the commonly used maximum likelihood point estimate. Specifically, the probability of recollection (*PR*) was estimated as $PR = (x + 0.5) / (n + 1)$, where x is the number of prompts that coincided with a memory and n is the total number of prompts received. The mean duration of recollection (*DR*) was calculated by doubling the mean time estimation for a given participant, assuming that, on average, the mid-point of a memory was interrupted by a prompt. The number of memories (*NM*) in a given temporal period (*TP)* was calculated for each participant as $NM_{TP} = PR \cdot TP / DR$, where TP / DR represents the total number of possible memories experienced in a temporal period. Taking the product of this value and the probability of recollection ($PR \cdot TP / DR$) captures the number of memories experienced in a specified temporal window, e.g., memories per hour (MpH). Importantly, given that memory durations are based on participant-estimations, recall durations and rates reported here are in terms of subjective (perceived) time. As such, the equation to calculate recall rates assumes that, for any one individual, the mean difference between actual and perceived memory durations does not appreciably vary according to memory type. The equation further assumes that, on average, any systematic difference between the actual and perceived duration of memories is similar to that found in other mental states. While such assumptions have yet to be tested, the results presented here provide verifiable hypotheses on the subjective rate of AM and PM recall.

All participants recorded at least one memory over the course of sampling. However, four participants (3.8% of total) did not report both an AM and PM. Thus, for these participants, duration estimates were not provided for both memory types. On such occasions, given the robust correlation between AM and PM duration (see Results: Fig. 2B,2E), the hourly recall rate for the memory type with an absent duration was calculated using the mean recall duration of the other memory type from that individual, adjusted based on the linear fit between AM and PM in the aggregate data. When applicable, this same procedure was used to assess recall rates across sampling intervals (i.e., within and across days of participation; see SI: Table S1). None of the conclusions or statistical significance of the results changed by including or excluding these computed data points.

Participant measures of recall probability, duration and hourly rate greater than three times the Inter-Quartile-Range (IQR) above the 75^{th} percentile, or less than three times the IQR below the 25^{th} percentile were considered outliers and excluded from analysis. To account for potential changes in recall across the life span, this procedure was applied separately to distinct age intervals, i.e., 18-19 (n=28), 20-29 (n=33), 30-39 (n=10), 40-49 (n=7), 50-59 (n=10), and 60-75 (n=18) years of age. Data from two participants were excluded for PM probability (and consequently PM MpH; 1.8% of total), from five for AM MpH (4.7% of total), and from three for PM MpH (2.8% of total). Two participants did not provide duration estimates of either memory type restricting their inclusion to probability analysis.

Regression analysis was used to statistically test the effect of age on each measure of recall. Subsequently, age-effects were further explored using the six age groups as described to identify outliers. Such division revealed an increase in the probability of PM recall among 50-59 yo (*M*=18%, *SD*=12%, *Mdn*=15%, Cohen's *d*=0.77) and 60-75 yo participants (*M*=22%, *SD*=18%, *Mdn*=14%, *d*=0.89) compared to younger age groups

(*M*=10%, *SD*=6%, *Mdn*=9%). Measures of AM recall probability and duration and PM duration were not meaningfully altered across age groups (see SI: Fig. S1). Thus, to describe the magnitude of age effects on recall, subjects were simply divided into two age ranges (18-49 yo: *M*=24.4, *SD*=9.1, 73% female, *n*=78; and 50-75 yo: *M*=61.7, *SD*=6.7, 79% female, *n*=28).

The younger group was predominately comprised of students enrolled at GMU (*n*=61), GMU faculty and staff (*n*=13), and individuals recruited from the local community (*n*=4). The older group was largely comprised of GMU faculty and staff (*n*=13), individuals enrolled in local life-long learning institutes (*n*=10), and those recruited from the local community (*n*=5); six subjects were retired/unemployed. Importantly, all age-effects reported here were unchanged when accounting for variation in these sampling pools. For instance, when comparing separately measures of recall across matched sampling pools (i.e., comparing younger and older GMU faculty and staff, and younger and older subjects recruited using local advertisements), older subjects showed a ~2 fold increase in the proportion of time engaged in PM (similar age-effects to that found overall), whereas engagement in AM was equivalent. Moreover, although not from the same sampling pool, the same conclusions were reached when comparing data from younger and older students (those enrolled at GMU and those enrolled at life-long learning institutes, respectively). Moreover, work status (currently working or unemployed/retired) was not associated with recall ($p > 0.10$). Together, these findings suggest that our results are not confounded by unintended or uncontrolled differences across subject pools.

Group sample sizes (younger age group: *n*=78; older age group: *n*=28) were reasoned to be sufficient based on one-hundred random samplings of twenty-eight younger participants which produced comparable measures of recall probability (AM: *M*=10.1%, 95% CI [9.9%,10.3%]; PM: *M*=10.3%, 95% CI [10.2%,10.5%]), duration (AM: *M*=32.1s, 95% CI [31.6s,32.6s]; PM: *M*=27.5s, 95% CI [26.9s,27.9s]), and rate (AM MpH: *M*=13.4, 95% CI [13.0,13.8]; PM MpH: *M*=16.9, 95% CI [16.5,17.2]) to those reflecting recall across all younger participants (Fig. 1). In particular, upon comparison of each of the one-hundred samplings to the overall data, we found that between 0% and 2% (*M*=0.8%) were statistically different, depending on memory type and recall dimension (Mean Cohen's *d*: AM probability: 0.13, PM probability: 0.12, AM duration: 0.11, PM duration: 0.14, AM MpH: 0.13, PM MpH: 0.11). This analysis is not definitive for recall probability, given the difference in variability between younger (AM *CV*=0.63; PM *CV*=0.59) and older (AM *CV*=0.88; PM *CV*=0.76) subjects. However, the consistency of the probability of AM recall across all ages, paired with the magnitude of the effect of age on the probability of PM recall (*d*=0.87; see Results) and the difference between AM and PM recall probability within older individuals (*d*=1.00), more than overcomes any underestimation of the sufficient number of older subjects.

Altogether, these factors supported the decision to terminate data collection for inclusion in the current work made subsequent to regularly performed data analysis, which invariably yielded equivalent conclusions based on relative group comparisons to those reported here. Moreover, we emphasize that the reliability of within-subject assessment is also ensured by the large numbers of recorded probes and days of participation per subject. In comparison, for example, using a similar sampling procedure, Klinger and Cox (1987) reported an average of 49 probes per subject, ~3 times fewer than the mean number of prompts recorded in our older age group.

Younger and older participants received calls over an equivalent number of days (younger: *M*=18, *SD*=6; older: *M*=21, *SD*=9). In addition, although younger subjects typically elected to receive calls slightly later in the day, the temporal window of daily prompts was quite similar across groups. On average, younger subjects received daily calls from 11:30am until 9pm, while older subjects received calls from 10:15am until 7:45pm. In contrast, older participants generally elected to receive a fewer number of total calls (younger: *M*=247, *SD*=89; older: *M*=142, *SD*=68). However, there was not a significant relationship between these sampling variables (e.g., total days of participation, duration of daily calling window, total calls received, and calls received per day) and measures of recall in either age group (see Results; also see Table S1). Thus, while age-associated variation in the sampling schedule is present, it does not confound the findings of this work.

Questionnaire

A separate sample of ninety eight college-aged subjects (18-36 years old; *M*=21.0; *SD*=4.0; 60% female) was enrolled from the George Mason University student body. No participants reported memory impairment. Subjects provided self-reported estimates of memory recall to asses if measures collected using experience sampling were expected and reproducible by questionnaire. Participants were instructed on AM and PM classifications as described for experience sampling. Subsequently, participants were given a questionnaire that collected estimations of the proportion of time engaged in AM and PM, and typical recall durations and hourly rates (the text of the questions is provided as SI). In addition to collecting direct estimates of typical hourly recall rates, we also derived these values from the recall probabilities and durations reported in the same questionnaire ("derived memories per hour" or dMpH), using the same computation as described for the experience sampling data. All protocols were approved by the George Mason University Human Subject Review Board. Informed consent was obtained prior to participation.

Values greater than three times the IQR above the 75th percentile, or less than three times the IQR below the 25th percentile for a given variable were considered outliers and excluded from data analysis. This resulted in removal of data from three participants for PM duration (and consequently PM dMpH; 3.1% of total), from three for AM MpH (3.1% of total), from six for AM dMpH (6.1% of total), and from an additional two for PM dMpH (2.0% of total). The remaining data were compared to those observed in the younger experience sampling age group (*M*=24.4 yo, *SD*=9.1, 73% female); results were equivalent when using a more precisely age-matched experience sampling population (18-36 yo; not reported).

Statistical procedures

Bivariate regression was performed to evaluate correlation between measures of recall and participant age, and between AM and PM. Independent and paired samples *t*-tests were conducted to assess differences between AM and PM recall within and across age groups, and across genders, sampling intervals, and collection methods. In cases of non-normal distributions, the Mann-Whitney U test and the Wilcoxon sign-ranked test were applied. To estimate effect size, Cohen's *d* was calculated for each group comparison. Statistical significance was interpreted using the criterion of $p < 0.05$. Statistically significant correlations were corroborated using data re-sampling procedures (Bishara and Hittner, 2012). False discovery rate correction was applied to control for spurious findings due to multiple comparisons (Benjamini and Hochberg, 1995). Statistical analyses were run using SPSS (IBM), Excel (Microsoft), and R (Dalgaard, 2008).

Results

The naturalistic occurrence of prospective memory increases with age

Each subject received on average ~220 prompts over the course of ~3 weeks (see Methods). Overall, participants experienced either an AM or PM ~23% of the time (Table 1; Fig. S2: see SI). Moreover, participant age was associated with an increase in the occurrence probability of PM ($r = 0.40$, $p < 0.001$), but not AM recall ($p > 0.10$).

Table* 1. *Naturalistic measurements of recall probability, duration, and hourly rate. *The mean, standard deviation (SD), median, and inter-quartile range (IQR) of AM and PM recall probability, duration and rate displayed here are from data collapsed across all ages (n is the number of subjects). Distributions of each recall dimension are plotted in Fig. S2.*

	Total (AM and PM)			AM			PM		
	Prob.	Dur.(s)	MpH	Prob.	Dur.(s)	MpH	Prob.	Dur.(s)	MpH
Mean	0.23	29.62	32.94	0.10	32.03	13.70	0.13	27.70	20.29
SD	0.14	16.33	27.02	0.07	18.86	14.40	0.11	15.38	18.25
Median	0.19	26.72	27.81	0.08	28.89	8.92	0.10	26.27	15.27
IQR	0.15	24.83	21.01	0.10	28.01	9.59	0.09	21.25	16.57
n	104	104	96	106	101	99	104	103	99

To further evaluate changes in recall across the life span, participants were separated into six age groups, i.e., 18-19 (*n*=28), 20-29 (*n*=33), 30-39 (*n*=10), 40-49 (*n*=7), 50-59 (*n*=10), and 60-75 (*n*=18) years of age. Based on these divisions, we observed an increase in the probability of PM recall among 50-59 yo (*M*=18%, *SD*=12%, *Mdn*=15%, *d*=0.77) and 60-75 yo participants (*M*=22%, *SD*=18%, *Mdn*=14%, *d*=0.89) compared to younger age groups (*M*=10%, *SD*=6%, *Mdn*=9%), whereas measures of AM recall probability and memory durations did not meaningfully differ across ages (Fig. S1). Thus, for subsequent description of the magnitude of age-effects on recall, subjects were separated into two age bins: younger (*Mean Age*=24 yo, *SD*=9, range: 18-49, *n*=78) and older (*Mean Age*=62 yo, *SD*=7, range: 50-75, *n*=28; see Methods). Younger subjects engaged in AM and PM equally often (~10% of the time; Fig. 1A). In contrast, older subjects were more than twice as likely to experience a PM at the time of a prompt (*M*=21%, *Mdn*=14%), both when compared to younger participants ($p < 0.05$, *d*=0.87) and to the likelihood of experiencing an AM ($p < 0.001$, *d*=1.00; Fig. 1A).

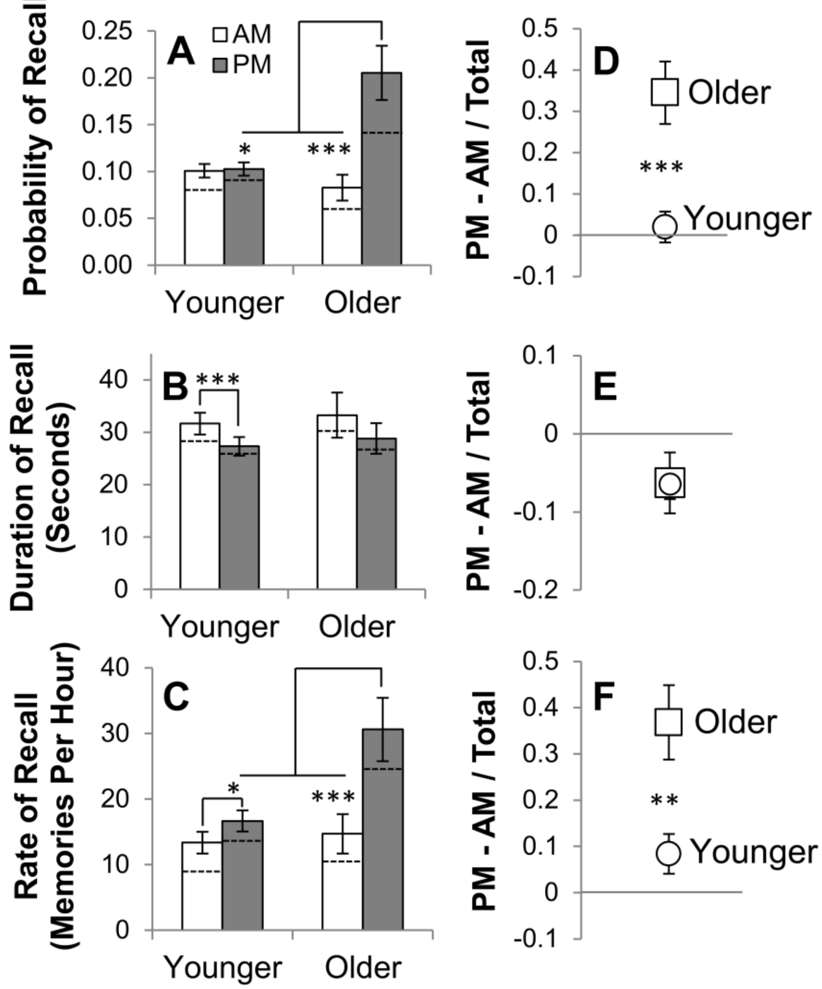


Figure 1. ***The probability and rate of prospection is increased in older adults.***
*(A) While the probability of AM recall was equivalent to the probability of PM recall in younger subjects (Mean Age=24 yo, SD=9), the probability of PM recall was enhanced in older subjects (Mean Age=62 yo, SD=7). (B) AMs were estimated to last slightly longer than PMs, showing no effects of age. (C) In general, the hourly rate of PM recall was higher than the rate of AM recall. However, this effect was strongest in the older age group, which experienced significantly more PMs. (D-F) These findings are further documented by the intra-subject difference between PM and AM recall. Data in all panels are presented as mean ± SEM. Dashed lines indicate median values. *p < 0.05; **p < 0.01; ***p < 0.001. Younger participants: AM Probability: n=78, PM Probability: n=76, AM Duration: n=77, PM Duration: n=76, AM MpH: n=72, PM MpH: n=73; Older participants: AM Probability: n=28, PM Probability: n=28, AM Duration: n=24, PM Duration: n=27, AM MpH: n=27, PM MpH: n=26.*

Each memory, on average, was perceived to have started ~15s prior to a prompt, corresponding to an estimated subjective duration of ~30s (Table 1; Fig. S2; see Methods). While recall durations were stable across age groups, the duration of AMs was mildly increased relative to PMs (younger group: $p < 0.001$, d=0.26; older group: p = 0.05, d=0.24; Fig. 1B). The within-individual mean and standard deviation of memory duration were positively correlated across all ages (AM: $r = 0.81$, $p < 0.001$; PM: r = 0.71, $p < 0.001$) yielding a stable coefficient of variation (younger group: AM CV=0.57, PM CV=0.62; older group: AM CV=0.57, PM CV=0.66). Additionally, recall probability and duration were largely independent measures, as suggested by weak correlations across memory types and age groups (younger: AM: R^2=0.08, PM: R^2=0.04; older: AM: R^2=0.01, PM: $R^2=10^{-6}$). Measures of recall probability and duration permitted estimation of the number of memories experienced in a given temporal period (see Methods). The number of prospective, but not autobiographical, memories experienced per hour (MpH) was positively correlated with participant age (PM: $r = 0.40$, $p < 0.001$; AM: $p > 0.10$). During any one subjective hour, younger subjects experienced slightly more PMs (M=16.6, Mdn=13.9) than AMs (M=13.3, Mdn=8.9; $p < 0.05$, d=0.24; Fig. 1C). Older subjects displayed an equivalent rate of AMs ($p > 0.10$), but their hourly rate of PM recall was considerably higher (M=30.6, Mdn=24.6; $p < 0.05$, d=0.70). In particular, one PM was estimated to occur every ~4 minutes in younger participants, compared to every ~2 minutes in older participants. Analyzing the relative differences between AM and PM on an individual basis also revealed these same age effects on recall (Fig. 1D-F).

Recollection is stable across a variety of conditions

While considerable inter- and intra-subject variability was observed for all measures (Table 1; Fig. S2), recall occurrence and duration were stable between genders ($p > 0.10$; Table S1: see SI). Moreover, results were equivalent when measures were computed separately for the first and second half of sampled moments, either within a given day ($p > 0.10$) or across the length of participation ($p > 0.10$). Similarly, there were no significant differences between weekdays and weekends ($p > 0.10$). In addition, neither the number of daily calls, nor the duration of the daily calling window affected recall measures in younger or older individuals ($p > 0.10$).

Measures of AM and PM are positively correlated; a relationship selectively altered with age

As AM and PM were sampled within the same subject, we further assessed how these cognitive phenomena co-vary (Fig. 2). Notably, measures of AM were strongly and positively correlated with measures of PM (probability: $r = 0.58$, $p < 0.001$; duration: $r = 0.85$, $p < 0.001$; MpH: $r = 0.57$, $p < 0.001$) in the young to mid-life participants. Moreover, the statistical significance of these results was unaltered when removing potential leverage points, as in Fig. 2C. Likewise, individuals with longer-lasting AMs also experienced longer-lasting PMs. The positive correlation between AM and PM duration was fully conserved in older subjects ($r = 0.85$, $p < 0.001$, Fig. 2E). However, the relationship between AM and PM occurrence was substantially weaker in the older group (probability: $p = 0.06$, $R^2 = 0.14$, Fig. 2D; MpH: $p > 0.10$, $R^2 = 0.08$, Fig. 2F).

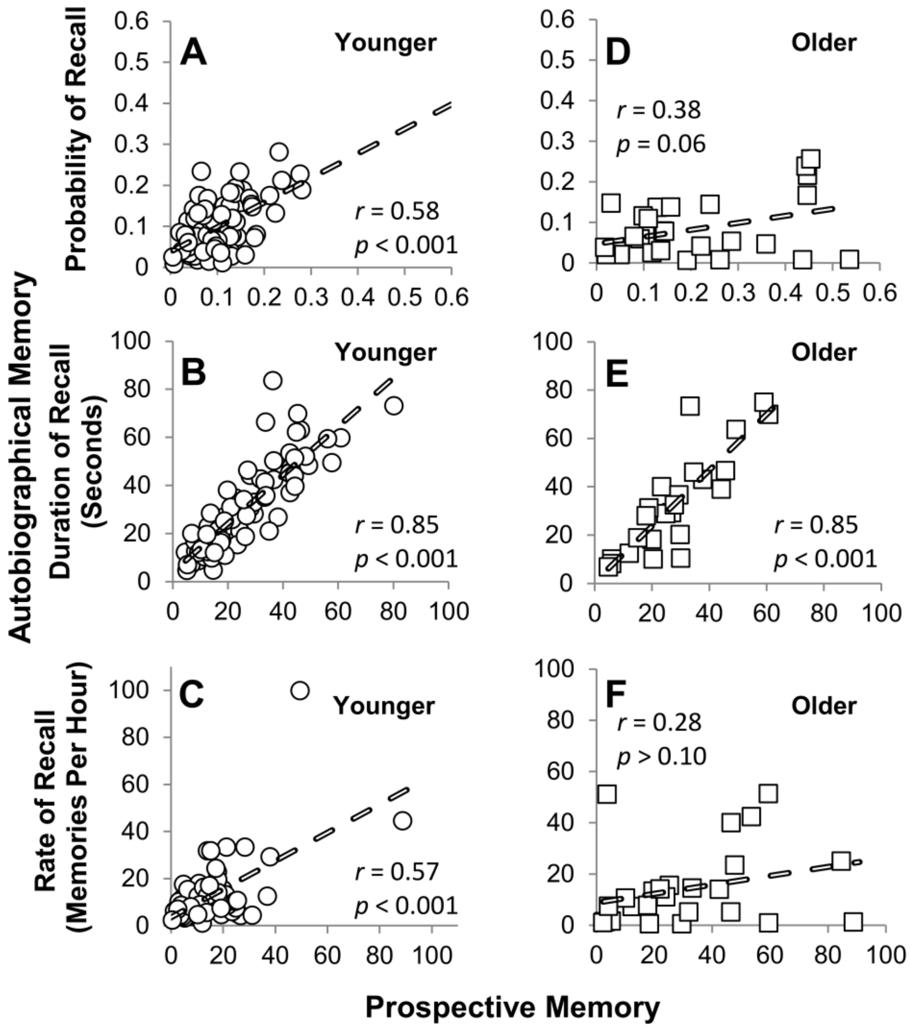


Figure 2. ***Measures of autobiographical and prospective recollection are positively correlated.***

AM and PM recall probabilities (A), durations (B), and hourly rates (C) were strongly and positively correlated in younger subjects (Mean Age=24 yo, SD=9). In older subjects (Mean Age=62 yo, SD=7), while a similar correspondence was observed between AM and PM durations (E), the relationship between AM and PM recall probabilities (D) and between AM and PM recall rates (F) was markedly weaker. Dashed lines are best linear fits. Younger participants: Probability: n=76, Duration: n=76, MpH: n=70; Older participants: Probability: n=28, Duration: n=24, MpH: n=26.

The appearance of "too high" hourly recall rates is consistent with bias detected by self-reports

Upon first inspection, both authors as well as the subjects who requested their results found the MpH rates computed from experience sampling measurements (Table 1; Fig. 1C) surprisingly high. To investigate this observation more extensively, we collected estimated assessments of recollection by questionnaire from a non-overlapping sample of young adults (*Mean Age*=21 yo, *SD*=4, range: 18-36, *n*=98). The goal was to determine the extent to which each of the experience sampling measures could be expected by intuitive introspection (Fig. S3: see SI). Questionnaire estimates of the proportion of time engaged in AM and PM were equivalent to those measured with experience sampling ($p > 0.10$; Fig. S3A). In contrast, estimated values of memory durations were significantly lower: about two-thirds of those reported during normal life ($p < 0.001$, d=0.56; Fig. S3B).

Together, these questionnaire estimates of recall probability and duration allow computation of a derived memory per hour rate ("dMpH"). Whereas the autobiographical dMpH (*M*=20.2, *Mdn*=10) was slightly elevated compared to that observed naturalistically ($p > 0.10$, d=0.37), the prospective dMpH (*M*=36.0, *Mdn*=24) was more than two-fold greater ($p < 0.01$, d=0.66; Fig. S3C). In striking contrast, when gauging MpH directly, the same questionnaire participants estimated that they typically experience 3.7 (*Mdn*=2) AMs and 5.4 (*Mdn*=4) PMs each hour. These values are ~31% of those observed by experience sampling (MpH; AM: $p < 0.001$, d=0.94; PM: $p < 0.001$, d=1.08), and ~16% of the dMpH values computed (for the same participants) from their questionnaire estimates of recall probability and duration (AM: $p < 0.001$, d=1.04; PM: p

< 0.001, d=1.09; Fig. S3C). Questionnaire estimates of AM and PM recollection were positively correlated (probability: $r = 0.53$, $p < 0.001$; duration: $r = 0.45$, $p < 0.001$; MpH: $r = 0.37$, $p < 0.001$; dMpH: $r = 0.36$, $p < 0.01$). Altogether, given the consistency of questionnaire estimates and naturalistic observations of recall probability, and inconsistency of rate estimates, the simplest explanation suggests that MpH values are underestimated by intuitive assessment. The striking incongruence between self-reported recall probability, duration, and rate may reveal the quantitative extent of this underlying cognitive bias.

Discussion

This research utilized experience sampling to measure the probability of AM and PM recall in younger (Mean Age=24 yo, SD=9) and older (Mean Age=62 yo, SD=7) subjects. Additionally, we collected real-time subjective estimations of memory durations. Together, these measures permitted computation of recall rates. In this work, AM refers to recollection of temporally-specific, personally-experienced past episodes; PM refers to recollection of to-be-experienced episodes, and includes first-person perspective thinking of future actions or events (Atance and O'Neill, 2001; see Methods).Thus, the resulting data provide a quantitative account, based on naturalistic measurements, of the everyday occurrence of past- and future-oriented episodic recollection among younger and older adults.

Younger participants engaged in AM and PM equally often (10% of the time). This result is consistent with prior studies showing equivalent proportions of general thoughts associated with past and future temporal periods (Klinger and Cox, 1987;

however also see Cameron et al., 1977). Extending this finding, we showed that, as the duration of a typical AM was slightly longer than a typical PM, the mean PM recall rate (~17/hour) was greater than the mean AM recall rate (~13/hour), in younger subjects. Although these observed recall rates appear high, as corroborated by comparison with self-reports, self-reported recall probabilities were equivalent to those observed naturalistically. Taken together with questionnaire-based estimates of memory duration that were lower than those estimated in real-time, we revealed a drastic internal incongruity between direct self-reports of MpH and the dMpH derived from recall probability and duration estimates in the same questionnaire. Altogether, these findings demonstrate the unreliability of intuitive assessment of recall occurrence that potentially explains why the naturalistic rate observations appear surprisingly high.

Naturalistic measures of AM and PM were strongly and positively correlated. Thus, while it might be expected that some people are more future-planning and others are more past-mulling, these data do not support this notion. Rather, they suggest that some individuals experience more of both AM and PM, while others spend less time engaging in memories in both temporal directions. This result is consistent with the hypothesis that past and future episodic thought cooperate, e.g., to support planning and decision making, at least in younger adults.

While the probability of AM recall in older participants was equivalent to that in younger subjects, older adults engaged in PM twice as frequently (~21% of the time). Moreover, the relationship between AM and PM occurrence in older subjects was substantially weaker than in younger participants. Thus, although in younger adults the probability of AM recall appears to be a good indicator of the probability of PM recall, and vice versa, the predictive power between memory types drastically decreases with age.

These observations are intriguing considering the changes in the temporal extents of one's personal past and future brought about by age. Specifically, the temporal period comprising the past expands with aging, while that comprising the perceived future shrinks. Thus, a priori, we might expect that older adults spend more time reminiscing past episodes at the expense of prospecting their future. On the contrary, our results suggest that while there does appear to be an age-related shift in the temporality of episodic recollection, that shift favors the future. Laboratory-based observations suggest that older adults recall PMs relatively restricted to the immediate future and AMs from a wider temporal range (Spreng and Levine, 2006). Therefore, we surmise that the foreseeable future of a typical older adult is comparatively dense with planned experiences, while their past is sparsely represented by autobiographical episodes. Research showing that older adults report fewer cued AMs compared to younger adults (Schlagman et al., 2009) also leaves open the possibility that recollection in older individuals is comprised of fewer unique, though well-rehearsed, memories. The fact that older (relative to younger) adults have a larger pool of retrievable past experiences, taken together with our finding of an equivalent AM recall rate across age groups, also suggests that on average the retrieval rate per AM is continuously lowered with increasing age; these data are in agreement with the notion that retrieval frequency may play an important role in age-related changes in episodic remembering (e.g., Cohen and Faulkner, 1989; Levine et al., 2002; Johnson, 1997).

What are the factors underlying the observed effects of age on memory frequency? In addition to potential age-related growth in the total number of planned occurrences, older adults may place a deliberate and comparatively greater focus on future events to ensure that intended actions are completed in a timely manner, which could result from (among other factors) task-related anxiety (Cockburn and Smith, 1994). This view relates to the hypothesis that older adults utilize external reminders more frequently than younger individuals (Cavanaugh, Grady and Perlmutter, 1983; however also see Phillips, Henry and Martin, 2008).

Alternatively, or additionally, our findings may result from age-related changes in the occurrence of mind wandering. Mind wandering refers to the engagement in stimulus-independent thought unrelated to the task at hand, and is a prominent feature of subjective experience (Killingsworth and Gilbert, 2010). Several findings suggest that thoughts associated with these wandering states are typically future oriented (Baird, Smallwood and Schooler, 2011; Smallwood, Nind and O'Connor, 2009), and linked to autobiographical planning (Baird et al., 2011; Cohen, 2013; Stawarczyk, Majerus, Maj, Van der Linden and D'Argembeau, 2011). Age effects on the occurrence of mind wandering have been found predominantly in studies carried out in the laboratory or those using retrospective self-reports (Jackson and Balota, 2012; Giambra, 1979, 1989). In particular, these studies suggest that the occurrence of off-task thought is decreased in older subjects. Given these results, we should expect that differential engagement in mind wandering across the life span does not explain our results; whether these previous findings translate to natural settings, however, largely remains an open question. As task demand is thought to regulate the frequency of task-unrelated thought (Smallwood and Schooler, 2006; Smallwood et al., 2009), if older and younger adults differentially engage in practiced, habitual, or cognitively demanding activities in natural settings, a proportional adjustment in their rates of mind wandering would be predicted. In line with this hypothesis, older (compared to younger) adults have reported experiencing less stress related to daily events (Schnitzspahn, Ihle, Henry, Rendell and Kliegel, 2011), and were more likely to be engaged in automatic activities during recall of a planned action (although they reported higher levels of concentration associated with those activities; Kvavilashvili and Fisher, 2007). We suggest that naturalistic measurement of the frequencies of PM thoughts (associated with unique in addition to rehearsed future experiences) as they relate to task focus and activity engagement, among younger and older subjects, is a valuable future endeavor that should clarify age-dependent changes in prospective recollection.

Our work may be relevant to studies of the aging and PM paradox, a finding that performance of prospective memory tasks (i.e., remembering to perform an action at a specified temporal interval) is generally impaired in older compared to younger subjects in laboratory settings but superior in natural settings (e.g., Rendell and Craik, 2000). Several factors have gained traction as to account for these findings, e.g., age- and setting-dependent changes in task motivation and the difficulty level of everyday activities (Aberle, Rendell, Rose, McDaniel and Kliegel, 2010; Kvavilashvili and Fisher, 2007; Niedźwieńska and Barzykowski, 2012; Schnitzspahn et al., 2011). Adding to this research, the current work suggests that the frequency at which older adults engage in future-oriented thought may also contribute to their enhanced naturalistic PM task performance. These factors (e.g., the degree of focus on future events, demand of ongoing activity, and task motivation) are not mutually exclusive and likely interact to regulate completion rates of prospective memory tasks. Future testing to tease apart their individual influences on the timely execution of intentions in natural settings is warranted.

Previous accounts of the occurrence of AM or PM are scarce and variable. Using the self-report diary technique, several studies have concentrated on the occurrence of involuntary AM recollection (memories that pop into one's mind without conscious effort) predominantly in young subjects. Reported averages of the frequency of these memories have varied from just a few in a typical day (Mace, 2004; Schlagman and Kvavilashvili, 2008; Schlagman et al., 2009) to more than ~20 (Rasmussen and Berntsen, 2011). In addition, Schlagman et al. (2009) suggested that the involuntary AM recall rate was moderately decreased among older adults. Rasmussen and Berntsen (2011) sampled both involuntary and voluntary (those deliberately retrieved) AMs, reporting that ~29 memories in total were retrieved in a single day. Using a diary study, D'Argembeau et al. (2011) estimated ~59 general future thoughts (not necessarily tied to a particular event) occurred daily. Kvavilashvili and Fisher (2007) found that for a single PM task, subjects reported on average ~2 recollections of the intention each day; a relatively high proportion of these thoughts occurred temporally proximal to the intention's execution.

While it is evident that the current study's estimates of retrieval frequencies are considerably higher than those provided by the diary method, making precise rate comparisons across studies is difficult, particularly because these previous designs did not include temporal estimates of daily participation (and were varied in their focus). Further confounding meaningful comparison, as our measures of memory duration were based on participant time estimation, values representing recall duration and consequently rate are in terms of perceived time. It should also be noted that our design sampled thoughts during participant-defined intervals. If these intervals were conducive to recollection, measures of memory occurrence would be overestimated. Any selection effect should be moderated, however, given that calling windows were continuous, did not vary from day to day, and, on average, covered a large portion of the waking day (from 10am until 9pm).

Other potential methodological distinctions that could account for a portion of this variability include the demand placed on participant introspection to document the occurrence of targeted thoughts, and how each of these thoughts is segmented. In particular, diary methods require subjects to continually monitor their thoughts in order to capture all targeted experiences, conceivably leading to some degree of underreporting. Intriguingly, in a lab-based study using random thought probes to measure the incidence of mind wandering, Schooler, Reichle and Halpern (2004) demonstrated that subjects were unaware of their engagement in these wanderings ~13% of the time, indicating intermittent lapses in meta-awareness; it is unclear, however, if these sampled thoughts would have passed entirely without being reported. It is also plausible that interrupted thoughts that are later revisited may be reported as a single tally in a diary (likely depending on the temporal delay). In contrast, upon using random prompts to sample momentary subjective experiences (e.g., as employed in our protocol), these same thoughts are more likely to contribute to a computed occurrence rate as separate and distinct events.

Considering the variability among estimates of the prevalence of subjective thought (cf. Cameron et al., 1977; Cameron, 1972; D'Argembeau et al., 2011; Gardner et al., 2012; Klinger and Cox, 1987; Kvavilashvili and Fisher, 2007; Rasmussen and Berntsen, 2011; Schlagman et al., 2009), we find it important to further emphasize several methodological choices of the current design and their implications (for a full discussion see additional text included as SI). Nonetheless, the main conclusions of this work, based on relative comparisons between memory types within and across age groups, should be robust to these choices. Thus, the reported naturalistic account of the probability, duration, and rate of AM and PM recall across the life span offers an important quantitative foundation of rarely studied dimensions of human memory.

Acknowledgements

We thank Dr. Matteo Mainetti for developing the automatic-dialing program and providing critical feedback. This research was supported in part by the Air Force Office of Scientific Research Award No: FA9550-10-1-0385, and the Office of Naval Research (ONR) Award No: 000141010198.

Supplemental Information

Table S1. *Recall is stable across genders and sampling intervals.*
Measurements of AM and PM recall (collapsed across all ages) are displayed as mean (standard deviation). [a]Defined as the 1st (Early in Day) or 2nd (Late in Day) half of sampled moments within each day of participation. [b]Defined as the 1st or 2nd half of all sampled moments across days of participation.

	AM			PM		
	Prob.	Dur.(s)	MpH	Prob.	Dur.(s)	MpH
Female, *n* = 79	0.10 (0.07)	31.49 (18.67)	14.20 (12.47)	0.14 (0.11)	27.70 (15.73)	21.50 (18.34)
Male, *n* = 27	0.08 (0.06)	33.69 (19.71)	12.29 (19.05)	0.10 (0.10)	27.70 (14.60)	16.92 (17.94)
Early in Day[a]	0.10 (0.07)	32.02 (20.54)	14.58 (18.58)	0.14 (0.11)	27.61 (15.47)	22.82 (21.47)
Late in Day[a]	0.10 (0.07)	31.99 (18.41)	14.67 (16.47)	0.12 (0.11)	28.08 (17.26)	19.71 (17.42)
1st Half[b]	0.10 (0.07)	33.21 (20.49)	14.90 (19.70)	0.14 (0.11)	28.35 (15.53)	22.16 (18.67)
2nd Half[b]	0.10 (0.07)	32.46 (20.41)	14.29 (14.20)	0.12 (0.11)	27.08 (16.99)	21.02 (20.51)
Weekdays	0.10 (0.07)	31.74 (18.86)	14.51 (18.33)	0.13 (0.11)	28.69 (16.42)	20.13 (18.28)
Weekends	0.10 (0.07)	32.02 (23.92)	16.26 (16.99)	0.14 (0.11)	26.89 (17.12)	23.26 (21.16)

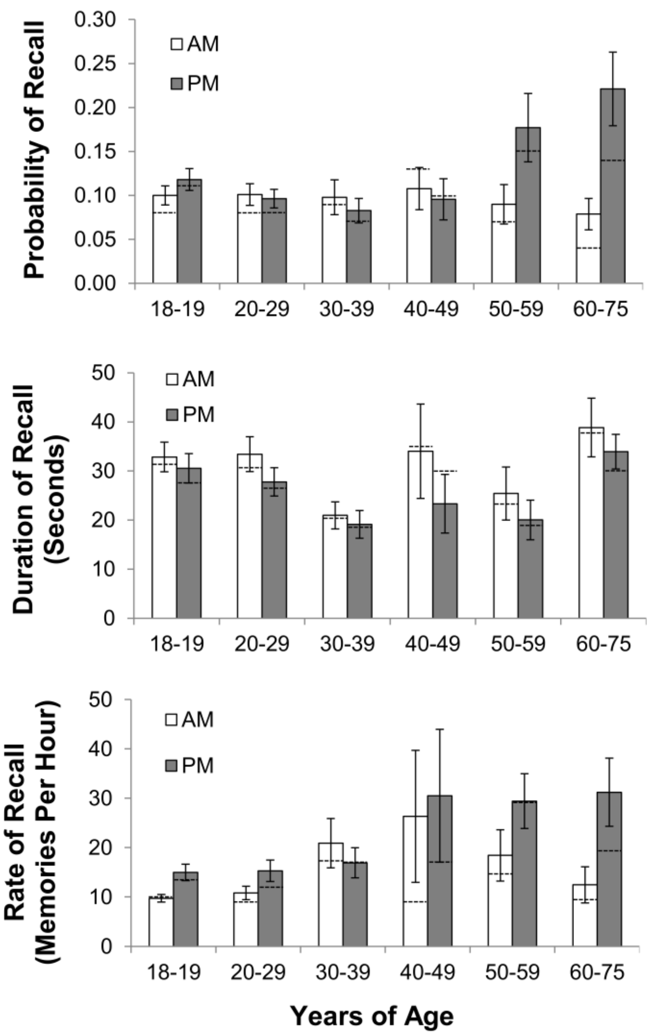


Figure S1. ***AM and PM recall probability, duration and hourly rate across age groups.*** *Recall probability, duration and rate of AM and PM are presented for each of several age ranges. Data are presented as mean ± one SEM. Dashed lines indicate median values. 18-19: n=28; 20-29: n=33; 30-39: n=10; 40-49: n=7; 50-59: n=10; 60-75: n=18.*

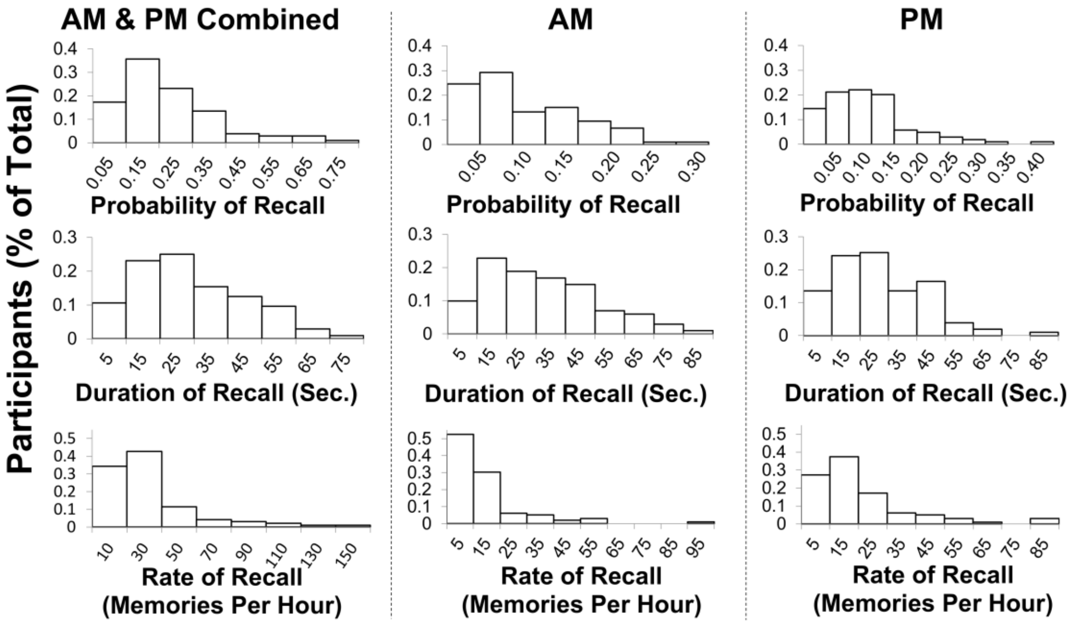


Figure S2. *Histograms of AM and PM recall probability, duration and hourly rate.* *Histograms of recall probability, duration and rate of AM and PM, presented individually and combined across memory type, feature mild to moderate right-tailed distributions. Data are collapsed across age groups.*

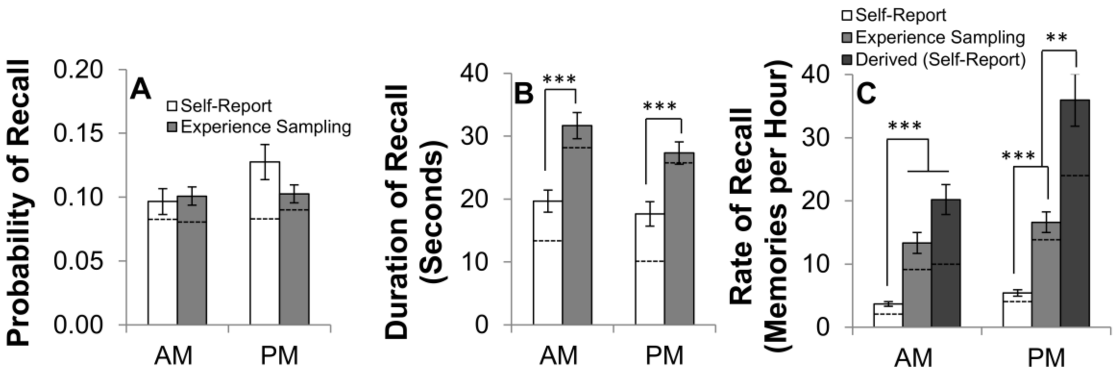


Figure S3. ***Self-reported measures of AM and PM recollection.***
*(A) Self-reported measures of the probability of AM and PM recall were equivalent to those observed by experience sampling. (B) In contrast, participants reported lower durations than those perceived naturalistically. (C) Combining self-reports of recall probability and duration, ~20 AMs and ~36 PMs should be expected to occur each hour (shown as derived MpH). Direct estimates of these values were far lower than both those calculated from experience sampling data and those derived from self-reports of recall probability and duration. Thus, while the recall rate observed in natural settings may strike as surprisingly high, this appearance may be attributed to a bias revealed by the clear inconsistency between derived and estimated MpH in the self-reports. Self-Reports: AM probability: n=98; PM probability: n=98; AM duration: n=98; PM duration: n=95; AM MpH: n=95; PM MpH: n=98; AM dMpH: n=92; PM dMpH: n=93. Experience sampling data are re-plotted from Fig. 1 (younger age group) to facilitate comparison across collection methods. Data are presented as mean ± one SEM. Dashed lines indicate median values. **p < 0.01; ***p < 0.001.*

Design considerations

Employing an equivalent sampling procedure as that used in the current work, we previously measured the occurrence and duration of AM in college-aged subjects (Gardner, Vogel, Mainetti and Ascoli, 2012). While the mean duration of AMs observed in the present study is consistent with that previously reported, the presently observed AM recall probability is lower (~10% vs. ~16%, d=0.68). It is unclear precisely what factor (or factors) contributed to the discrepancy across experiments. One glaring methodological difference is the inclusion of PM as a targeted cognitive phenomenon in the present work. It is plausible that certain memories defined as PMs (e.g., those thoughts that recalled an intention, formed in the past, to act in the future) may have been documented as AMs in the prior study (e.g., see Schlagman, Kliegel, Schulz and Kvavilashvili, 2009). Further research is needed to compare the results of these distinct designs directly, particularly when considering the variability of individual accounts employing disparate research methods of the prevalence of subjective thought (cf. Cameron, 1972; Cameron, Desai, Bahador and Dremel, 1977; D'Argembeau, Renaud and Van der Linden, 2011; Gardner et al., 2012; Klinger and Cox, 1987; Kvavilashvili and Fisher, 2007; Rasmussen and Berntsen, 2011; Schlagman et al., 2009).

Consequently, we find it prudent to highlight and discuss several methodological choices and limitations of the current research. For instance, as in earlier reports (Gardner et al., 2012), it was up to the participant to identify and classify the mental states each prompt interrupted. To minimize the inherent subjective variability of these decisions, pre-study instruction was designed to ensure that each participant had an accurate and comparable understanding of AM, PM, and the goals of the experience sampling procedure. Post-study participant debriefs of recorded memories were in agreement with these prescribed guidelines, supporting the effectiveness of this protocol (see Methods). Moreover, given that measures of recall were equivalent within and across days of participation (Table S1), it appears that subjects were able to reliably remember and apply memory classifications to sampled moments across the experiment. PM, as targeted in the current work, included both thoughts associated with goal-directed planning in addition to those which reflect future-oriented episodic musings (see Methods). Although this is a subtle distinction and there is considerable overlap among and interactions between these types of thought, they may play distinct functional roles. Further work will help determine the individual contributions of these specific thoughts to naturalistic prospection.

Values representing recall duration and rate are in terms of subjective (perceived) time, and therefore may differ from objective accounts of recollection duration. Notably, it has been suggested that time perception may change with age (Carrasco, Bernal and Redolat, 2001; Coelho et al., 2004). However, when asked to spontaneously estimate the duration of relatively short activities (a theoretically similar task to that used in the current design), younger and older adults showed equivalent time perception (Coelho et al., 2004). Moreover, participant replication of a ten second interval revealed age differences only after task repetition (Carrasco et al., 2001). These studies leave open the possibility that age-effects on subjective time might confound recall duration and rate comparisons, but average effects are likely to be mild.

Calculation of mean recall durations did not account for interruptions of recorded memories. If a memory coincided with a prompt, its estimated duration was doubled, assuming that on average a call coincided with its mid-point. In other words, this calculation assumes that a sampled memory, on average, would be engaged for the same duration following the prompt as it was preceding the prompt. As everyday distractions undoubtedly would have interrupted a proportion of these sampled thoughts following a prompted moment, to some degree our design would overestimate memory duration and underestimate memory rate. A different source of systematic error might offset this potential bias: if the intervals that participants selected to receive random prompts were conducive to recollection, to some degree measures of memory occurrence would tend to be inflated. However, any selection effect should be tempered given that the daily calling schedule was continuous, did not vary from day to day, and covered a substantial fraction of the typical waking day (see Methods).

Instructional scripts

Autobiographical Memory (AM)

Definition: An autobiographical memory is a recollection of an episode from your personal past; a memory of something that you have personally experienced in your lifetime. This memory could be of an event that occurred from the moment you were born to the most recent second you lived. The event in the memory is typically less than 3 hours and is specific to a time and place. The actual memory of the event is typically short in time (1 to 60 seconds); it is like a Kodak moment.

Autobiographical Memory Examples:

NO (Not An) AM Example: You remember a 3 hour-long trip to your grandmother's house. Specifically, you recall thinking how long the car ride felt.

Why Not? This memory would not be considered an AM because the memory is of an event that was 3 hours long.

YES (An) AM Example: You remember once on a 3 hour-long ride to your grandmother's house, you saw a cow for the first time. You can see the cow again through your mind's eye and remember what you felt (e.g., curious) and the smell of manure.

Why? This is an AM because the event that you are remembering is specific to a moment in your life. This event only lasted a few minutes. It did not stretch out over hours, days, or weeks. Also, re-experiencing the feeling and smell of the event is a sign of mental time travel, a hallmark of AMs.

YES (An) AM Example: You remember the moment your dog came home after being lost for two days. You are re-living the feeling of relief that your dog is safe and the feeling of him licking your face.

Why? This is an AM because you are re-living some portions of your event (e.g., seeing your dog come home, emotional and physical feelings). Also, this memory is of an event that happened to you personally.

NO (Not An) AM Example: You remember a story about a person that lost their dog.

Why Not? This is not an AM because it is not an event that happened to you. This is a memory of an event from someone else's life.

NO (Not An) AM Example: You remember that last week, every time you turned your head you got a shooting pain down your spine. You can feel what this pain was like.

Why Not? This is not an AM because the event is repetitive. You have experienced this event several times in your life and are not recalling a specific instance or time.

YES (An) AM Example: You remember that last week, every time you turned your head you got a shooting pain down your spine. Specifically, you recall the first time this happened, when someone called your name and you jerked your neck to see who it was.

Why? Though this is a memory of an event that has happened several times in your past, you are remembering one specific instance. It is specific to a particular time.

Prospective Memory (PM)

Definition: A prospective memory is the recollection of a task or event that is to occur in the personal future; that is, remembering an intention to perform an action in the future. The memory can be of a future task or event seconds or decades away from the present moment. The memory must also be of personal relevance. For example, the memory of the date of an upcoming event is not a PM. However, if you are recalling the date to remember to attend the event, the thought is a PM.

Prospective Memory Examples:

NO (Not A) PM Example: You recall that your friend's phone number is 703-999-6643.

Why Not? This is not a recollection of an action that is to happen in the personal future.

YES (A) PM Example: You left your cell phone at home and are now remembering to repeat your friend's number in your head so you can correctly dial it when you get to another phone later in the day.

Why? You are remembering to repeat a factual piece of information with the intention to perform an action in the future.

YES (A) PM Example: You recall earlier today when you wrote yourself a note to pick up groceries on the way home from work. The note didn't help and you forgot the groceries. As a lesson learned, you tell yourself to remember to use another method in the future, e.g., placing the note on the dashboard of your car to better remind you to stop at the store.

Why? The prospective memory is remembering to use another strategy for staying on-top of the errand.

NO (Not A) PM Example: You recall earlier today when you wrote yourself a note to pick up groceries on the way home from work. The note didn't help and you forgot the groceries.

Why Not? This is purely an autobiographical memory. No aspect of the future is being thought about.

NO (Not A) PM Example: You remember that you have an assignment due later in the week.

Why Not? You are just recalling a fact. You are not thinking about intentions to act in the future.

YES (A) PM Example: You remember you have an assignment due later in the week. However, you can't start it at the present moment. You think about what you are going to do to complete the assignment when you have the time.

Why? This situation involves thinking about an intention to perform an action in the future.

Self-report questions

Please read the following questions and estimate your answers as best you can.

Estimate how much time you spend experiencing *autobiographical* memories in one given hour. _______________

Estimate how much time you spend experiencing *prospective* memories in one given hour. _______________

On average, how many seconds would you say you spend recalling a single episode from your past. In other words, typically how long does one *autobiographical* memory last (in seconds)? ________________

On average, how many seconds would you say you spend recalling a single intention to act in your future. In other words, typically how long does one *prospective* memory last (in seconds)? ____________________

How many *autobiographical* memories would you say you experience in one given hour during your waking day? _________________

How many *prospective* memories would you say you experience in one given hour during your waking day? ________________

CHAPTER FIVE: A SECONDARY WORKING MEMORY CHALLENGE PRESERVES PRIMARY PLACE STRATEGIES DESPITE OVERTRAINING

[Gardner, R. S., Uttaro, M. R., Fleming, S. E., Suarez, D. F., Ascoli, G. A. and Dumas, T. C. (2013). A secondary working memory challenge preserves primary place strategies despite overtraining. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *20*(11), 648–656. doi:10.1101/lm.031336.113] [Used with permission from the publisher]

Abstract

Learning by repetition engages distinct cognitive strategies whose contributions to performance are adjusted with experience. Early in learning, performance relies upon flexible, attentive strategies. With extended practice, inflexible, automatic strategies emerge. This transition is thought fundamental to habit formation and applies to human and animal cognition. In the context of spatial navigation, place strategies are flexible, typically employed early in training, and rely on the spatial arrangement of landmarks to locate a goal. Response strategies are inflexible, become dominant after over-training, and utilize fixed motor sequences. While these strategies can operate independently, they have also been shown to interact. However, since previous work has focused on single-choice learning, if and how these strategies interact across sequential choices remains unclear. To test strategy interactions across sequential choices, we utilized various two-choice spatial navigation tasks administered on the Opposing T’s maze, an apparatus for rodents that permits experimental control over strategy recruitment. We found that when a second choice required spatial working memory, the transition to response navigation on the first choice was blocked. Control experiments specified this effect to the cognitive

aspects of the secondary task. In addition, response navigation, once established on a single choice, was not reversed by subsequent introduction of a secondary choice reliant on spatial working memory. These results demonstrate that performance strategies interact across choices, highlighting the sensitivity of strategy use to the cognitive demands of subsequent actions, an influence from which over-trained rigid actions may be protected.

Introduction

Whenever a subject engages in a repetitive task or behavior, with practice, several aspects of performance undergo modification. Not only does overall performance tend to improve or become more effective, but the underlying cognitive strategies change. Specifically, the formation of habits is thought to result from an incremental progression away from the use of flexible, attentive performance strategies to the engagement of inflexible, automatic strategies. This strategy transition is observed in various cognitive domains, e.g., spatial navigation (Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999; Schmitzer-Torbert, 2007), instrumental and skill learning (Balleine and O'Doherty, 2010; Derusso et al., 2010; De Kleine and Van der Lubbe, 2011), and language interpretation and execution (Ullman, 2004), and conserved across species (Packard and McGaugh, 1996; Packard, 1999; Schmitzer-Torbert, 2007). Although the transition is robust, several factors internal and external to the subject alter the onset of automatic behaviors (Restle, 1957; McDonald et al., 2004; Packard, 2009). These factors seem to exert their effects by differentially engaging dissociable learning and memory systems which interact to recruit a distinct performance strategy at select time-points during learning. Thus, an understanding of these interactions has important implications for decision-making at large.

Much of the current evidence for the nature of these interactions stems from studies of spatial navigation. In these studies, flexible/attentive strategies, termed place (locale) strategies, rely on calibrated spatial information associated with an expectancy of an outcome (Tolman, 1948; O'Keefe and Nadel, 1978); whereas inflexible/automatic strategies, termed response (taxon/praxic) strategies, rely on fixed motor sequences. During single-solution tasks, for which the use of spatial landmarks or a fixed motor sequence (but not both) are relevant to successful navigation, place and response navigational strategies show reciprocal interference and are thought to compete. For example, performance is diminished when an abundance of spatial cues (conducive to place learning) are presented during a task requiring response navigation (Packard, 1987). In dual-solution tasks, for which the use of spatial landmarks and a fixed motor sequence are both suitable for successful performance, navigational strategies can work together to coordinate an appropriate action (Hamilton et al., 2004). Typical dual solution designs include periodic probe trials to estimate the reliance on place and response navigation. Probe trials intermittently spaced across training have reliably demonstrated the strategy transition in spatial navigation, showing a greater reliance on place strategies early in training and response strategies after extensive repetition (Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999; Schmitzer-Torbert, 2007).

Previous work has predominantly focused on strategy interactions during a single isolated decision, which precludes an understanding of strategy interplay during serial-choice learning. To investigate strategy interactions across sequential decisions, using spatial navigation in rodents as a model, we designed and built the Opposing T's (OpT) maze, which permits experimental control over the strategies used on sequential intersections. We found that inclusion of a two-choice task that required attentive navigation, i.e., sustained use of spatial working memory (McDonald and White, 1993; Devan et al., 2011), at the second choice prevented the transition to response navigation at the first choice. However, this effect was not observed in two-choice control tasks. Moreover, when secondary attentive training was initiated after extensive pre-training on a single dual-solution turn, its influence on primary strategy engagement was lost, i.e., fixed motor responses persisted indefinitely. These findings suggest that performance strategies interact across serial decisions and identify conditions that modulate the influence of subsequent choices on strategy recruitment.

Results

Rats were pseudo-randomly assigned to one of four tasks (Win-Shift, n=16; Win-Stay, n=14; Win-Win, n=13; and Plus, n=20), which differed in their cognitive requirement to find a food reward. For each task, rats underwent five weeks of daily training sessions administered on a sequential two-choice apparatus (the OpT maze; Fig. 1). Win-Shift, Win-Stay, and Win-Win tasks utilized two consecutive choice points (primary and secondary; Fig. 1C). On these three tasks, the primary choice invariably utilized a dual-solution turn, i.e., a turn that could be solved by reference memory engaging either place or response navigation. The learning demands on the secondary choice varied by task (Fig. 1D). The Win-Shift task required sustained attentive navigation on the second choice, i.e., flexible actions reliant on working memory for recent spatial exploration (e.g., McDonald and White, 1993). This was accomplished by requiring a shift strategy from an immediately preceding run on which a random secondary arm was rewarded. In contrast, the secondary choice on the Win-Stay task (similar to the primary turn) could be solved by reference memory, engaging either place or response navigation. This was accomplished by invariably rewarding a secondary arm; found either by its fixed location or a fixed route. The Win-Win task demanded the same sensory-motor experiences as Win-Shift and Win-Stay tasks, but did not require secondary learning or memory, i.e., both secondary arms were rewarded. Between-task comparisons permitted isolation of the unique effects of secondary learning strategies on the primary place-to-response transition that occurs with over-training. Inclusion of a dual-solution single-choice task (Plus) provided a control condition to which the effects on primary strategy of secondary tasks could be compared more generally (Fig. 1E). An additional group of rats (P-WSh, n=14) was pre-trained on the single-choice Plus task for two weeks prior to initiation of secondary Win-Shift training (see methods for full detail on each task).

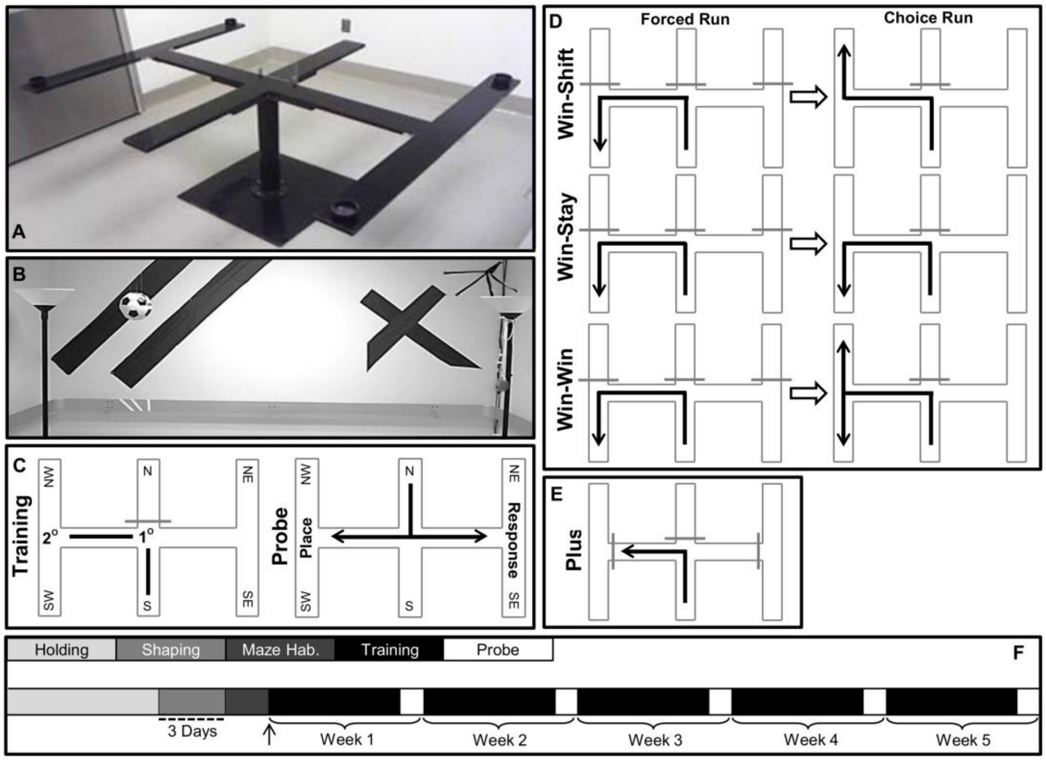


Figure 1. ***The OpT maze allows investigation of distinct navigational strategies across serial choice points.***

Training was administered on the Opposing T's (OpT) maze (A), which permitted sequential choice points (primary: 1^o; and secondary: 2^o; C) that could vary in strategic demand. The maze was positioned in a room with a moderately-rich extra-maze environment (B). Each rat was assigned to one of four tasks (Win-Shift, Win-Stay, Win-Win and Plus). During training, rats were started from the south arm (C). Win-Shift, Win-Stay, and Win-Win tasks utilized primary and secondary choice points (C). The primary choice could be solved using either a place or a response strategy. However, the strategies needed to solve the secondary choice varied by task. Each trial was comprised of two paired runs: forced followed by choice (D). On the forced run, a blockade on the secondary choice (pseudo-randomly positioned for Win-Shift and Win-Win tasks; invariably positioned for the Win-Stay task) forced the animal to enter the open (rewarded) arm. On the choice run, no blockades were present on the secondary choice and a reward was positioned according to specific task demands (D). On the choice run, the Win-Shift task rewarded entry into the secondary arm blocked on the forced run (requiring spatial working memory). The Win-Stay task rewarded the same arm as on the forced run (solvable by reference memory). The Win-Win task rewarded both secondary arms (not requiring learning or memory). The Plus task consisted of a single dual-solution turn on the primary choice point (E), and restricted access to secondary arms. On every seventh day after the start of training, a probe (five in total) was administered to identify strategy selection (C, F). Arrows indicate the location of food rewards (D-E;

Froot Loop cereal halves). The experimental timeline (F) shows the progression of each rat through the experiment prior to (i.e., Holding, Shaping, Maze Habituation) and following the onset of training (indicated by the arrow).

Primary choice accuracy and reward latency are equivalent between tasks

On the primary choice, all rats reduced arm entry errors over the course of six days of training ($F(3.67,260.35) = 52.79$, $p < 0.001$), reaching greater than 90% accuracy (Fig. 2A; Supplemental Information: SI Fig. 1A). While accuracy in movement toward the food reward was maintained (>90%) with continued training, rats on the Win-Win task displayed an increased number of errors compared to rats on the Plus condition ($p < 0.01$). However, on average only a 3% performance decrement drove this effect. Reward latency also decreased with training ($F(4.40,312.30) = 42.80$, $p < 0.001$) in a similar fashion for each task (Fig. 2B; SI Fig. 1B). Taken together, differences in secondary learning demands did not meaningfully alter performance on the primary choice.

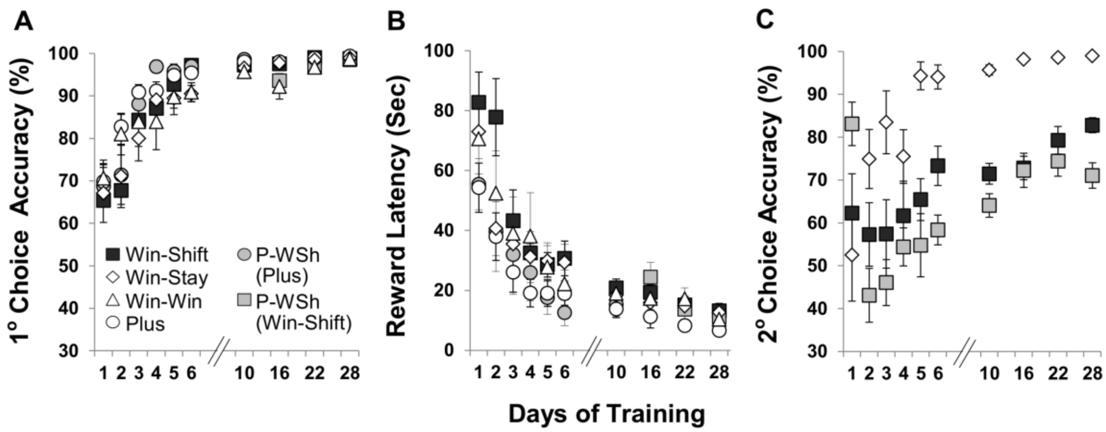


Figure 2. ***Rates of learning on primary and secondary choice points and reward latency by task.***
Task accuracy on the primary choice increased with training and reached asymptotic values after the first week (A). Across tasks, rats maintained greater than 90% accuracy on the primary choice with over-training (A). Reward latency was reduced throughout training with no task-specific effects (B). Secondary task accuracy increased with training (C). The Win-Stay task was acquired more quickly than the Win-Shift task with (P-WSh) or without Plus pre-training. While Win-Stay rats reached asymptotic performance, Win-Shift performance continued to increase with extended training with similar learning rates in Win-Shift and P-WSh conditions. Despite an equivalent learning rate, Plus pre-training significantly increased the number of Win-Shift errors across training sessions. Data are shown as mean ± SEM. After the first six days of training, values presented are averaged across each training week. The Win-Shift phase of the P-WSh group (P-WSh: Win-Shift) begins on training week three (A-B). Rats per condition: Win-Shift, n=16; Win-Stay, n=14; Win-Win, n=13; Plus, n=20; P-WSh, n=14.

Secondary choice accuracy differs according to task assignment

All rats reduced secondary choice errors over the course of training ($F(11.66,431.45) = 6.87$, $p < 0.0001$). However, significant task differences in secondary accuracy were observed ($F(2,37) = 67.12$, $p < 0.001$). Win-Stay assigned rats quickly learned the secondary choice task during the first training week (>90% accuracy), and maintained this level of accuracy during subsequent training sessions (Fig. 2C). Moreover, the learning rate demonstrated on the secondary choice mimicked that observed on the primary choice (compare Win-Stay in Fig. 2A and 2C), suggesting they were learned in parallel. Rats also performed significantly above chance on the Win-Shift task, with (P-WSh: reaching ~74% accuracy; $p < 0.0001$) or without Plus pre-training (Win-Shift: reaching ~83% task accuracy; $p < 0.0001$). At the same time, both conditions showed increased secondary errors compared to that observed in the Win-Stay condition ($p < 0.05$). Furthermore, Plus pre-training significantly increased the number of Win-Shift errors ($p < 0.05$). However, secondary Win-Shift accuracy in the P-WSh group, although reduced overall, increased with training at a rate similar to that observed in the Win-Shift group ($p > 0.10$; Fig. 2C).

Secondary task assignment differentially affects primary strategy use across training sessions

The strategy used by a rat to navigate the primary choice was identified (probed) every seventh day following the onset of training by starting the rat from the opposite (north) arm to that during training (south; Fig. 1C, 1F). On these probe runs, a place strategy was defined as entry into the side of the maze rewarded during training (indicating reliance on distal spatial cues), and a response strategy as use of the same body turn as rewarded during training (Fig. 1C; see Methods). After six days of training, each task produced a similar proportion of rats using a place strategy, i.e., ~42% ($p > 0.10$; Fig. 3A; SI Fig 2). As expected, response strategies on the dual solution Plus task increased significantly following continued training (e.g., from 55% after six days training to 90% after thirty days; $\chi^2(4) = 9.94$, $p < 0.05$; see Gardner et al., 2013: SI Video A), reflecting the transition from attentive to automatic performance. In contrast, secondary Win-Shift training maintained the number of rats using primary place strategies throughout the experiment (Fig. 3A; SI Fig. 2; see Gardner et al., 2013: SI Video B). Although the strategy transition was delayed in Win-Stay and Win-Win conditions, response strategies significantly increased by the end of training (Win-Stay: $\chi^2(4) = 11.23$, $p < 0.05$; Win-Win: $\chi^2(4) = 13.39$, $p < 0.05$), reaching comparable levels to that observed in the Plus task. Accordingly, on the final probe, significantly more Win-Shift assigned rats used a place strategy (69%) compared to those assigned to Plus (10%; $p < 0.0005$), Win-Stay (14%; $p < 0.005$) and Win-Win (15%; $p < 0.005$) conditions.

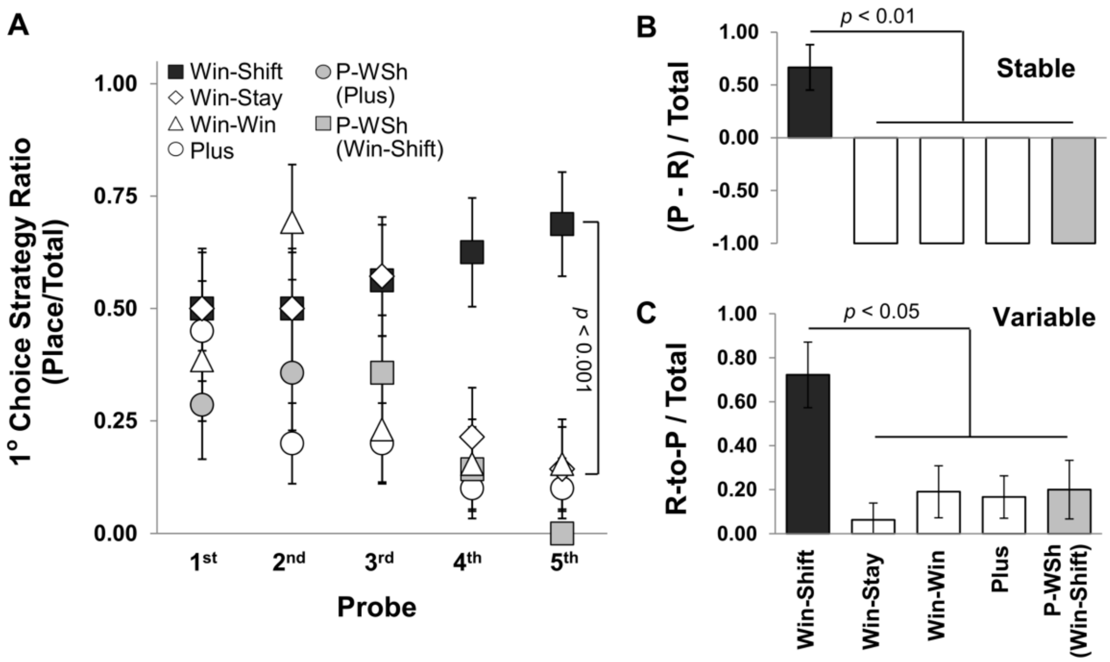


Figure 3. ***Win-Shift task assignment blocks the transition to response navigation on the primary choice.***

The incidence of primary response strategies increased with over-training in all but the Win-Shift task. This is illustrated by plotting the primary choice strategy ratio (Place Rats/Total Rats ± SEM) for each task across probe runs (A: Win-Shift, n=16; Win-Stay, n=14; Win-Win, n=13; Plus, n=20; P-WSh, n=14). Strategy data from the Win-Shift phase of the P-WSh group (P-WSh: Win-Shift) begins with probe three. These data show that Win-Shift task assignment modifies primary strategy recruitment in favor of place navigation. In addition, Plus pre-training negated Win-Shift's influence over primary strategy. Rats were divided into stable place (P), stable response (R), variable place-to-response (P-to-R), and variable response-to-place (R-to-P) strategy profiles. The proportion of rats in each strategy profile is plotted for each task (stable rats: Win-Shift, n=6; Win-Stay, n=4; Win-Win, n=2; Plus, n=5; P-WSh, n=5; variable rats: Win-Shift, n=9; Win-Stay, n=10; Win-Win, n=11; Plus, n=15; P-WSh, n=9).Variable rats were weighted according to the extent of their transition (see Methods). The ratio of rats using stable (B) to variable (C) strategies was equivalent between tasks, however more place rats and more response-to-place rats were found in the Win-Shift task. Data compiled across probe runs from the P-WSh condition (shown in Fig. 3B-C) represent those collected during the Win-Shift training phase.

These data suggest that the transition to response navigation with repeated training is prevented by the cognitive demands of the Win-Shift task, i.e., spatial working memory, rather than additional sensory-motor experiences or secondary reference memory.

To determine if secondary Win-Shift training was able to increase reliance on primary place strategies following the emergence of response navigation, a group of rats (P-WSh) was trained on the Plus task for two weeks prior to the onset of Win-Shift training. As expected, two-weeks of Plus pre-training induced relative reliance on response strategies at the primary choice; a strategy profile consistent with that observed from rats assigned to the Plus task: ($p > 0.05$; Fig. 3A; SI Fig. 2). However, Plus pre-training also blocked the effects of secondary Win-Shift training on strategy as shown by the maintained use of response navigation throughout the experiment. Consequently, like Plus, Win-Stay and Win-Win tasks, after thirty days of training, the P-WSh condition produced significantly fewer rats using a primary place strategy as compared to the Win-Shift task without Plus pre-training ($p < 0.0001$; Fig. 3A; SI Fig. 2). Moreover, this response dominant strategy profile of P-WSh rats after three weeks of secondary Win-Shift training remained stable after an additional two weeks of training (SI Fig. 2), showing no effect of extended secondary spatial working memory training on primary strategy.

To better understand the task-specific aggregate strategy data across training sessions, individual rats were grouped into two broad strategy groups: ‘stable’ and ‘variable’ (Fig. 3B-C). Rats using a stable strategy invariably displayed the same strategy on each and every probe run (Fig. 3B). The remaining rats were grouped as having a variable strategy profile, and used each strategy at least once. Stable rats were subdivided according to place or response tendencies. Variable rats were subdivided according to the direction of their strategy transition across training sessions, i.e., from place-to-response or from response-to-place (Fig. 3C; see Methods). Each task produced an equivalent proportion of stable to variable rats ($p > 0.10$). However, the Win-Shift task produced a greater number of both stable rats using a place strategy ($\chi^2(4) = 12.38$, $p < 0.01$) and variable rats showing a response-to-place strategy transition ($\chi^2(4) = 9.81$, $p < 0.05$). In addition, within-task comparisons revealed five times the number of stable place (as compared to stable response) rats, and two times the number of variable response-to-place (as compared to variable place-to-response) rats in the Win-Shift group. These findings suggest that primary place navigation observed in rats assigned to the Win-Shift task (while maintained throughout training) is predominantly the cumulative result of rats utilizing a stable place strategy and those transitioning from a response to a place strategy. Thus, although the Win-Shift primary strategy ratio (place/total) on average remains close to chance levels (Fig. 3A; SI Fig. 2), this result is not due to random arm entry on probe runs. Individual strategy groups did not differentiate primary ($p > 0.10$) or secondary ($p > 0.10$; Table 1) choice accuracy or reward latency ($p > 0.10$), showing that task performance is robust to strategy engagement (Table 1). Altogether, these data suggest that secondary spatial working memory training maintains reliance on place strategies on a primary choice despite over-training, but its effect is prevented by two weeks of previous single-choice dual-solution training.

Table 1. ***Strategy does not differentiate secondary choice accuracy.***
Secondary task accuracy (percent correct), shown as mean (standard deviation), is equivalent across strategy groups. [*]*Strategy combinations (Place,Response) observed across the five probe runs administered during secondary task training.* [†]*Non-significant (p > 0.10) Pearson's correlation coefficients show the lack of a relationship between strategy and accuracy within each secondary task.*

Task	*5,0	4,1	3,2	2,3	1,4	0,5	Total	†*r*
Win-Stay	--	95 (2)	97 (3)	93 (3)	94 (--)	95 (1)	95 (2)	0.07
Win-Shift	74 (10)	80 (2)	--	73 (6)	73 (7)	77 (7)	74 (7)	0.07
P-WSh	--	--	--	--	66 (8)	70 (--)	68 (8)	0.24

Vicarious trial and error is selectively associated with place navigation

Vicarious trial and error (VTE) is a term used to characterize back and forth movement at a choice point (Muenzinger and Gentry, 1931; Muenzinger, 1938; Tolman, 1948). As VTE is thought to be involved in deliberative spatial navigation (Papale et al., 2012; van der Meer et al., 2012), we hypothesized that VTE would be selectively increased during place navigation and in rats assigned to the Win-Shift task. We defined VTE as the number of partial arm crosses at the primary choice point prior to arm entry and measured these values on probe runs. Such measurement revealed that the proportion of place strategies used by a rat was directly correlated with the amount of VTE ($r = 0.58$, $p < 0.0001$; Fig. 4), a trend mimicked when looking at individual strategy groups ($F(3,50) = 7.54$, $p < 0.01$; Fig. 4 inset). Like strategy recruitment, the amount of VTE observed after six days of training was equivalent between tasks ($p > 0.10$). However, after thirty days of training, there was significantly more VTE observed in Win-Shift trained rats ($p < 0.05$; e.g., see Gardner et al., 2013: SI Videos A-B, compare VTE during

Plus to VTE during Win-Shift probe runs). Considering VTE on place and response runs separately for each task (as correlation analysis between VTE and strategy is confounded by task assignment), we found a robust increase (~61%) in VTE during place as compared to response navigation; although considerable variability in the magnitude of this increase was observed across tasks (Win-Shift = 66%; Win-Stay = 150%; Win-Win = 20%; Plus = 20%; P-WSh = 51%).

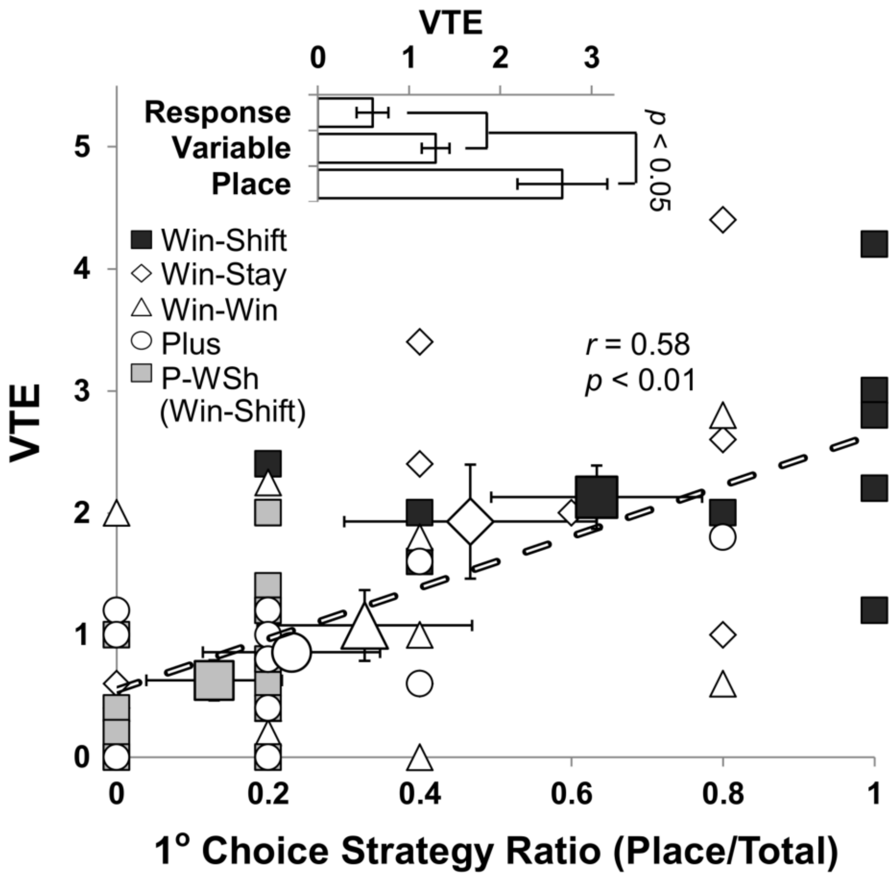


Figure 4. ***Vicarious trial and error is associated with place navigation and preserved during secondary Win-Shift training.***

Mean vicarious trial and error (VTE) was directly correlated with mean strategy ($r = 0.58$, $p < 0.01$); VTE was increased in rats that displayed relatively greater reliance on place navigation. Likewise, a greater number of VTE events was found in place compared to variable and response strategy groups (mean ± SEM; inset). Larger symbols reflect mean strategy and VTE (± SEM) across rats according to task assignment, illustrating a between-task gradient in VTE and place-based navigation (ranked in decreasing order of mean VTE: Win-Shift, n=12; Win-Stay, n=9; Win-Win, n=11; Plus, n=13; P-WSh, n=14). Data shown from the P-WSh condition represent those collected during the Win-Shift training phase.

While the scoring method we used may not be sensitive to fine head movements and does not account for pausing (Hu et al., 1997), we also measured the duration of time rats spent in distinct segments of the maze on probe runs. The results from this analysis were consistent with the results obtained from tallies of partial arm crosses, showing that the mean time spent in the primary choice-point was significantly greater on place (10.58 seconds) than on response runs (5.76 seconds; $F(1,317) = 9.34$, $p < 0.0001$); this effect was not observed on other maze segments, e.g., start or goal arms ($p > 0.10$), Together, these data demonstrate a robust correspondence between VTE and place navigation.

Discussion

This research set out to examine interactions between performance strategies during serial-choice learning. In particular, we investigated the effects of secondary tasks with specified cognitive requirements on the transition from attentive to automatic performance that occurs with repeated practice and experience. The findings reported here suggest the presence of a transfer effect of strategy during serial navigation, and outline conditions that modulate its occurrence. Specifically, we identified a strictly behavioral treatment that prevents, but does not reverse, the transition from the use of place (attentive) to response (automatic) strategies. Confirming previous research, we demonstrated that over-training on a single-choice, dual-solution task (Plus) increased the incidence of response navigation (Fig. 3A; SI Fig. 2; Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999). In contrast, we found that a two-choice task that required attentive performance (Win-Shift: spatial working memory) at the second intersection completely blocked the transition from place to response navigation at the first (primary; Fig. 3A; SI Fig. 2). In addition, control two-choice tasks that could be solved using reference memory and/or response navigation, but that overlapped considerably in many behavioral aspects (Win-Stay, Win-Win), delayed but did not prevent the onset of response strategies with over-training. Together, these results suggest that while secondary reference memory and/or additional sensory-motor experiences may delay the strategy transition, secondary spatial working memory training is able to preserve the reliance on place performance at a primary choice.

The extent and rate of secondary learning was decreased in the Win-Shift task compared to control tasks (Fig. 2C). Therefore, we cannot eliminate secondary learning and task difficulty as factors influencing primary strategy recruitment. Indeed, uncertainty has been shown to modify the use of performance strategies (Derusso et al., 2010; Sullivan et al., 2012). However, it appears that uncertainty, learning phase, and/or task difficulty on the secondary choice are not principal factors driving the effects on primary strategy observed here. In particular, although rats in the Win-Shift condition did not reach asymptotic secondary performance, they significantly increased accuracy across training sessions (Fig. 2C). Moreover, secondary task accuracy did not predict primary strategy in the aggregate (compare Win-Shift from Fig. 2C and Fig. 3A; SI Fig. 2) or individual data (Table 1). In fact, while there was a non-significant increase in primary place strategies with continued Win-Shift training, a significant transition to response navigation was observed in control tasks, although secondary performance was increased with training on all tasks.

Similarly, our experimental design did not isolate the effects of working memory for spatial locations as opposed to working memory for sequences of actions or turns. For example, although the Win-Shift task requires an animal to attend to relevant task information (Zola-Morgan and Squire, 1985), and recruits the same neural systems implicated in place/spatial strategies (McDonald and White, 1993; Devan et al., 2011; Packard et al., 1989; Packard and McGaugh, 1996; Packard, 1999), it is possible that some Win-Shift animals strictly relied on working memory for recently used routes (as opposed to spatial locations) to find the goal. Thus, it would be interesting to test the distinct contributions of secondary spatial memory and working memory (in addition to task difficulty) on the regulation of primary strategy recruitment. A design that includes additional conditions, e.g., fixed, higher-order serial tasks that vary in difficulty, and a cued non-spatial working memory task (McDonald and White, 1993), would help disentangle these factors to better understand their individual influences on strategy engagement during serial decision-making.

The interaction between attentive and automatic performance strategies across decision-points observed here complements human studies showing transfer effects of working memory training to unpracticed tasks that utilize similar brain structures and cognitive processes (Klingberg, 2010; Brehmer et al., 2012). Moreover, in agreement with working memory experiments that measure capacity (Huang-Pollock and Karalunas, 2010), these results suggest that working memory demand is inversely associated with implicit skill acquisition or automatic performance. To determine if secondary Win-Shift training directly inhibits primary response engagement and/or directly facilitates primary place engagement will require further investigation.

VTE is a marker of deliberative spatial processing in the rat. We found VTE to selectively co-occur with place-dependent learning strategies on a dual-solution turn, a finding consistent with Schmidt et al. (2013) who showed that VTE is increased during explicit place compared to response training. Additionally, VTE behaviors were strongly associated with the prevention of the place-to-response transition by secondary Win-Shift training. As the Win-Shift task requires ongoing updating and application of recent spatial information, these results support the hypothesis that VTE is more akin to an active search mechanism (Johnson et al., 2012; Papale et al., 2012), rather than a passive exploratory behavior.

Together, these findings support the notion that secondary working memory training has potential value to increase attentive performance on targeted tasks and skills despite extensive practice and experience (e.g., Di Nocera et al., 2006; Youmans and Ohlsson, 2008; Hagewoud et al., 2010; Friederich and Herzog, 2011; Gillan et al., 2011; He et al., 2011; Reichenbach et al., 2011; Hogarth et al., 2012). Consistent with this hypothesis, delay discounting has been shown to improve both with working memory training (Bickel et al., 2011) and instruction that shifts attention to focus on later rather than immediate rewards (Radu et al., 2011).

However, our data also suggest that automatic performance resulting from repeated practice may be less responsive to such treatment. The increase in rats transitioning from a response to a place strategy in the Win-Shift task presents the possibility that secondary working memory training is able to increase the use of attentive strategies after sustained reliance on fixed motor behaviors. However, we did not see this effect after two-weeks of Plus pre-training, which produced a relative reliance on response navigation (Fig. 3A; SI Fig. 2). These discrepant findings are potentially explained by methodological differences between the Win-Shift task with (P-WSh) and without Plus pre-training. Specifically, the Plus task rewards the primary turn at the end of the east or west arm. In contrast, the Win-Shift task demands primary and secondary choices, but grants a single reward at the end of the 'correct' secondary arm (see Methods). Therefore, prior learning in the Plus configuration could elicit compartmentalization of the primary and secondary choice points into discrete tasks that rely on discrete strategies, an effect that persists across Win-Shift training sessions. Alternatively, as rats pre-trained on the Plus task reached asymptotic performance and predominantly showed response strategies on the primary choice after two weeks of training, the influence of subsequent Win-Shift training on primary strategy may have been blocked by previous training and/or response interference. This notion suggests there may be a sensitive period during which secondary Win-Shift training can act upon primary strategy recruitment.

In accordance with this latter explanation, compared to Win-Shift success when training began at the onset of the experiment, secondary Win-Shift accuracy was decreased following Plus pre-training, demonstrating some degree of performance interference. This result is consistent with findings from a dual-solution task that showed reduced performance during subsequent strategy reversal training (on the same choice point) after over-training (Hicks, 1964). Extending this research, our findings suggest that strategy interference may occur across choice points, during which the use of previously utilized strategy/choice-point combinations (e.g., the use of an automatic strategy on an initial decision) impairs acquisition of a subsequent task requiring the alternate strategy (e.g., a subsequent attentive task). Further testing will help determine the degree to which primary strategy, as opposed to non-specific learning, influences secondary single-solution learning. At large, our data suggest that administration of a secondary working memory task may preserve attentive decision making, and is most effective when administered from the onset of task acquisition. That is, once a task is over-trained, even prolonged secondary working memory training may not be an effective means to modify the underlying performance strategy. The initiation of secondary Win-Shift training only varied across two time points, i.e., at the start of the experiment or following two weeks of Plus pre-training. Therefore, it is unclear if the effects of secondary spatial working memory training on primary strategy are 'all or none' or graded. For example, it is possible that as behaviors become more rigid with increased training periods, the efficacy of working memory training is gradually reduced. Identifying the nature of these effects will help resolve whether secondary working memory training may be effective at increasing the use of attentive performance strategies in specified over-trained behaviors. Similarly, this research did not test whether response navigation, as determined on probe runs, was habitual. However, previous research using a modified dual-solution procedure paired with reward devaluation showed that after ten days of over-training (40 runs/day), performance was habitual (Smith et al., 2012). Therefore, the results reported here on the malleability of the transition to response navigation may be applicable to the transition to habitual behaviors with repeated performance.

In rodents, a robust double dissociation is observed between a hippocampal-mediated place strategy and a dorsolateral striatal-mediated response strategy (Packard et al., 1989; Packard and McGaugh, 1996; Packard, 1999; Yin and Knowlton, 2004); with similar dissociations noted in humans (Corkin, 1968; Knowlton et al., 1996; Hartley et al., 2003; Squire, 2004; Bohbot et al., 2007; Balleine and O'Doherty, 2010; Banner et al., 2011; Wit et al., 2012). Moreover, several factors that modulate strategy recruitment are shown to act differentially on neural structures implicated in place and response navigation (Packard, 1987; McDonald et al., 2004; Packard, 2009). In light of these findings, and those showing that successful Win-Shift performance and VTE are dependent upon place-based brain regions (McDonald and White, 1993; Hu and Amsel, 1995; Hu et al., 1997; Hu et al., 2006; Devan et al., 2011), it is tempting to discuss the results of this research in terms of task-specific engagement of place and response neural circuits. We put forth a hypothesis that maintenance of a place strategy on the primary choice point (despite over-training) during Win-Shift training is due to specified secondary recruitment of the hippocampus and supporting place-based neural structures (e.g., dorsomedial striatum: Yin and Knowlton, 2004). As Win-Stay and Win-Win tasks produce a comparable albeit transient effect on strategy, the simplest explanation would assert that these tasks also exert their effects through a time-limited hippocampal-mediated mechanism. In support of this idea, learning of the second turn of a two-turn response-based task was disrupted in a transgenic line of mice lacking N-methyl-Daspartate (NMDA) receptor GluN1 subunits in CA1 of the hippocampus (Rondi-Reig et al., 2006); a finding consistent with the role of hippocampus in sequence learning and memory (Kesner et al., 2002). This two-turn response-based task was administered for approximately two-weeks. Therefore, it is unclear if a more extensive training period would permit secondary performance to be supplemented by an alternative neural mechanism and learning strategy. Our findings that show an increase in primary response navigation with extended training on Win-Stay and Win-Win tasks, however, support this notion.

Consistent with the hypothesis that the effects of secondary Win-Shift training on primary strategy reliance are mediated by the hippocampus, spatial training has been shown to increase gray matter selectively in the hippocampus (Maguire et al., 2006; Lerch et al., 2011), a factor directly associated with increased reliance on place strategies (Bohbot et al., 2007). Further work will need to clarify the degree to which secondary training can influence strategy reliance in subjects with a response predisposition, e.g., as a result of aging (Bohbot et al., 2012), neuroanatomical variability (Bohbot et al., 2007; Wit et al., 2012), genetic polymorphism (Banner et al., 2011), stress (Packard, 2009; Schwabe et al., 2010), and/or sleep deprivation (Hagewoud et al., 2010), rather than repeated practice and experience. Nonetheless, this research opens a promising line of future investigation. We emphasize, however, that even slight alterations to training designs can significantly modify the neural structures supporting a particular task. Therefore, behavioral results obtained using the OpT maze in conjunction with measures of neural activity, plasticity, and causality will help further our understanding of the neural regulation of and the precise relationship between attentive and automatic memory systems.

Methods

Animals

Male Long-Evans Hooded rats (Charles River Laboratories, Wilmington, MA), 275-500 grams, were pseudo-randomly assigned to one of four tasks (Win-Shift, n=16; Win-Stay, n=14; Win-Win, n=13; and Plus, n=20; Fig. 1D-E). In addition, a separate group of rats (***P-WSh***, n=14) was given two weeks of pre-training on the Plus task followed by training on the Win-Shift task. All methods were carried out in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the George Mason University Institutional Animal Care and Use Committee.

Apparatus

Tasks were administered on the OpT maze (Opposing T's; Fig. 1), an adaptation of the Plus maze (e.g., Tolman et al., 1946, see also Pol-Bodetto et al., 2011) with secondary arms attached to and bisected by the ends of the east and west arm segments of the 'plus'. Thus, it is comprised of four primary maze segments, i.e., north, east, south, west, and four secondary maze segments, i.e., northeast, northwest, southeast, and southwest (Fig. 1C), each measuring 23.5 (length) x 4.5 (width) x 0.75 (height) inches built off square choice points measuring 4.5 (length) x 0.75 (height) inches. Each arm segment permits attachment of a removable blockade (Plexiglas) and an opaque, circular food (reward) cup measuring 1.25 inches high and 2.5 inches in diameter. The maze is constructed from pine boards painted flat black and coated with a clear lacquer. It swivels 360 degrees on its cylindrical base which elevates the maze 28 inches. The testing room was environmentally-enriched (Fig. 1B) with three-dimensional objects and geometric shapes hanging on the walls and ceiling. Furthermore, lighting was provided by four floor lamps placed in the north-east, north-west, south-east, and south-west corners of the room each using thirty-watt bulbs.

Habituation

Upon import to our animal facility, rats were housed (2-3 per cage) for a minimum of seven days. Rats were then individually housed and brought to 85% of their free-feeding weight through caloric restriction over an additional seven days. Concurrent with caloric restriction, rats were handled five minutes daily (Holding in Fig. 1F; see Packard and McGaugh, 1996). Following this period, rats were habituated to Froot Loop cereal (FL, Kellogg) by placing three FL halves in their home cage. Rats were shaped daily to take and consume a FL half from a reward cup positioned at the end of a rectangular table in a room distinct from that used for training (Shaping in Fig. 1F). After consuming the FL half for at least three consecutive days and under 180 seconds on the final two days, rats were habituated to the OpT maze (Maze Hab. in Fig. 1F). On each day, for two days, maze habituation provided five minutes of maze exploration starting the rat from the south arm with the north arm blocked (e.g., Packard and McGaugh, 1996). During maze habituation, single FL halves were placed in all reward cups, with one cup positioned at each possible reward site (at the ends of the East and West arm segments for the Plus task, and at the ends of the four secondary arm segments for Win-

Shift, Win-Stay, and Win-Win tasks). The experimenter recorded the rat's first arm entry. The side recorded on day two of maze habituation was identified as the rat's arm/turn preference.

Tasks

After maze habituation, training began and proceeded daily for five weeks (except for every seventh day when a strategy probe was administered; Fig. 1F). During training, the north arm was blocked and the animal was started from the south arm (Fig. 1C). Twelve runs were given daily with a three-minute run maximum and thirty-second inter-run-interval during which the rat was kept in the holding cage behind the south arm. Throughout training, the animal was rewarded to the side of the maze (east or west) opposite its preference (e.g., Tolman et al., 1946) identified during day two of maze habituation. The reward consisted of a single FL half placed in the reward cup positioned at the end of the appropriate arm segment according to task-specific training rules. During days one and two of training, the rat could retrace its steps to find the goal after entering a non-rewarded arm (e.g., Ritchie et al., 1950). With the exception of the P-WSh condition, for which this self-correction method was also applied to the secondary arms for the first two days of Win-Shift training, on day three and onward, after entering a non-rewarded arm, the rat was removed from the maze. Between runs, the maze was pseudo-randomly rotated 180 degrees and wiped down with water to minimize use of intra-maze cues. On the first two runs of the first day of training, a trail of four FL halves led the rat from the choice point to the baited reward cup. Arm entries and reward latency were documented by the experimenter with the aid of a silent-operation stopwatch.

Win-Shift, Win-Stay, and Win-Win tasks (Fig. 1D) included two consecutive choice points (primary and secondary: Fig. 1C). The primary choice could be learned using either a place or response strategy and rely on reference memory. This primary turn was equivalent to that in the Plus task. However, to retrieve the food reward, each rat was required to navigate the secondary arms. Reward cups were placed at the end of each secondary arm. Each trial (six daily) consisted of a pair of runs (e.g., Dumas et al., 2004). One of the secondary arms was blocked on the first run leading the rat to find a reward on the open arm (forced run). On the second run, neither arm was blocked, requiring a choice to be made between the two open arms (choice run).

The learning demands of the secondary choice varied according to the assigned task. On the choice run, the Win-Shift task (e.g., McDonald and White, 1993) rewarded entry into the previously blocked arm (assigned pseudo-randomly), requiring working memory. In contrast, the Win-Win task rewarded entry into either arm and thus did not require secondary learning. For the Win-Stay task, on forced runs, the blocked arm was invariably positioned opposing the start arm; and the same previously rewarded arm (on the forced run) was again rewarded on choice runs. Thus, the Win-Stay task, by design, was a serial dual-solution task and could be solved by reference memory.

The ***Plus*** task (Fig. 1E) consisted of a single (primary) choice for which rats were rewarded when making a consistent turn from the south starting position (constants throughout training; e.g., Ritchie et al., 1950; Hicks, 1964). Thus, like the primary choice on each secondary task, the Plus task could be learned using long-term memory engaging either a place or response strategy. Entry into secondary arms was restricted by Plexiglas blockades at the ends of the east and west arms. Reward cups were placed at the ends of the east and west arms in front of the blockades.

To assess the strategies used on the primary choice across training sessions, on every seventh day a probe consisting of a single run was administered starting the animal from the arm opposite to the training start arm (Fig. 1C, 1F; Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999). A place strategy was identified if the rat entered the same arm rewarded during training. A response strategy was recorded if the rat made the same turn rewarded during training (Fig. 1C; see Gardner et al., 2013: SI Videos A-B for examples of Plus and Win-Shift training and probe runs). Five total probes were administered for Win-Shift, Win-Stay, Win-Win and Plus groups (Fig. 1F). Seven total probes were given in the P-WSh condition: two during Plus pre-training, and five during Win-Shift training.

Data analysis

Maze habituation, training, and probe sessions were videotaped for subsequent analysis. During maze habituation, training and probe runs, an arm entry was defined by the full body of the rat excluding the tail. Rats were excluded from analysis if less than 75% accurate on the primary choice or displaying a reward latency mean greater than two minutes on days prior to probe runs. This procedure excluded five rats (6% of total).

As rats were prevented from self-correcting after the initial two-day interval, task accuracy on the secondary choice for Win-Shift, Win-Stay and P-WSh conditions was calculated only if a rat was successful on the primary choice both during forced and choice runs. This procedure permitted orthogonal analyses of primary and secondary choice accuracy. Since the Win-Win task rewarded both secondary arms on choice runs, by design, rats in this condition could not make an incorrect choice and were thus excluded from secondary choice accuracy analyses.

Rats using a ‘stable’ strategy invariably displayed the same strategy on each probe run. The remaining rats were grouped as having ‘variable’ strategy use. Stable rats were subdivided according to place or response biases. Variable rats were subdivided according to the direction of their strategy transition, i.e., from place-to-response or from response-to-place. For each variable rat, this designation was determined by dummy coding place (1) and response (0) strategies and calculating the slope of the line of best fit across probes one through five. The sign of this line determined each variable rat’s group; a negative slope indicated a transition from place to response navigation and a positive slope indicated a transition from response to place navigation. In the P-WSh condition, the line of best fit was calculated from the five probe runs administered during the Win-Shift phase of training (data were equivalent when accounting for all probes given both during Plus pre-training and Win-Shift training phases). For analysis, variable data (Fig. 3C) were weighted by the absolute value of the slope of the line of best fit. This approach provided greater influence to those rats with more robust transitions (data were equivalent when using un-weighted counts of variable transitions).

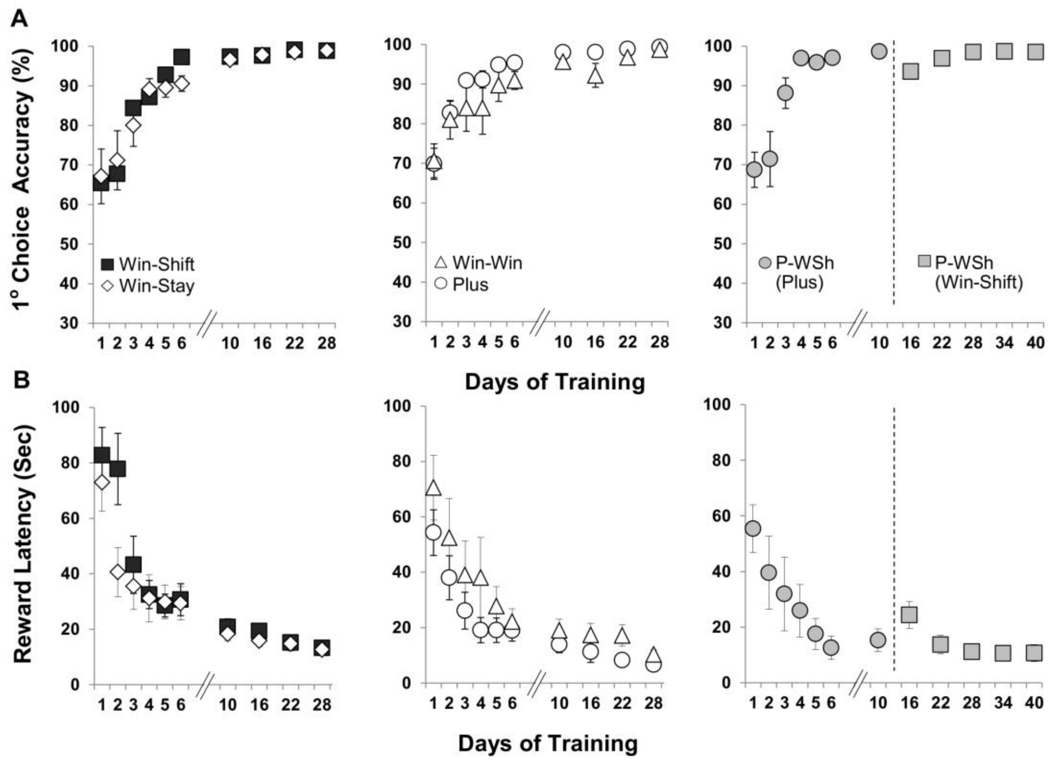
VTE was quantified on the primary choice point of probe runs using experimenter-scored analysis of recorded videos. Similar to previous studies (e.g., Hu and Amsel, 1995) measuring VTE by the number of head orientations to visual stimuli prior to action selection, we defined VTE as the number of unique partial head/body entries into the open arms surrounding the primary choice-point prior to full body entry into either the east or west arm. This value was tallied for each probe session for each rat. The duration of time spent in distinct portions of the maze on probe runs was also scored from videos by an experimenter with the aid of a stopwatch. Three regions of interest were applied to segment the maze: the starting arm, the primary choice point (defined by the central square), and the east and west arm segments. A subset of video files (~22% in total) distributed across Win-Shift, Win-Stay, Win-Win and Plus tasks was corrupted during back-up. Thus, not all probe runs were analyzed for VTE events or exploration of maze segments. However, complete sets of intact video files (i.e., video of all probe runs for a given rat) from animals representative of aggregate data in each task (Win-Shift, n=12; Win-Stay, n=9; Win-Win, n=11; Plus, n=13; P-WSh, n=14) permitted unbiased statistical comparison of VTE and arm segment exploration. In addition to arm entry designations on probe runs, all findings reliant on experimenter-scored video analysis, were corroborated through further video analysis by researchers blind to secondary task assignment and research hypotheses.

Data compiled across probe runs from the P-WSh condition (i.e., Fig. 3B-C; Fig. 4) represent those collected during the Win-Shift training phase (weeks three through seven). Unless otherwise noted, data are presented as mean ± standard error of the mean. A mixed model ANOVA was utilized to assess task, training duration, and strategy group effects on primary and secondary choice accuracy, reward latency, and VTE. Single-sample *t* tests were run to determine if secondary Win-Shift accuracy was above chance levels. Fisher's exact test, the chi-square test for homogeneity, Cochran's *Q* test and a mixed model logistical regression with binary variables, a procedure fit with the generalized estimating equation, were utilized to compare strategy use across tasks and training duration. Pearson's *r* was calculated to assess correlation between strategy and secondary choice accuracy, and strategy and VTE. Statistical significance was interpreted using the criterion of $p < 0.05$. All post-hoc tests were corrected for type-1 error inflation using the Bonferroni technique. The Greenhouse-Geisser correction was applied to violations of sphericity (for full details on these statistical tests see Davis, 2002; Sheskin, 2011).

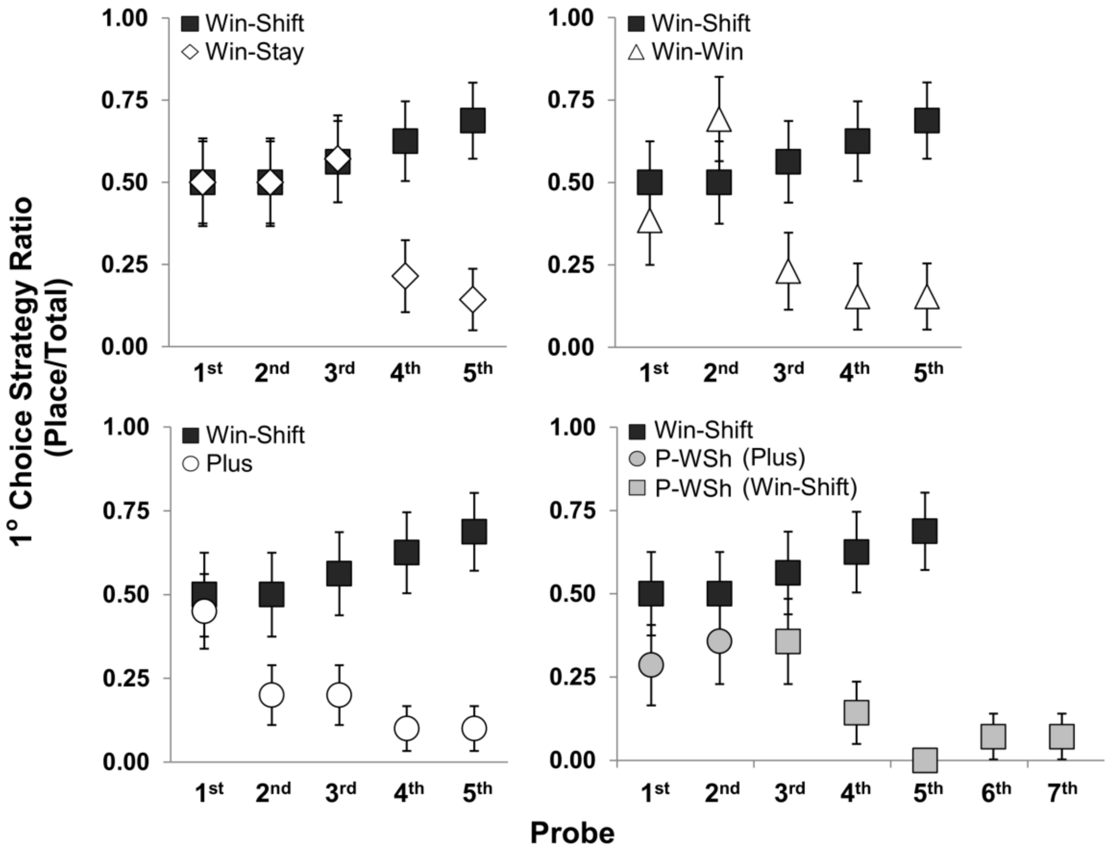
Acknowledgements

We thank Asish Gulati, Iris Talebi, Mai-Phuong Hoang, Christopher Rudnicky, and Christina Pehlivanis for their help conducting behavioral experiments and with data analysis. This research was funded by the Office of Naval Research: ONR#000141010198, the Air Force Office of Scientific Research (AFOSR): Award No. FA9550-10-1-0385, and the Krasnow Institute for Advanced Study.

Supplemental Information



SI Figure 1. ***Primary choice performance.***
Plotting primary choice accuracy (A) and reward latency (B) illustrates enhanced performance on the primary choice point with repeated training sessions (data re-presented from Fig. 2A-B). Data are shown in multiple panels to facilitate examination of task-specific effects. All data from the P-WSh group are presented (two weeks of Plus pre-training, and five weeks of Win-Shift training).



SI Figure 2. ***Strategy use on the primary choice.***
Plotting task-specific primary strategy ratios (Place/Total) across probe runs shows the prevention of the transition to response navigation in the Win-Shift group despite extensive training, an effect not observed in control tasks (data re-presented from Fig. 3A). Data are shown in multiple panels to facilitate examination of task-specific effects. Similarly, each panel includes data from the Win-Shift task to illustrate differential effects of task-assignment on primary strategy. All data from the P-WSh group are presented (two probes from the Plus pre-training phase, and five probes from the Win-Shift training phase).

APPENDIX: THE TRANSITION FROM ATTENTIVE TO AUTOMATIC PERFORMANCE CORRELATES WITH CHANGES IN HIPPOCAMPAL AND STRIATAL ARC EXPRESSION

Abstract

The strategies utilized to effectively perform a given task change with practice and experience. During a spatial navigation task, with relatively little training, performance is typically attentive enabling an individual to locate the position of a goal by relying on spatial landmarks. These (place) strategies require an intact hippocampus. With task repetition, performance becomes automatic; the same goal is reached using a fixed response or sequence of actions. These (response) strategies require an intact striatum. The current work aims to understand the activation patterns across these neural structures during the strategy transition. This was accomplished by region-specific measurement of activity-dependent immediate early gene expression among rats trained to different degrees on a dual-solution task (i.e., a task that can be solved using either place or response navigation). As expected, rats increased their reliance on response navigation with extended task experience. In addition, dorsal hippocampal expression of the immediate early gene Arc was considerably reduced in rats that used a response strategy late in training (as compared to those that used a place strategy early in training). In line with these data, vicarious trial and error, a behavior associated with hippocampal metabolism, also decreased with task repetition. Although Arc mRNA expression in

dorsal medial or lateral striatum alone did not correlate with strategy/training stage, the ratio of expression in the medial striatum to that in the lateral striatum was relatively high among rats that used a place strategy early in training compared to over-trained response rats. Altogether, these results identify specific changes in the relative activation of distinct neural systems that may underlie the experience-dependent emergence of response navigation (and/or indicate strategy recruitment). Limitations of the study are discussed.

Introduction

Upon engaging in a previously unfamiliar task, attentive performance is typically required to accomplish a desired outcome. After extensive practice, performance becomes fixed and automatic. This experience-dependent transition from the use of attentive to automatic performance strategies is largely studied in the context of spatial navigation; attentive (place) strategies rely on memory of the position of spatial landmarks to flexibly locate a goal, whereas automatic (response) strategies rely on a series of fixed movements that comprise an inflexible route. This strategy transition is readily observed in numerous species, including humans (Schmitzer-Torbert, 2007) and rodents (Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999), suggesting the value of model systems to investigate its underlying neural mechanics.

In rodents, the plus (cross) maze (e.g., Tolman et al., 1946) is a simple apparatus consisting of four arms built off a central square commonly used to study the place-to-response transition. In particular, on a dual-solution task (one that can be solved using place or response navigation; Ritchie et al., 1950; Hicks, 1964), animals are trained in a room with an enriched extra-maze environment to find a reward in a static location (e.g., the west arm) from a consistent start position (e.g., the south arm). After task acquisition, to identify which strategy is dominant at any given moment, a single trial (probe) is administered starting each animal from the opposite position to that used during training. The new position puts at odds the route associated with each navigational strategy and thus permits simple identification of the primary mode of performance.

Using this dual-solution plus maze design coupled with reversible neural inactivation techniques, Packard and McGaugh (1996) demonstrated a double dissociation between the expression of place and response strategies and their neural correlates. The expression of place navigation required the dorsal hippocampus and not the dorsolateral striatum, and the expression of response navigation required the dorsolateral striatum and not the hippocampus (see also Packard, 1999; Packard, Hirsh and White, 1989). Further studies showed that the dorsal striatum was functionally heterogeneous suggesting that the medial region was required for flexible spatial navigation and the lateral region for fixed response navigation (Yin and Knowlton, 2004).

How the activation of these distinct navigational systems relates to strategy recruitment at different stages of training, however, remains relatively unexplored. Is hippocampal activation highest early in training when attentive strategies dominate? Does activity within the dorsolateral striatum rise only after extensive training coinciding with the emergence of automatic strategies? Is strategy engagement predicted by the relative activation across neural structures?

The current work begins to address these questions by measuring the expression of the immediate early gene (IEG) Arc/Arg 3.1, a proposed marker of neural activity (e.g., Guzowski, McNaughton, Barnes and Worley, 1999; Pinaud and Tremere, 2006; Vazdarjanova et al., 2006). This was done in the hippocampus and striatum of rats trained for either a brief or a protracted schedule on a dual-solution task. Thus, this design permits evaluation of the relative activation across neural structures as it relates to strategy engagement at distinct time points of training. Complementary to this approach, we examined experience-dependent changes in vicarious trial and error (VTE). VTE refers to the tendency for rats to pause at a choice point, and look back and forth toward potential destinations (Muenzinger and Gentry, 1931; Muenzinger, 1938; Tolman, 1948). As VTE is associated with deliberation (Papale et al. 2012; Gardner et al, 2013; van der Meer et al. 2012; Schmidt et al., 2013) and correlated with hippocampal activation (Hu, Xu and Gonzalez-Lima, 2006), this behavioral measure may further clarify the degree to which attentive systems are recruited across task repetition. Although the data presented here are preliminary, in part, due to a small number of samples for which Arc expression was quantified (see Methods and Discussion), they provide verifiable hypotheses on the temporal dynamics of neural activation that underlies the experience-dependent emergence of response navigation.

Methods

Animals

Sixty-four male Long-Evans rats (275-500 grams) were used for experimentation. These rats were bred in-house at George Mason University (*n*=47) or ordered from

Charles River Laboratories (Wilmington, MA; *n*=17: ~175 grams at the time of arrival). Until beginning the experiment, animals were housed two to three per cage. All methods were carried out in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the George Mason University Institutional Animal Care and Use Committee.

Apparatus

Rats were trained to find a food reward (Froot Loop cereal; Kellog) on a maze positioned in a room with a rich heterogeneous extra-maze environment (Fig. 1). We used a previously described multi-choice maze (the Opposing T's: OpT maze) set in a plus maze configuration (see Gardner et al., 2013; minor modifications are detailed). Briefly, four arm segments (north, east, south, and west) built off a central square (choice point) were utilized. The south and north arms were potential starting positions, with identical opaque start boxes affixed to the ends of each arm. Reward cups were placed at the ends of the east and west arms. On any given run, three of the four arms were accessible, which restricted the maze to a "T" shape. The south, east, and west arms were open during training; the north, east, and west arms were open during probe trials. To limit falls during maze runs without restricting visual access to the extra-maze environment, clear Plexiglas railings were attached at the ends of each of the four arm segments.

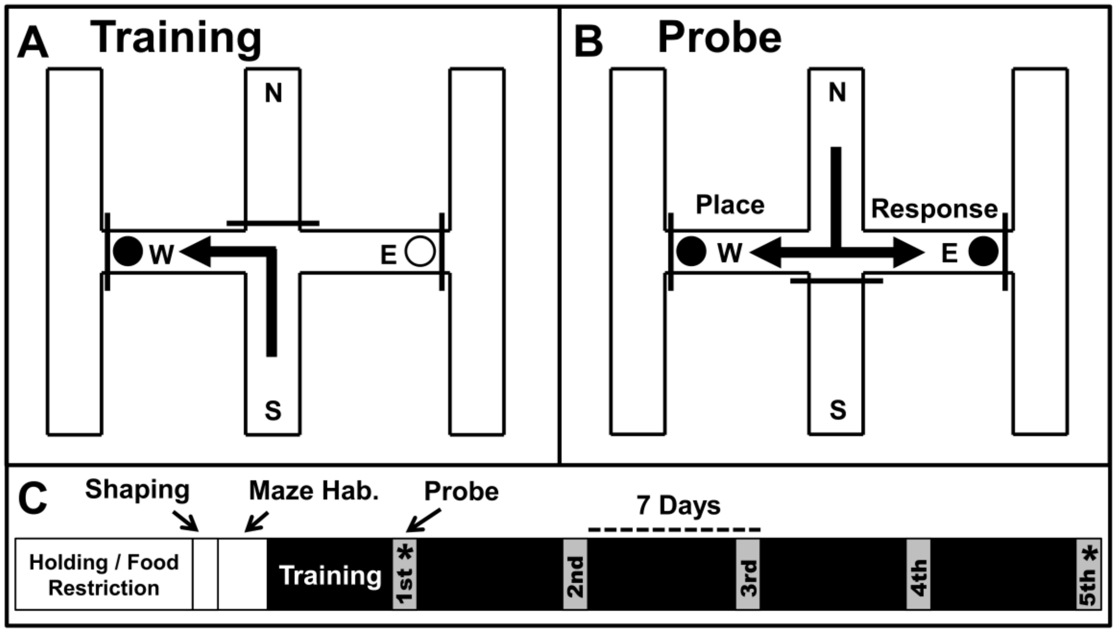


Figure 1. ***Behavioral testing and experimental time line.***
*(A) Rats were trained from a consistent starting position (e.g., south) to find a food reward (e.g., west; filled circles indicate rewarded food cups) in a room with a heterogeneous extra-maze environment. (B) To probe the dominant strategy (place or response) at various training stages, the rat was started from the opposite (e.g., north) arm. A place strategy was identified as entry into the arm previously rewarded (e.g., west). A response strategy was identified as the use of the previously rewarded turn (e.g., left). (C) The experimental time line is illustrated. Prior to training, rats were food-restricted, habituated to the experimenter, shaped to approach and consume Froot Loop cereal from reward cups, and given two days to explore the maze. To identify the changing reliance on place and response navigation with task repetition, a probe trial was administered after task acquisition, and subsequently every seventh day. *Rats were sacrificed five minutes after the first or fifth probe trial and processed for in situ hybridization targeting the immediate early gene Arc (see Methods for full detail).*

Behavioral training

Food restriction and habituation were largely implemented as previously described (Gardner et al., 2013). Rats were individually-housed and maintained at 85% of their free-feeding weight throughout the experiment. While rats were brought down to their target weight, they were handled for five minutes each day. After being held a minimum of seven days and after meeting their target weight, rats were habituated to Froot Loop (FL) cereal (Fig. 1C: shaping) and given five minutes to explore the maze for each of two days (Fig. 1C: maze habituation).

The day after maze habituation, a one-time pre-training trial was administered after which training commenced. During these trials, an animal was placed in the south start box. After ~10 seconds, the front door to the box was remotely raised, using a pulley, providing the rat access to the maze. If a rat did not exit the start box after 180 seconds, the experimenter placed the animal on the maze directly in front of the box and closed its door to restrict re-entry. During all trials, the experimenter stood ~3 feet behind the south arm. The pre-training trial was unrewarded (no food was placed on the maze). On this trial, the arm first entered (with the full body excluding the tail) determined an animal's turn preference, and the opposite arm/turn was rewarded (with half of a piece of FL cereal) on all subsequent training trials for a given subject. This procedure ensured that all animals explored both arms at least once over the course of the experiment.

Following the pre-training trial, and during all daily training sessions thereafter, six training trials were administered with an inter-trial-interval of ~45 seconds. Rats were started from a consistent arm (south arm) and removed once they found the food reward. For any given rat, as the locations of the food and starting position were unchanged across and within days of training, the task could be solved relying on landmarks to find the position of the food (e.g., going west) or a fixed body turn (e.g., turning left; Fig. 1).

Rats were allowed to self-correct and re-trace their steps to find the food reward. The time to find the reward and the first arm entered were recorded. Correct trials were identified as those during which the rat entered the rewarded arm and reached the FL reward without entering the unrewarded arm. Training sessions proceeded daily, for a minimum of four days, and until a rat correctly navigated the maze on all six daily trials of the most recent training day and at least ten of the twelve trials over the most recent two days.

Probing strategy dominance

The day after meeting these performance criteria, a single probe trial was administered in place of training (Fig. 1). On the probe, all conditions were identical to training except that the rat was started from the north arm (the opposite of that used for training) and both the east and west arms were rewarded (as neither choice was "wrong"). If an animal attended to environmental landmarks and entered the same arm as that rewarded during training, it was documented to have used a place strategy. If the rat used the same turn rewarded during training (and thus entered the opposite arm), it was documented to have used a response strategy (Fig. 1B). The day following the probe trial, training re-commenced for an additional four weeks with additional probe trials given every seventh day (in lieu of training on those days; Fig. 1C). All training and probe trials were videotaped for analysis. Vicarious trial and error (VTE) was quantified on each probe run by summing the number of times a rat paused at the decision point of the maze and the number of times it partially entered the east and west arms prior to making a full entry (e.g., Gardner et al., 2013).

Tissue collection

A subset of rats was sacrificed after the first probe trial (n=37) and the remaining (n=27) animals were sacrificed following the fifth probe trial (Fig. 1C). Specifically, five minutes after the end of the applicable probe, the rat was taken to a surgery room and immediately anesthetized with isofluorane, and perfused intra-cardially with ice-cold 4% paraformaldehyde mixed in DEPC-treated 1X phosphate-buffered saline solution. This time line was used to capture strong nuclear Arc mRNA expression among neuronal ensembles active during maze exposure (Guzowski et al., 1999; Guzowski et al., 2006). Following perfusion, each brain was excised and placed in the paraformaldehyde solution overnight at 4°C. Brains were then submerged in a 30% sucrose solution mixed in DEPC-treated water for a minimum of four days.

Brain sectioning

To assess changes in Arc mRNA expression in dorsal hippocampus and dorsal striatum during the transition from place to response navigation, brains from four rats that used a place strategy on their first probe and brains from four rats that used a response strategy on their fifth probe were processed for in situ hybridization (ISH). Brains were sectioned (30 microns) in the coronal plane at -25°C. Sections were mounted on slides, air-dried overnight and subsequently stored at -70°C. For each brain, four sections were collected on each of at least four slides (Colorfrost Plus Slides: Fisher Scientific). To facilitate comparison across brain regions within a given animal, each slide contained two sections of dorsal striatum (between 1.0mm and -1.0mm from bregma; Fig. 2B) and two sections from dorsal hippocampus (between -2.5mm and -4.5mm from bregma; Fig. 2A).

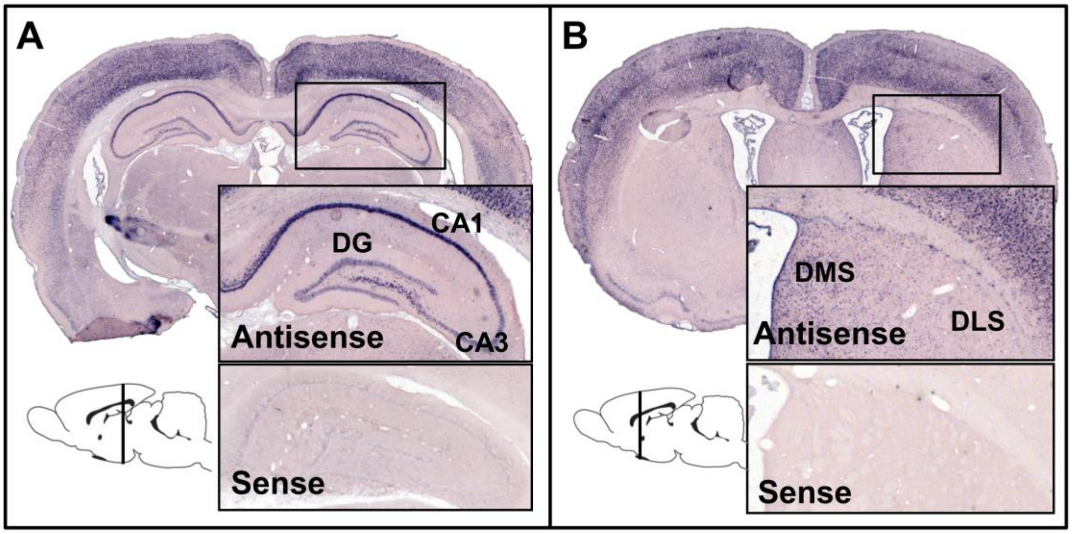


Figure 2. ***In situ hybridization targeting the immediate early gene Arc.***
Arc mRNA expression was measured from tissue sections of (A) dorsal hippocampus (between -2.5mm and -4.5mm from bregma) and (B) dorsal striatum (between 1.0mm and -1.0mm from bregma). Typical colorimetric ISH (relying on the AP-BCIP/NBT system; see Methods) results are presented for antisense and sense riboprobe application and together demonstrate the specificity of the experimental technique. Neuronal sub-fields measured include CA1, CA3 and dentate gyrus (DG) of the hippocampus and medial (DMS) and lateral (DLS) striatum.

Colorimetric ISH targeting the mRNA transcribed by the immediate early gene Arc/Arg 3.1 was performed using full-length digoxigenin-labeled riboprobes, detected using the AP-BCIP/NBT system (Kessler et al., 1990; Wehr et al., 2009).

Probe generation

The rat Arc DNA sequence inserted into a pBluescript plasmid was maintained in colonies of Escherichia coli. Plasmids were isolated and linearized with restriction enzymes (New England Biolabs) XhoI (for subsequent synthesis of sense probe) or EcoRI (for antisense probe). The effectiveness of the restriction digests were verified using gel electrophoresis. Full-length riboprobes were synthesized using T3 (sense) or T7 (antisense) RNA polymerase (Roche), with digoxigenin-labelled UTPs (DIG RNA labelling kit: Roche). Briefly, a mixture of fully cut DNA template, RNA polymerase, RNase inhibitor, and a mix of DIG-labelled NTPs, was incubated at 37°C for two hours. To degrade the DNA template, a fifteen minute incubation of this mixture in the presence of DNase I (Roche) followed. Synthesized RNA probe was extracted using spin columns (OmegaBiotek), verified using gel electrophoresis, and stored at -70°C.

In situ hybridization

For an ISH experiment, probe was diluted (1:40) and denatured for ten minutes in hot (70°C) hybridization buffer (Sigma). This solution was added to each thawed section (~250µl/slide), which was surrounded by a lipid layer (Pap Pen Liquid Blocker: Ted Pella) to prevent loss of solution. Subsequently, cover glass (Fisher) was placed atop each slide. Slides were incubated in a humidified chamber (with 1X sodium chloride sodium citrate, 50% formamide) overnight at 62°C. Slides were washed in 1X sodium chloride sodium citrate, 50% formamide for 15 minutes at 62°C after which the coverslips were removed. An additional three 30 minutes washes followed. Slides were transferred to maleic acid buffer, 0.1% Tween 20 and washed three times for 30 minutes each at room temperature. Blocking solution (1X maleic acid buffer, 20% sheep serum, 20% blocking reagent: Roche) was added to each slide (250 µl/slide), which was placed in a humidified chamber (with 1X phosphate buffered saline solution) for one hour. This mixture was then discarded after which to each slide anti-digoxigenin antibody conjugated to alkaline phosphatase (Roche) diluted (1:1500) in blocking solution was added (~250 µl/slide).

Slides were re-placed in the humidified chamber and incubated overnight at room temperature. Four five minutes washes in maleic acid buffer and two ten minute washes in alkaline phosphatase staining buffer (0.1M Sodium Chloride, 0.05M Magnesium Chloride, 0.1M Tris pH 9.5, 0.1% Tween 20) followed. Slides were incubated for four hours at 37°C in a mixture of 10% (w/v) high molecular-weight polyvinyl alcohol, 8% Levamisole Solution (Vector Laboratories), 3.5% NBT (Roche), 2.6% BCIP (Roche) made up in alkaline phosphatase staining buffer. The colorimetric reaction was stopped by washing slides in phosphate-buffered saline solution (0.1% Tween 20; two ten minute washes) and deionized water (two ten minute washes). Slides were dehydrated in a graded series of ethanol dilutions (70%-100%), cleared in xylenes (Fisher), and permanently cover slipped. All pre-hybridization solutions were DEPC-treated (RPI corp) and all working surfaces and instruments were maintained RNase free. Each ISH experiment included all comparative conditions. Unless stated otherwise, reagents were obtained from Sigma (St. Louis, MO).

Image analysis

Images of sections were collected using bright-field microscopy (Olympus AX70) under 2X magnification, and saved as TIF files for analysis using imageJ (http://imagej.nih.gov/ij). Images of sections incubated with sense riboprobe were used to confirm the specificity of the probe to bind Arc mRNA, and generally showed little staining relative to those incubated with antisense probe (Fig. 2A-B). ISH experiments for which slides incubated with sense probe displayed equivalent staining to those incubated with anti-sense probe (indicating non-selective binding) were excluded form analysis.

Arc mRNA expression, quantified as the mean grayscale intensity corrected for background, was measured from images of sections incubated with antisense riboprobe. In particular, expression was measured in the pyramidal cell layer of CA1 and CA3, in the granule cell layer of dentate gyrus (DG), and dorsolateral and dorsomedial striatum. Regions of interest are presented in Figure 3.

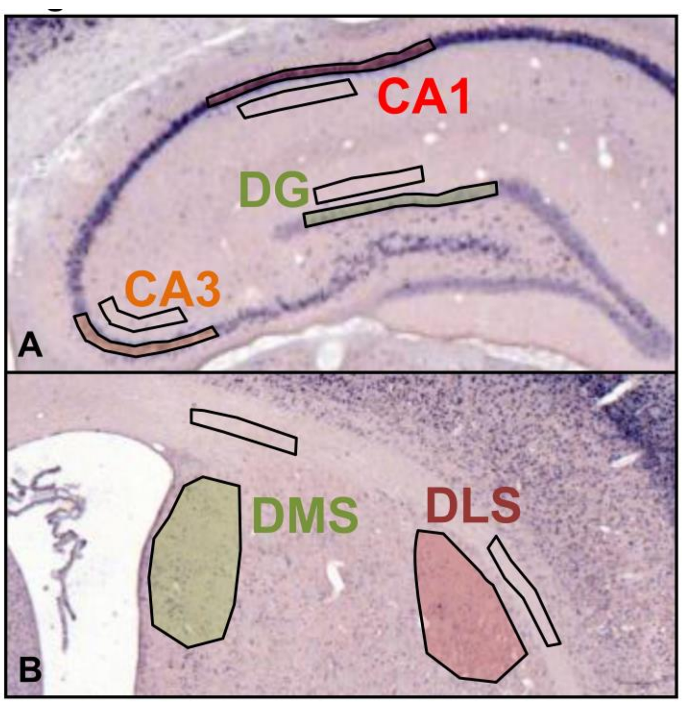


Figure 3. ***Hippocampal and striatal regions of interest.***
(A) Dorsal hippocampal Arc mRNA expression was measured in the pyramidal cell layer of CA1 and CA3, and in the granule cell layer of dentate gyrus (DG). (B) Dorsal striatal Arc mRNA expression was measured in medial (DMS) and lateral (DMS) sub-fields. (A-B) Exemplar regions of interest (shaded) and background regions (selected in areas that contained little to no punctate staining) are illustrated.

Background regions were selected from proximal areas containing little to no punctate staining (Fig. 3). Regional boundaries of each substructure were verified using the adult rat brain atlas (Paxinos and Watson, 2008). In addition, hippocampal regions of interest were drawn to ensure their inclusion in the applicable cell body layer. To facilitate consistent measurement, the surface area sampled from a given neural structure was maintained across sections. Image analysis was performed by a researcher blind to the experimental condition with which an image was associated.

Statistics

ANOVA was conducted to assess the effects of training duration on choice accuracy, reward latency, and VTE. Logistical regression with binary variables, fit with generalized estimating equations (Davis, 2002), was utilized to compare strategy use across training duration. Independent-sample t-tests were conducted to determine effects of training stage/strategy on Arc expression, and strategy groups on VTE. For select comparisons, Cohen's *d* was computed to estimate effect size. Statistical significance was interpreted using the criterion of $p < 0.05$. False discovery rate correction was applied to multiple comparisons (Benjamini and Hochberg, 1995). Statistical analyses were performed using SPSS (IBM), Excel (Microsoft), and R (Dalgaard, 2008).

Results

Rats were trained daily (6 trials/day) to perform a dual-solution task on a plus maze (Fig. 1). With training, animals typically reduced arm entry errors, and reduced their latency to locate the reward ($p < 0.001$; Fig. 4). For example, on the first day of training rats entered the rewarded arm first on ~60% of the trials taking ~57 seconds to reach the reward. In contrast, by the fourth training day, rats infrequently entered the unrewarded arm (less than 15% of the time), taking ~30 seconds to find the goal. Rats were trained for a minimum of four days, and until they reached 100% trial accuracy on the most recent training day and at least ~67% accuracy on the previous training day (see Methods). About 47% of rats reached these criteria by the fourth day, ~77% by the fifth day, and ~87% by the sixth day.

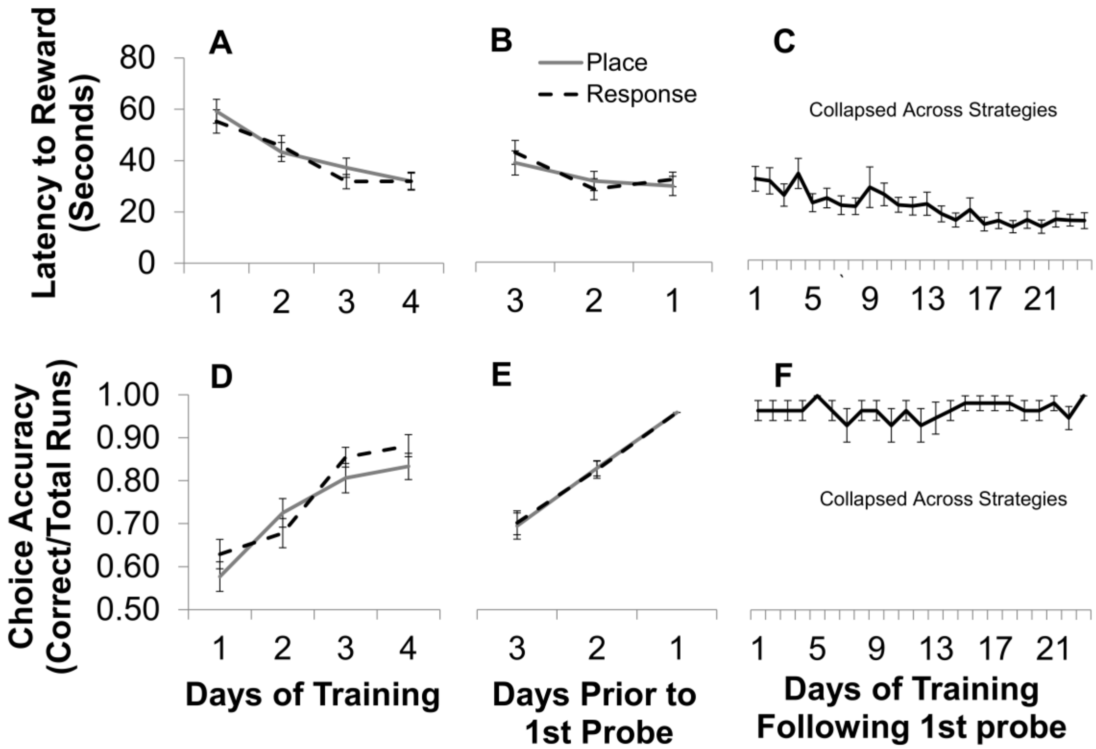


Figure 4. *Dual-solution task performance.*
Rats reduced their latency to find the goal (A-C) and arm entry errors (D-F) across days of training. (A-B, D-E) Equivalent learning curves between rats that used place strategies compared to those that used response strategies on the first probe suggest strategy reliance did not influence task performance. (C, F) High levels of performance were maintained with continued training. Rats that utilized place navigation on the first probe: n= 33; Rats that utilized response navigation on the first probe: n= 31. Error bars indicate ± one standard error of the mean.

The day after meeting these criteria, to identify the dominant strategy at that learning stage, the rat was started from the opposite arm to that used during training. If the rat relied on the spatial arrangement of extra-maze cues to locate the position of the reward, and thus entered the arm rewarded during training, it was classified as using a place strategy. If the rat relied on a specific motor turn and thus made the practiced turn, entering the arm unrewarded during training, it was classified as using a response strategy (Fig. 1). On this first probe, about half of the rats used a place strategy (Fig. 5A). Moreover, the learning rate (expressed by accuracy or reward latency) was equivalent between those rats that used a place strategy and those that used a response strategy on the first probe trial, indicating strategy reliance did not influence performance (Fig. 4). After the first probe, training re-commenced with subsequent probes given every seventh day (for a total of five probes across testing; Fig. 1C). Task performance, on average, was maintained at a high level throughout the remainder of the experiment; choice accuracy remained greater than ~95%, whereas the latency to locate the reward continued to decrease until reaching asymptotic values around the end of the fourth week of testing (Fig. 4). Moreover, with task repetition, rats gradually transitioned to almost exclusive reliance on response navigation. On the fourth probe, ~85% of animals used a response strategy, a proportion that increased to ~93% on the fifth and final probe (Fig. 5A).

Individual strategy analysis (see Gardner et al., 2013 for additional detail) of animals that completed all five probe trials revealed that the majority of rats transitioned to the use of a response strategy across the duration of the experiment (56%). Moreover, 26% of rats relied on response navigation throughout testing. In contrast, only one rat maintained throughout testing the use place navigation, and only two (7%) transitioned to rely on place navigation. Two animals did not fit these strategy classifications: one used a response strategy on all but the third probe, whereas the other consistently alternated between strategies.

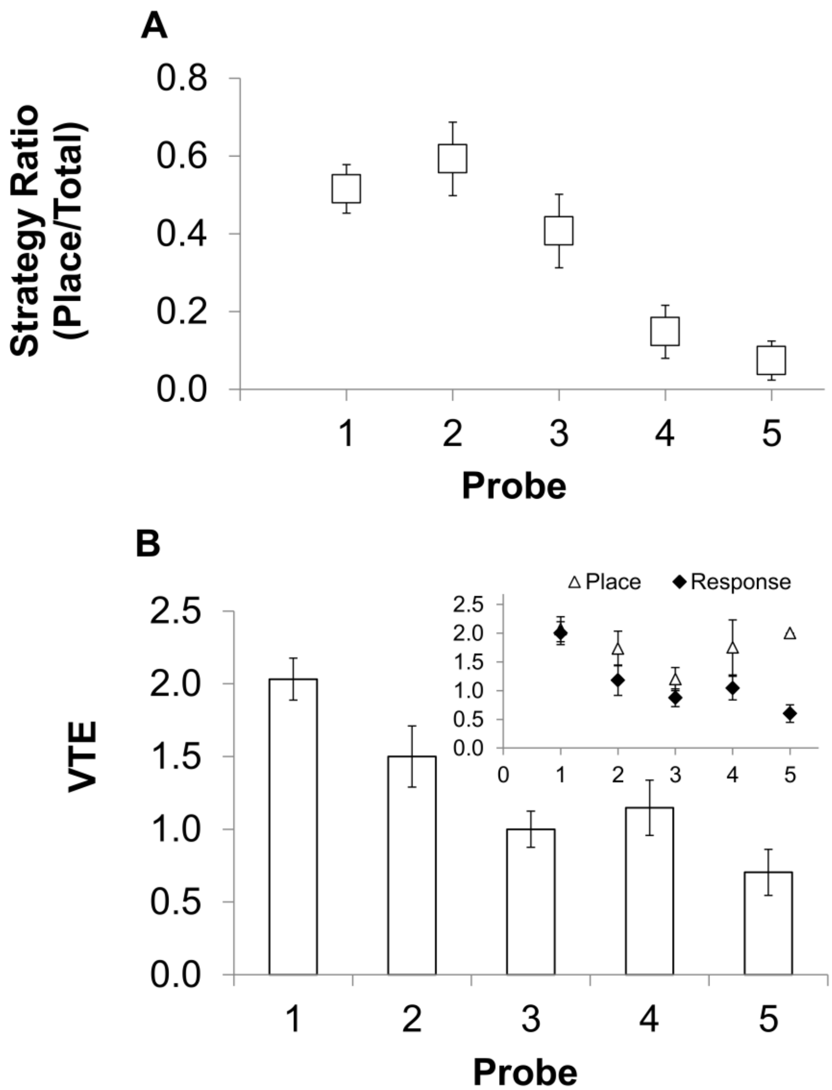


Figure 5. ***Strategy reliance and vicarious trial and error across training sessions.*** *(A) On the first probe, about half (52%) of rats relied on spatial (place) navigation. With continued training, response navigation dominated, as shown by an almost exclusive reliance on response strategies (93%) by the fifth and final probe. (B) Similar to the transition from place to response navigation, vicarious trial and error (VTE) decreased from the first to last probe trial. (B inset) However, this experience-dependent reduction in VTE appeared to be modulated by strategy engagement; place strategies were associated with a less severe drop that, after extensive training, remained relatively close to values measured early in training. Error bars indicate ± one standard error of the mean.*

On these probe runs, we also measured VTE, a proposed index of deliberation (Papale et al., 2012; Gardner et al., 2013; van der Meer et al., 2012; Schmidt et al., 2013). VTE was quantified by counting the number of pauses and discrete head orientations toward potential choices that occurred at the decision point of the maze (see Methods; Gardner et al., 2013). VTE occurrence was highest on the first probe trial, and decreased with continued task experience (Fig. 5B). In particular, VTE decreased two-fold comparing the fifth probe to the first probe. Thus, the occurrence of VTE was generally higher during place trials and lower during response trials.

Extending these findings, the magnitude of this strategy-dependent distinction appeared to be a function of training stage. Interestingly, VTE was equally frequent during place and response trials early in testing (e.g., probe 1; Fig. 5B Inset). Moreover, VTE events were reduced in an experience-dependent manner. However, while VTE events associated with response navigation remained relatively low late in training (i.e., probes 4-5), VTE events associated with place navigation were closer to that observed early in training (Fig. 5B Inset). Consistent with these findings, mean VTE across all probe sessions was relatively prominent in rats that used a place strategy throughout testing ($M = 2.5$, $n = 1$) or that transitioned to the use of place navigation ($M = 2.1$, $n = 2$), as compared to rats that transitioned to response navigation ($M = 1.3$, $SD = 0.5$, $n = 15$) or relied upon a response strategy throughout the experiment ($M = 0.8$, $SD = 0.7$, $n = 7$; $p < 0.05$).

All conclusions on strategy recruitment and VTE across training sessions were maintained when restricting analysis to those rats that completed all five probes and training sessions (i.e., upon excluding rats that were sacrificed after the first probe trial; data not shown; see Methods).

Colorimetric ISH using the AP-BCIP/NBT system (e.g., Kessler et al., 1990; Wehr et al., 2009) targeting Arc/Arg3.1 mRNA was performed to evaluate patterns of neural activation across dorsal hippocampus and dorsal striatum in rats that used a place strategy early in training and those that used a response strategy late in training (see Methods). This ISH procedure leaves a bluish-purple precipitate localized to cells that contain Arc mRNA. Given that Arc mRNA has relatively low basal expression levels, which increase quickly and dramatically in an activity-dependent manner (Guzowski et al., 1999; Pinaud and Tremere, 2006) selectively in alpha-CaMKII-expressing neurons within the striatum and pyramidal cell layer of the hippocampus (Vazdarjanova et al., 2006), colored (Arc positive) cells should indicate the principal neurons that were activated during recent experiences. Thus, measurement of the color intensity within a particular brain region (used here as an index of regional neural activation) should correspond to the number of principal neurons activated during maze traversal (probe trials).

Hippocampal Arc mRNA expression was relatively high in brains of rats that used a place strategy on the first probe as compared to those that used a response strategy on the fifth probe (Fig. 6A). This effect, while present in all hippocampal sub-regions, was strongest in CA1 and DG ($p < 0.01$; $d = 3.7$ and $d = 2.4$, respectively); the distinction in

Arc expression between strategies and did not reach statistical significance in CA3 (p > 0.10; d = 0.9). These data suggest that hippocampal activation is high early in training when animals are deliberative, and declines with repeated practice that coincides with the emergence of response navigation. In contrast, although there was a mild training stage-dependent increase in medial (d = 0.6) and decrease (d = 0.7) in lateral sub-fields, striatal Arc mRNA expression appeared relatively stable across conditions (Fig. 6B; p > 0.10). With the exception of a positive association between Arc expression in CA1 and DG (r = 0.83, p < 0.05), inter- and intra-regional correlations were non-significant, possibly due to the low sample size of rats processed for ISH (n=8).

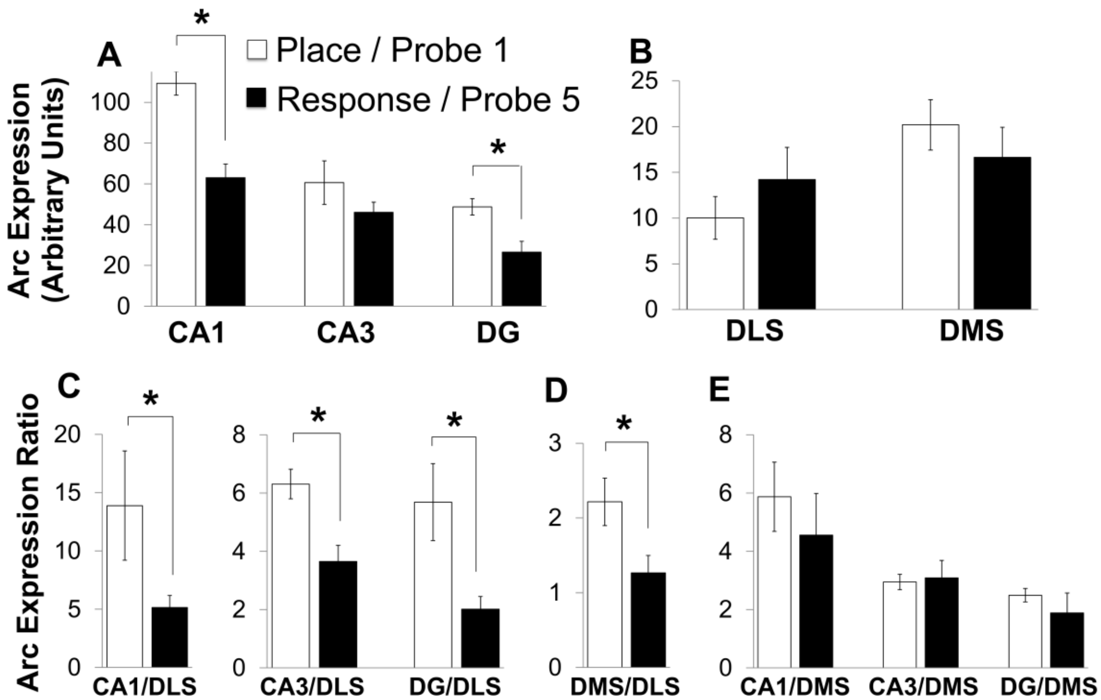


Figure 6. ***Distinct patterns of Arc expression within hippocampus and striatum correlate with the experience-dependent emergence of response navigation.*** *Arc mRNA expression in the dorsal hippocampus (CA1; CA3; dentate gyrus: DG) and dorsal striatum (medial: DMS; lateral: DLS) was quantified as the mean grayscale intensity (corrected for background; see Methods; Fig. 2-3). Expression among neural structures from rats that used a place strategy on the first probe and those that used a response strategy on the fifth probe is displayed. (A) Hippocampal Arc mRNA expression is high early in training when attentive performance dominates, and declines with task repetition which coincides with the emergence of response navigation. In contrast, Arc mRNA expression in medial and lateral striatal fields remains relatively stable across testing. (C-E) However, on an individual level, the ratio of expression in hippocampal to DLS (C), but not to DMS (E), and (D) the ratio of expression in medial to lateral striatal sub-fields discriminates strategy/training stage. Error bars indicate ± one standard error of the mean. *$p < 0.05$.*

Upon initial assessment, the number of VTE events on probe runs was not significantly correlated with hippocampal Arc mRNA expression ($p > 0.10$). However, we found enhanced expression levels in CA1 and DG ($p < 0.05$; $d = 2.5$ and 1.8, respectively) from rats that demonstrated some degree of VTE compared to those that did not (data not shown). These data are in agreement with the role of the hippocampus in deliberative navigation as assessed by VTE. Given that VTE and strategy were correlated, we calculated semi-partial correlations to assess their independent relationships to Arc expression. In both CA1 and DG, strategy appeared to explain a considerably larger amount of variation in Arc expression ($r = 0.64$ and 0.58, respectively) than VTE ($r = 0.17$ and 0.14, respectively).

To explore the possibility that relative activation across brain regions could discriminate strategy engagement (or learning stage), we also computed the ratios of Arc mRNA expression in hippocampal to striatal sub-regions, and expression in medial to lateral striatal sub-fields for each animal (Fig. 6C-E). As expected, the mean ratio of Arc expression in each hippocampal sub-field to that in the lateral striatum was significantly greater in rats that used a place strategy compared to those that used a response strategy ($p < 0.01$). In contrast, the expression ratios of hippocampal sub-fields to medial striatum were equivalent between conditions ($p > 0.10$).

In addition, the ratio of expression in medial to lateral striatum was increased in rats that used a place strategy early in training compared to rats that used a response strategy late in training ($p < 0.01$). Moreover, the size of this effect (Cohen's $d = 1.7$) was almost three-fold larger than that found when considering either medial or lateral striatum separately. Together, these findings suggest that relative activity patterns across place and response neural systems may be good predictors of the dominant mode underlying performance across task repetition, and/or indicate the onset of automatic (or habitual) behaviors.

Discussion

This research begins to identify activation patterns across hippocampus and striatum that correspond to the experience-dependent transition from attentive (place) to automatic (response) performance. This was accomplished by quantifying expression levels of the activity-dependent IEG Arc among rats that used a place strategy early in training and those that used a response strategy after extensive training on a dual-solution plus-maze task. Comparison across conditions identifies the changing degree of neuronal activation as related to the strategy transition. Complementary to this approach, we measured vicarious trial and error across training sessions to evaluate the changing degree of task-related deliberation.

Rats reached a high level of performance on the dual-solution task relatively quickly (~5 days). Moreover, at this training stage, about half of the rats relied on place navigation. Conversely, after extensive practice, rats transitioned to almost exclusive reliance on response navigation. This finding completely replicates studies showing an experience-dependent transition from place to response navigation (Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999). Extending these findings, as suggested by levels of Arc mRNA expression, activation of the hippocampus was relatively strong early in training and declined considerably after task repetition (coinciding with the emergence of response navigation). In contrast, Arc expression of the medial and lateral striatum, rose and fell, respectively, on a more modest scale. Interestingly, the ratio of Arc expression among structures implicated in place navigation (hippocampus and medial striatum) to those implicated in response navigation (lateral striatum) computed on an individual level significantly discriminated strategy engagement/training stage.

These data are generally consistent with studies showing that the relative patterns of acetylcholine release (as measured by in vivo micro-dialysis) within the hippocampus and lateral striatum correlate with strategy reliance across training (Chang and Gold, 2003). Moreover, our findings are in line with the notion that these systems compete to control behavior; although we did not find direct evidence for this hypothesis (e.g., as would be indicated by an inverse correlation between activation among these functionally disparate neural structures), we suggest that further data collection will be a valuable endeavor to clarify inter-regional interactions as they relate to decision making.

In contrast to Chang and Gold (2003), who found levels of acetylcholine release in the hippocampus were maintained throughout training (independent of strategy engagement), we found an experience-dependent reduction in hippocampal Arc expression. These discrepant findings may result from differences in the training schedule applied across experiments. In particular, Chang and Gold (2003) restricted testing to a single training day, whereas our testing duration was an average of thirty-four days. It is plausible that the activation of the hippocampus appreciably declines only after task repetition across many days of training, e.g., resulting from synaptic or cellular plasticity mechanisms acting on protracted time scales. As the firing of place cells is generally thought to be stable across days of testing (Muller, Kubie and Ranck, 1987; Thompson and Best, 1990; Barnes, Suster, Shen and McNaughton, 1997), however, it is also possible that Arc mRNA expression becomes uncoupled from neuronal firing across training. Although this does not appear to be the case during brief yet repeated bouts of spatial exploration across several days of testing (Guzowski et al., 2006), it is unclear if uncoupling occurs during a prolonged and incentivized task. From this perspective, and in consideration of the role of Arc in synaptic plasticity and memory consolidation (Guzowski et al., 2000), a reduction of hippocampal Arc expression late (compared to early) in training may not reflect a reduction of context-specific neural representation or information processing but rather a reduction of context-specific plasticity.

VTE behaviors were prominent early in testing and tended to decrease with training, an effect that mirrors the strategy transition. These results complement prior studies showing VTE is associated with place navigation (e.g., Gardner et al., 2013; Schmidt et al., 2013) and suggest, on average, a reduction of hippocampal engagement with task experience. However, this experience-dependent reduction of deliberative behavior also appeared to be modulated by strategy expression. Early in testing, VTE was similarly observed on place and response navigation probe trials. After extended training, however, VTE associated with place (compared to response) navigation was maintained comparatively close to levels observed early in training. Although the number of subjects that used a place strategy late in training was small (probe 4: n=4, probe 5: n=2), our findings suggest two potential relationships between VTE and strategy engagement. First, it is feasible that rats that show high levels of deliberation are relatively resistant to the strategy transition. Second, as the task becomes over-trained (and automatic response-based actions become increasingly ingrained in the animal), the degree of deliberation associated with the expression of rival (i.e., spatial) strategies is enhanced; this effect may be particularly heightened when the actions associated with place and response strategies are incompatible; e.g., during probe trials. We suggest further study to understand the changing dynamics between deliberation and the expression of navigational strategies across task repetition.

We stress that there are several notable limitations of the current work which should be addressed to verify and illuminate our findings. For example, the data presented here on Arc expression are restricted to a small number of rats that either used a place strategy early in training (n=4) or that used a response strategy after extensive training (n=4). Thus, any distinction between Arc expression across conditions may result from differences in strategy recruitment and/or training duration. Likewise, we did not collect tissue from naïve control animals and therefore could not determine the degree to which brain regions were activated from basal levels in response to training and strategy engagement. We employed a single-label colorimetric ISH design. This procedure makes counts of Arc-positive cells relatively difficult. Nonetheless, our findings qualitatively replicated previous studies that quantified the number of neurons expressing Arc after novel maze exploration; these experiments showed that the proportion of hippocampal neurons positive for Arc was highest in CA1, followed by CA3, and DG, and that the proportion of striatal neurons positive for Arc was higher in medial than lateral sub-fields (see Fig. 6; Vazdarjanova et al., 2006). This convergence between studies suggests that our measures may closely correlate with cell counts. However, we also suggest that fluorescent Arc mRNA labeling performed in conjunction with a nuclear stain that facilitates counts of cells positive for Arc will provide a distinct and potentially more precise assessment of mRNA expression within neural sub-fields. Moreover, given findings suggesting a functional distinction along the anterior-posterior axis of the dorsal striatum (e.g., Yin and Knowlton, 2004), precise measurement of Arc expression along this axis as it relates to strategy recruitment is warranted.

Altogether, this work provides distinct hypotheses on the changing patterns of activation within the hippocampus and striatum that underlies alterations in decision-making that occur with task experience. Although these findings are preliminary, the sizable effects of strategy/training-stage on region-specific Arc expression (which complement behavioral findings on VTE) highlight the value of this approach to study the neural mechanics of the transition from attentive to automatic performance.

Acknowledgements

This research was funded by the Office of Naval Research: ONR#000141010198, the Air Force Office of Scientific Research (AFOSR): Award No. FA9550-10-1-0385, and the Krasnow Institute for Advanced Study. We thank Christopher Rudnicky, Daniel Suarez, Nadira Robinson-Burton, Dana Achmar, Natalia Lattanzio, Jimmy Preston, David Lucente, Mariam Nasir, and Harmain Rafi for their assistance with behavioral testing and in situ hybridization experiments.

REFERENCES

Aberle, I., Rendell, P. G., Rose, N. S., McDaniel, M. A., and Kliegel, M. (2010). The age prospective memory paradox: young adults may not give their best outside of the lab. *Dev Psychol* 46, 1444–1453. doi:10.1037/a0020718.

Addis, D. R., Musicaro, R., Pan, L., and Schacter, D. L. (2010). Episodic simulation of past and future events in older adults: Evidence from an experimental recombination task. *Psychol Aging* 25, 369–376. doi:10.1037/a0017280.

Addis, D. R., Wong, A. T., and Schacter, D. L. (2008). Age-related changes in the episodic simulation of future events. *Psychol Sci* 19, 33–41. doi:10.1111/j.1467-9280.2008.02043.x.

Altgassen, M., Rendell, P. G., Bernhard, A., Henry, J. D., Bailey, P. E., Phillips, L. H., and Kliegel, M. (2014). Future thinking improves prospective memory performance and plan enactment in older adults. *Q J Exp Psychol (Hove)*, 1–30. doi:10.1080/17470218.2014.956127.

Aradillas, E., Libon, D. J., and Schwartzman, R. J. (2011). Acute loss of spatial navigational skills in a case of a right posterior hippocampus stroke. *J. Neurol. Sci.* 308, 144–146. doi:10.1016/j.jns.2011.06.026.

Ascoli, G. A., and Samsonovich, A. V. (2008). Science of the Conscious Mind. *Biol Bull* 215, 204–215.

Atance, C. M., and O’Neill, D. K. (2001). Episodic future thinking. *Trends Cogn. Sci. (Regul. Ed.)* 5, 533–539.

Baddeley, A. D., and Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior* 9, 176–189. doi:10.1016/S0022-5371(70)80048-2.

Baird, B., Smallwood, J., and Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Conscious Cogn* 20, 1604–1611. doi:10.1016/j.concog.2011.08.007.

Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.

Banner, H., Bhat, V., Etchamendy, N., Joober, R., and Bohbot, V. D. (2011). The brain-derived neurotrophic factor Val66Met polymorphism is associated with reduced functional magnetic resonance imaging activity in the hippocampus and increased use of caudate nucleus-dependent strategies in a human virtual navigation task. *Eur. J. Neurosci.* 33, 968–977. doi:10.1111/j.1460-9568.2010.07550.x.

Barnes, C. A., Suster, M. S., Shen, J., and McNaughton, B. L. (1997). Multistability of cognitive maps in the hippocampus of old rats. *Nature* 388, 272–275. doi:10.1038/40859.

Barsalou, L. W. (1988). "The content and organization of autobiographical memories," in *Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory*, eds. U. Neisser and E. Winograd (Cambridge University Press), 193–243.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. 2nd ed. Cambridge University Press.

Bäuml, K.-H., Hanslmayr, S., Pastötter, B., and Klimesch, W. (2008). Oscillatory correlates of intentional updating in episodic memory. *Neuroimage* 41, 596–604. doi:10.1016/j.neuroimage.2008.02.053.

Bayley, P. J., Hopkins, R. O., and Squire, L. R. (2006). The fate of old memories after medial temporal lobe damage. *J. Neurosci.* 26, 13311–13317. doi:10.1523/JNEUROSCI.4262-06.2006.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289–300. doi:10.2307/2346101.

Berntsen, D. (2002). Tunnel memories for autobiographical events: central details are remembered more frequently from shocking than from happy experiences. *Mem Cognit* 30, 1010–1020.

Berntsen, D., and Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Mem Cognit* 32, 789–803.

Berntsen, D., and Jacobsen, A. S. (2008). Involuntary (spontaneous) mental time travel into the past and future. *Conscious Cogn* 17, 1093–1104. doi:10.1016/j.concog.2008.03.001.

Berntsen, D., and Rubin, D. C. (2004). Cultural life scripts structure recall from autobiographical memory. *Mem Cognit* 32, 427–442.

Berntsen, D., and Rubin, D. C. (2012). *Understanding autobiographical memory: theories and approaches*. New York: Cambridge University Press.

Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., and Baxter, C. (2011). Remember the future: working memory training decreases delay discounting among stimulant addicts. *Biol. Psychiatry* 69, 260–265. doi:10.1016/j.biopsych.2010.08.017.

Bishara, A. J., and Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods* 17, 399–417. doi:10.1037/a0028087.

Bluck, S., and Alea, N. (2011). Crafting the TALE: construction of a measure to assess the functions of autobiographical remembering. *Memory* 19, 470–486. doi:10.1080/09658211.2011.590500.

Bluck, S., Alea, N., Habermas, T., and Rubin, D. C. (2005). A TALE of Three Functions: The Self?Reported Uses of Autobiographical Memory. *Social Cognition* 23, 91–117. doi:10.1521/soco.23.1.91.59198.

Bohbot, V. D., Iaria, G., and Petrides, M. (2004). Hippocampal function and spatial memory: evidence from functional neuroimaging in healthy participants and performance of patients with medial temporal lobe resections. *Neuropsychology* 18, 418–425. doi:10.1037/0894-4105.18.3.418.

Bohbot, V. D., Lerch, J., Thorndycraft, B., Iaria, G., and Zijdenbos, A. P. (2007). Gray Matter Differences Correlate with Spontaneous Strategies in a Human Virtual Navigation Task. *J. Neurosci.* 27, 10078–10083. doi:10.1523/JNEUROSCI.1763-07.2007.

Bohbot, V. D., McKenzie, S., Konishi, K., Fouquet, C., Kurdi, V., Schachar, R., Boivin, M., and Robaey, P. (2012). Virtual navigation strategies from childhood to senescence: evidence for changes across the life span. *Front Aging Neurosci* 4, 28. doi:10.3389/fnagi.2012.00028.

Bohn, A., and Berntsen, D. (2010). The Reminiscence Bump Reconsidered: Children's Prospective Life Stories Show a Bump in Young Adulthood. *Psychological Science* 22, 197–202. doi:10.1177/0956797610395394.

Botzung, A., Denkova, E., and Manning, L. (2008). Experiencing past and future personal events: Functional neuroimaging evidence on the neural bases of mental time travel. *Brain and Cognition* 66, 202–212. doi:10.1016/j.bandc.2007.07.011.

Brehmer, Y., Westerberg, H., and Bäckman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Front Hum Neurosci* 6, 63. doi:10.3389/fnhum.2012.00063.

Brewer, W. F. (1988). “Memory of randomly sampled autobiographical events,” in *Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory*, eds. U. Neisser and E. Winograd (Cambridge: Cambridge University Press), 21–89.

Brewer, W. F. (1986). “What is autobiographical memory?,” in *Autobiographical Memory*, ed. D. C. Rubin (Cambridge: Cambridge University Press), 25–49.

Cameron, P. (1972). The Generation Gap: Time Orientation. *The Gerontologist* 12, 117–119. doi:10.1093/geront/12.2_Part_1.117.

Cameron, P., Desai, K. G., Bahador, D., and Dremel, G. (1977). Temporality across the life-span. *Int J Aging Hum Dev* 8, 229–259.

Carrasco, M. C., Bernal, M. C., and Redolat, R. (2001). Time estimation and aging: a comparison between young and elderly adults. *Int J Aging Hum Dev* 52, 91–101.

Catal, L. L., and Fitzgerald, J. M. (2004). Autobiographical memory in two older adults over a twenty-year retention interval. *Mem Cognit* 32, 311–323.

Cavanaugh, J. C., Grady, J. G., and Perlmutter, M. (1983). Forgetting and use of memory aids in 20 to 70 year olds everyday life. *Int J Aging Hum Dev* 17, 113–122.

Chang, Q., and Gold, P. E. (2003). Switching memory systems during learning: changes in patterns of brain acetylcholine release in the hippocampus and striatum in rats. *The Journal of neuroscience* 23, 3001.

Chu, S., and Downes, J. J. (2000). Long live Proust: the odour-cued autobiographical memory bump. *Cognition* 75, B41–50.

Cockburn, J., and Smith, P. T. (1994). Anxiety and errors of prospective memory among elderly people. *British Journal of Psychology* 85, 273–282. doi:10.1111/j.2044-8295.1994.tb02523.x.

Coelho, M., Ferreira, J. J., Dias, B., Sampaio, C., Pavão Martins, I., and Castro-Caldas, A. (2004). Assessment of time perception: the effect of aging. *J Int Neuropsychol Soc* 10, 332–341. doi:10.1017/S1355617704103019.

Cohen, A.-L. (2013). Attentional decoupling while pursuing intentions: a form of mind wandering? *Front Psychol* 4, 693. doi:10.3389/fpsyg.2013.00693.

Cohen, G. (1998). “The effects of aging on autobiographical memory,” in *Autobiographical memory: Theoretical and applied perspectives*, eds. C. P. Thompson, D. J. Herrmann, D. Bruce, J. D. Read, D. G. Payne, and M. P. Toglia (Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers), 105–123.

Cohen, G., and Faulkner, D. (1989). Age differences in source forgetting: effects on reality monitoring and on eyewitness testimony. *Psychol Aging* 4, 10–17.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language* 53, 594–628. doi:10.1016/j.jml.2005.08.005.

Conway, M. A., and Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychol Rev* 107, 261–288.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia* 6, 255–265. doi:10.1016/0028-3932(68)90024-9.

Corkin, S., Amaral, D. G., González, R. G., Johnson, K. A., and Hyman, B. T. (1997). H. M.’s medial temporal lobe lesion: findings from magnetic resonance imaging. *J. Neurosci* 17, 3964–3979.

Crovitz, H. F., and Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society* 4, 517–518.

Csikszentmihalyi, M., and Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *J. Nerv. Ment. Dis.* 175, 526–536.

Csikszentmihalyi, M., Larson, R., and Prescott, S. (1977). The ecology of adolescent activity and experience. *J Youth Adolesc* 6, 281–294. doi:10.1007/BF02138940.

Dalgaard, P. (2008). *Introductory Statistics with R*. Springer.

D’Argembeau, A., Renaud, O., and Van der Linden, M. (2011). Frequency, characteristics and functions of future-oriented thoughts in daily life. *Applied Cognitive Psychology* 25, 96–103. doi:10.1002/acp.1647.

Davis, C. S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. Springer Science & Business Media.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research* 50, 61–77.

Derusso, A. L., Fan, D., Gupta, J., Shelest, O., Costa, R. M., and Yin, H. H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Front Integr Neurosci* 4. doi:10.3389/fnint.2010.00017.

Devan, B. D., Hong, N. S., and McDonald, R. J. (2011). Parallel associative processing in the dorsal striatum: segregation of stimulus-response and cognitive control subregions. *Neurobiol Learn Mem* 96, 95–120. doi:10.1016/j.nlm.2011.06.002.

Dumas, T. C., Powers, E. C., Tarapore, P. E., and Sapolsky, R. M. (2004). Overexpression of calbindin D28k in dentate gyrus granule cells alters mossy fiber presynaptic function and impairs hippocampal-dependent memory. *Hippocampus* 14, 701–709. doi:10.1002/hipo.10210.

Ezzyat, Y., and Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychol Sci* 22, 243–252. doi:10.1177/0956797610393742.

Fitzgerald, J. M., and Lawrence, R. (1984). Autobiographical memory across the life-span. *J Gerontol* 39, 692–698.

Friederich, H.-C., and Herzog, W. (2011). Cognitive-behavioral flexibility in anorexia nervosa. *Curr Top Behav Neurosci* 6, 111–123. doi:10.1007/7854_2010_83.

Fromholt, P., and Larsen, S. F. (1991). Autobiographical memory in normal aging and primary degenerative dementia (dementia of Alzheimer type). *J Gerontol* 46, P85–91.

Fromholt, P., Mortensen, D. B., Torpdahl, P., Bender, L., Larsen, P., and Rubin, D. C. (2003). Life-narrative and word-cued autobiographical memories in centenarians: comparisons with 80-year-old control, depressed, and dementia groups. *Memory* 11, 81–88.

Galton, F. (1879). Psychometric Experiments. *Brain* 2, 149–162. doi:10.1093/brain/2.2.149.

Gardner, R. S., Uttaro, M. R., Fleming, S. E., Suarez, D. F., Ascoli, G. A., and Dumas, T. C. (2013). A secondary working memory challenge preserves primary place strategies despite overtraining. *Learn. Mem.* 20, 648–656. doi:10.1101/lm.031336.113.

Gardner, R. S., Vogel, A. T., Mainetti, M., and Ascoli, G. A. (2012). Quantitative measurements of autobiographical memory content. *PLoS ONE* 7, e44809. doi:10.1371/journal.pone.0044809.

Giambra, L. M. (1979). Sex differences in daydreaming and related mental activity from the late teens to the early nineties. *Int J Aging Hum Dev* 10, 1–34.

Giambra, L. M. (1989). Task-unrelated-thought frequency as a function of age: a laboratory study. *Psychol Aging* 4, 136–143.

Gibbs, B. R., and Rude, S. S. (2004). Overgeneral Autobiographical Memory as Depression Vulnerability. *Cognitive Therapy and Research* 28, 511–526. doi:10.1023/B:COTR.0000045561.72997.7c.

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., and de Wit, S. (2011). Disruption in the Balance Between Goal-Directed Behavior and Habit Learning in Obsessive-Compulsive Disorder. *Am J Psychiatry* 168, 718–726. doi:10.1176/appi.ajp.2011.10071062.

Guzowski, J. F., Lyford, G. L., Stevenson, G. D., Houston, F. P., McGaugh, J. L., Worley, P. F., and Barnes, C. A. (2000). Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. *J. Neurosci.* 20, 3993–4001.

Guzowski, J. F., McNaughton, B. L., Barnes, C. A., and Worley, P. F. (1999). Environment-specific expression of the immediate-early gene Arc in hippocampal neuronal ensembles. *Nat. Neurosci.* 2, 1120–1124. doi:10.1038/16046.

Guzowski, J. F., Miyashita, T., Chawla, M. K., Sanderson, J., Maes, L. I., Houston, F. P., Lipa, P., McNaughton, B. L., Worley, P. F., and Barnes, C. A. (2006). Recent behavioral history modifies coupling between cell activity and Arc gene transcription in hippocampal CA1 neurons. *Proc. Natl. Acad. Sci. U.S.A.* 103, 1077–1082. doi:10.1073/pnas.0505519103.

Hagewoud, R., Havekes, R., Tiba, P. A., Novati, A., Hogenelst, K., Weinreder, P., Van der Zee, E. A., and Meerlo, P. (2010). Coping with sleep deprivation: shifts in regional brain activity and learning strategy. *Sleep* 33, 1465–1473.

Hamilton, D. A., Rosenfelt, C. S., and Whishaw, I. Q. (2004). Sequential control of navigation by locale and taxon cues in the Morris water task. *Behav. Brain Res.* 154, 385–397. doi:10.1016/j.bbr.2004.03.005.

Hartley, T., Maguire, E. A., Spiers, H. J., and Burgess, N. (2003). The well-worn route and the path less traveled: distinct neural bases of route following and wayfinding in humans. *Neuron* 37, 877–888.

Hashtroudi, S., Johnson, M. K., and Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychol Aging* 5, 119–126.

He, J., Becic, E., Lee, Y.-C., and McCarley, J. S. (2011). Mind wandering behind the wheel: performance and oculomotor correlates. *Hum Factors* 53, 13–21.

Hicks, L. H. (1964). Effects of overtraining on acquisition and reversal of place and response learning. *Psychological Reports* 15, 459–462. doi:10.2466/pr0.1964.15.2.459.

Hogarth, L., Balleine, B. W., Corbit, L. H., and Killcross, S. (2012). Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Ann. N. Y. Acad. Sci.* doi:10.1111/j.1749-6632.2012.06768.x.

Holdstock, J. S., Mayes, A. R., Cezayirli, E., Aggleton, J. P., and Roberts, N. (1999). A comparison of egocentric and allocentric spatial memory in medial temporal lobe and Korsakoff amnesics. *Cortex* 35, 479–501.

Howes, J. L., and Katz, A. N. (1992). Remote memory: recalling autobiographical and public events from across the lifespan. *Can J Psychol* 46, 92–116.

Huang-Pollock, C. L., and Karalunas, S. L. (2010). Working memory demands impair skill acquisition in children with ADHD. *J Abnorm Psychol* 119, 174–185. doi:10.1037/a0017862.

Hu, D., and Amsel, A. (1995). A simple test of the vicarious trial-and-error hypothesis of hippocampal function. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5506–5509.

Hu, D., Griesbach, G., and Amsel, A. (1997). Development of vicarious trial-and-error behavior in odor discrimination learning in the rat: relation to hippocampal function? *Behav. Brain Res.* 86, 67–70.

Hu, D., Xu, X., and Gonzalez-Lima, F. (2006). Vicarious trial-and-error behavior and hippocampal cytochrome oxidase activity during Y-maze discrimination learning in the rat. *Int. J. Neurosci* 116, 265–280. doi:10.1080/00207450500403108.

Hupbach, A., Gomez, R., Hardt, O., and Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learn Mem* 14, 47–53. doi:10.1101/lm.365707.

Huttenlocher, J., Hedges, L. V., and Bradburn, N. M. (1990). Reports of elapsed time: bounding and rounding processes in estimation. *J Exp Psychol Learn Mem Cogn* 16, 196–213.

Iaria, G., Petrides, M., Dagher, A., Pike, B., and Bohbot, V. D. (2003). Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: variability and change with practice. *J. Neurosci.* 23, 5945–5952.

Jackson, J. D., and Balota, D. A. (2012). Mind-wandering in younger and older adults: Converging evidence from the sustained attention to response task and reading for comprehension. *Psychology and Aging* 27, 106–119. doi:10.1037/a0023933.

James, L. E., Burke, D. M., Austin, A., and Hulme, E. (1998). Production and perception of “verbosity” in younger and older adults. *Psychol Aging* 13, 355–367.

James, W. (1950). *The Principles of Psychology, Vol. 1*. Reprint edition. New York: Dover Publications.

Jansari, A., and Parkin, A. J. (1996). Things that go bump in your life: explaining the reminiscence bump in autobiographical memory. *Psychol Aging* 11, 85–91.

Janssen, S. M. J., Chessa, A. G., and Murre, J. M. J. (2005). The reminiscence bump in autobiographical memory: effects of age, gender, education, and culture. *Memory* 13, 658–668. doi:10.1080/09658210444000322.

Janssen, S. M. J., Rubin, D. C., and St Jacques, P. L. (2011). The temporal distribution of autobiographical memory: changes in reliving and vividness over the life span do not explain the reminiscence bump. *Mem Cognit* 39, 1–11. doi:10.3758/s13421-010-0003-x.

Jeffreys, H. (1961). *Theory of Probability*. 3rd ed. Oxford University Press, Oxford.

Johannessen, K. B., and Berntsen, D. (2010). Current concerns in involuntary and voluntary autobiographical memories. *Conscious Cogn* 19, 847–860. doi:10.1016/j.concog.2010.01.009.

Johnson, A., Varberg, Z., Benhardus, J., Maahs, A., and Schrater, P. (2012). The hippocampus and exploration: dynamically evolving behavior and neural representations. *Front Hum Neurosci* 6, 216. doi:10.3389/fnhum.2012.00216.

Johnson, M. K. (1997). Source monitoring and memory distortion. *Philos Trans R Soc Lond B Biol Sci* 352, 1733–1745.

Johnson, M. K., Foley, M. A., Suengas, A. G., and Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *J Exp Psychol Gen* 117, 371–376.

Johnson, M. K., Kahan, T. L., and Raye, C. L. (1984). Dreams and reality monitoring. *J Exp Psychol Gen* 113, 329–344.

Johnson, M. K., O'Connor, M., and Cantor, J. (1997). Confabulation, memory deficits, and frontal dysfunction. *Brain Cogn* 34, 189–206. doi:10.1006/brcg.1997.0873.

Johnson, M. K., and Raye, C. L. (1981). Reality monitoring. *Psychological Review* 88, 67–85. doi:10.1037/0033-295X.88.1.67.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., and Stone, A. A. (2004). A survey method for characterizing daily life experience: the day reconstruction method. *Science* 306, 1776–1780. doi:10.1126/science.1103572.

Kesner, R. P., Gilbert, P. E., and Barua, L. A. (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. *Behav. Neurosci.* 116, 286–290.

Kessler, C., Höltke, H. J., Seibl, R., Burg, J., and Mühlegger, K. (1990). Non-radioactive labeling and detection of nucleic acids. I. A novel DNA labeling and detection system based on digoxigenin: anti-digoxigenin ELISA principle (digoxigenin system). *Biol. Chem. Hoppe-Seyler* 371, 917–927.

Killingsworth, M. A., and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science* 330, 932. doi:10.1126/science.1192439.

Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A Neostriatal Habit Learning System in Humans. *Science* 273, 1399–1402. doi:10.1126/science.273.5280.1399.

De Kleine, E., and Van der Lubbe, R. H. J. (2011). Decreased load on general motor preparation and visual-working memory while preparing familiar as compared to unfamiliar movement sequences. *Brain Cogn* 75, 126–134. doi:10.1016/j.bandc.2010.10.013.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends Cogn. Sci. (Regul. Ed.)* 14, 317–324. doi:10.1016/j.tics.2010.05.002.

Klinger, E., and Cox, W. M. (1987). Dimensions of Thought Flow in Everyday Life. *Imagination, Cognition and Personality* 7, 105–128. doi:10.2190/7K24-G343-MTQW-115V.

Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.

Kopelman, M. D., Wilson, B. A., and Baddeley, A. D. (1989). The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *J Clin Exp Neuropsychol* 11, 724–744. doi:10.1080/01688638908400928.

Kvavilashvili, L., and Fisher, L. (2007). Is time-based prospective remembering mediated by self-initiated rehearsals? Role of incidental cues, ongoing activity, age, and motivation. *J Exp Psychol Gen* 136, 112–132. doi:10.1037/0096-3445.136.1.112.

Lerch, J. P., Yiu, A. P., Martinez-Canabal, A., Pekar, T., Bohbot, V. D., Frankland, P. W., Henkelman, R. M., Josselyn, S. A., and Sled, J. G. (2011). Maze training in mice induces MRI-detectable brain shape changes specific to the type of learning. *Neuroimage* 54, 2086–2095. doi:10.1016/j.neuroimage.2010.09.086.

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., and Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging* 17, 677–689. doi:10.1037/0882-7974.17.4.677.

Linton, M. (1986). “Ways of searching and the contents of memory,” in *Autobiographical Memory*, ed. D. C. Rubin (Cambridge: Cambridge University Press), 50–67.

Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: the Proustian view is incorrect. *Appl. Cognit. Psychol.* 18, 893–899. doi:10.1002/acp.1020.

Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory* 13, 874–884. doi:10.1080/09658210444000485.

Maguire, E. A., Woollett, K., and Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus* 16, 1091–1101. doi:10.1002/hipo.20233.

Manns, J. R., Hopkins, R. O., and Squire, L. R. (2003). Semantic memory and the human hippocampus. *Neuron* 38, 127–133.

McCormack, P. D. (1979). Autobiographical memory in the aged. *Canadian Journal of Psychology/Revue canadienne de psychologie* 33, 118–124. doi:10.1037/h0081705.

McDaniel, M. A., and Einstein, G. O. (2007). *Prospective memory an overview and synthesis of an emerging field*. Thousand Oaks, Calif.: SAGE Publications.

McDonald, R. J., Devan, B. D., and Hong, N. S. (2004). Multiple memory systems: the power of interactions. *Neurobiol Learn Mem* 82, 333–346. doi:10.1016/j.nlm.2004.05.009.

McDonald, R. J., and White, N. M. (1993). A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav. Neurosci.* 107, 3–22.

Van der Meer, M., Kurth-Nelson, Z., and Redish, A. D. (2012). Information processing in decision-making systems. *Neuroscientist* 18, 342–359. doi:10.1177/1073858411435128.

Mello, E. W., and Fisher, R. P. (1996). Enhancing Older Adult Eyewitness Memory with the Cognitive Interview. *Appl. Cognit. Psychol.* 10, 403–417. doi:10.1002/(SICI)1099-0720(199610)10:5<403::AID-ACP395>3.0.CO;2-X.

Milner, B. (1962). Les troubles de la me´moire accompagnant des le´sions hippocampiques bilate´rales. In Physiologie de l_hippocampe (pp. 257–272). Paris: Centre National de la Recherche Scientifique. English translation: B. Milner & S. Glickman (Eds.). Princeton: Van Nostrand, 1965 (pp. 97–111).

Milner, B., Squire, L. R., and Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron* 20, 445–468.

Mitzner, T. L., Boron, J. B., Fausset, C. B., Adams, A. E., Charness, N., Czaja, S. J., Dijkstra, K., Fisk, A. D., Rogers, W. A., and Sharit, J. (2010). Older adults talk technology: Technology usage and attitudes. *Computers in Human Behavior* 26, 1710–1721. doi:10.1016/j.chb.2010.06.020.

Morillon, B., Kell, C. A., and Giraud, A.-L. (2009). Three stages and four neural systems in time estimation. *J. Neurosci.* 29, 14803–14811. doi:10.1523/JNEUROSCI.3222-09.2009.

Muenzinger, K. F. (1938). Vicarious Trial and Error at a Point of Choice: I. A General Survey of its Relation to Learning Efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology* 53, 75–86. doi:10.1080/08856559.1938.10533799.

Muenzinger, K. F., and Gentry, E. (1931). Tone discrimination in white rats. *Journal of Comparative Psychology* 12, 195–206. doi:10.1037/h0072238.

Muller, R. U., Kubie, J. L., and Ranck, J. B. (1987). Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. *J. Neurosci.* 7, 1935–1950.

Nadel, L., Campbell, J., and Ryan, L. (2007). Autobiographical memory retrieval and hippocampal activation as a function of repetition and the passage of time. *Neural Plast.* 2007, 90472. doi:10.1155/2007/90472.

Nadel, L., Samsonovich, A., Ryan, L., and Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* 10, 352–368. doi:10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D.

Niedźwieńska, A., and Barzykowski, K. (2012). The age prospective memory paradox within the same sample in time-based and event-based tasks. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* 19, 58–83. doi:10.1080/13825585.2011.628374.

Di Nocera, F., Fabrizi, R., Terenzi, M., and Ferlazzo, F. (2006). Procedural errors in air traffic control: effects of traffic density, expertise, and automation. *Aviat Space Environ Med* 77, 639–643.

O'Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. First Edition. Oxford University Press, USA.

Packard, M. G. (1987). *Differential Roles of Hippocampus and Caudate Nucleus in Memory: Selective Mediation of "cognitive" and "associative" Learning*. McGill University.

Packard, M. G. (2009). Exhumed from thought: basal ganglia and response learning in the plus-maze. *Behav. Brain Res* 199, 24–31. doi:10.1016/j.bbr.2008.12.013.

Packard, M. G. (1999). Glutamate infused posttraining into the hippocampus or caudate-putamen differentially strengthens place and response learning. *Proceedings of the National Academy of Sciences of the United States of America* 96, 12881.

Packard, M. G., and McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory* 65, 65–72.

Packard, M., Hirsh, R., and White, N. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *The Journal of Neuroscience* 9, 1465 –1472.

Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S., and Redish, A. D. (2012). Interactions between deliberation and delay-discounting in rats. *Cogn Affect Behav Neurosci* 12, 513–526. doi:10.3758/s13415-012-0097-7.

Paxinos, G., and Watson, C. (2008). *The Rat Brain in Stereotaxic Coordinates: Compact, Sixth Edition*. 6 edition. Amsterdam; Boston; London: Academic Press.

Peeters, F., Nicolson, N. A., Berkhof, J., Delespaul, P., and deVries, M. (2003). Effects of daily events on mood states in major depressive disorder. *J Abnorm Psychol* 112, 203–211.

Phillips, L. H., Henry, J. D., and Martin, M. (2008). "Adult aging and prospective memory: The importance of ecological validity," in *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives*, eds. M. Kliegel, M. A. McDaniel, and G. O. Einstein (New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates), 161–185.

Pillemer, D. B. (1992). "Remembering personal circumstances: A functional analysis," in *Affect and accuracy in recall* Emory Symposia in Cognition. (Cambridge University Press). Available at: http://dx.doi.org/10.1017/CBO9780511664069.013.

Pinaud, R., and Tremere, L. A. eds. (2006). *Immediate Early Genes in Sensory Processing, Cognitive Performance and Neurological Disorders*. 2006 edition. New York, NY: Springer.

Piolino, P., Desgranges, B., Benali, K., and Eustache, F. (2002). Episodic and semantic remote autobiographical memory in ageing. *Memory* 10, 239–257. doi:10.1080/09658210143000353.

Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., and Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature* 414, 546–550. doi:10.1038/35107080.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., and Gabrieli, J. D. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology* 13, 564–574.

Pol-Bodetto, S., Jeltsch-David, H., Lecourtier, L., Rusnac, N., Mam-Lam-Fook, C., Cosquer, B., Geiger, K., and Cassel, J.-C. (2011). The double-H maze test, a novel, simple, water-escape memory task: acquisition, recall of recent and remote memory, and effects of systemic muscarinic or NMDA receptor blockade during training. *Behav. Brain Res* 218, 138–151. doi:10.1016/j.bbr.2010.11.043.

Radu, P. T., Yi, R., Bickel, W. K., Gross, J. J., and McClure, S. M. (2011). A mechanism for reducing delay discounting by altering temporal attention. *J Exp Anal Behav* 96, 363–385. doi:10.1901/jeab.2011.96-363.

Rah, M. J., Walline, J. J., Lynn Mitchell, G., and Zadnik, K. (2006). Comparison of the experience sampling method and questionnaires to assess visual activities in pre-teen and adolescent children*. *Ophthalmic and Physiological Optics* 26, 483–489. doi:10.1111/j.1475-1313.2006.00372.x.

Rasmussen, A. S., and Berntsen, D. (2011). The unpredictable past: spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Conscious Cogn* 20, 1842–1846. doi:10.1016/j.concog.2011.07.010.

Rathbone, C. J., Moulin, C. J. A., and Conway, M. A. (2008). Self-centered memories: the reminiscence bump and the self. *Mem Cognit* 36, 1403–1414. doi:10.3758/MC.36.8.1403.

Reichenbach, J., Onnasch, L., and Manzey, D. (2011). Human performance consequences of automated decision aids in states of sleep loss. *Hum Factors* 53, 717–728.

Remondes, M., and Schuman, E. M. (2004). Role for a cortical input to hippocampal area CA1 in the consolidation of a long-term memory. *Nature* 431, 699–703. doi:10.1038/nature02965.

Rendell, P. G., and Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Appl. Cognit. Psychol.* 14, S43–S62. doi:10.1002/acp.770.

Restle, F. (1957). Discrimination of cues in mazes: a resolution of the place-vs.-response question. *Psychol Rev* 64, 217–228.

Ritchie, B. F., Aeschllman, B., and Peirce, P. (1950). Studies in spatial learning.’. *The Journal of comparative and physiological psychology*, 73.

Ritchie, T. D., Skowronski, J. J., Walker, W. R., and Wood, S. E. (2006). Comparing two perceived characteristics of autobiographical memory: Memory detail and accessibility. *Memory* 14, 471–485. doi:10.1080/09658210500478434.

Robinson, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology* 8, 578–595. doi:10.1016/0010-0285(76)90020-7.

Rondi-Reig, L., Petit, G. H., Tobin, C., Tonegawa, S., Mariani, J., and Berthoz, A. (2006). Impaired sequential egocentric and allocentric memories in forebrain-specific-NMDA receptor knock-out mice during a new task dissociating strategies of navigation. *J. Neurosci.* 26, 4071–4081. doi:10.1523/JNEUROSCI.3408-05.2006.

Rosenbaum, R. S., Köhler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E., Gao, F., and Tulving, E. (2005). The case of K.C.: contributions of a memory-impaired person to memory theory. *Neuropsychologia* 43, 989–1021. doi:10.1016/j.neuropsychologia.2004.10.007.

Rubin, D. C. (1988). *Autobiographical Memory*. Cambridge University Press.

Rubin, D. C. (1982). On the retention function for autobiographical memory. *Journal of Verbal Learning and Verbal Behavior* 21, 21–38. doi:10.1016/S0022-5371(82)90423-6.

Rubin, D. C. (2000). The distribution of early childhood memories. *Memory* 8, 265–269. doi:10.1080/096582100406810.

Rubin, D. C., and Berntsen, D. (2009). The frequency of voluntary and involuntary autobiographical memories across the life span. *Mem Cognit* 37, 679–688. doi:10.3758/37.5.679.

Rubin, D. C., Boals, A., and Berntsen, D. (2008). Memory in posttraumatic stress disorder: properties of voluntary and involuntary, traumatic and nontraumatic autobiographical memories in people with and without posttraumatic stress disorder symptoms. *J Exp Psychol Gen* 137, 591–614. doi:10.1037/a0013165.

Rubin, D. C., Groth, E., and Goldsmith, D. J. (1984). Olfactory cuing of autobiographical memory. *Am J Psychol* 97, 493–507.

Rubin, D. C., Rahhal, T. A., and Poon, L. W. (1998). Things learned in early adulthood are remembered best. *Mem Cognit* 26, 3–19.

Rubin, D. C., Schrauf, R. W., and Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Mem Cognit* 31, 887–901.

Rubin, D. C., and Schulkind, M. D. (1997a). Distribution of important and word-cued autobiographical memories in 20-, 35-, and 70-year-old adults. *Psychol Aging* 12, 524–535.

Rubin, D. C., and Schulkind, M. D. (1997b). The distribution of autobiographical memories across the lifespan. *Mem Cognit* 25, 859–866.

Rubin, D. C., Schulkind, M. D., and Rahhal, T. A. (1999). A Study of Gender Differences in Autobiographical Memory: Broken Down by Age and Sex. *Journal of Adult Development* 6, 61–71. doi:10.1023/A:1021676309064.

Rubin, D. C., and Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review* 103, 734–760. doi:10.1037/0033-295X.103.4.734.

Rubin, D. C., Wetzler, S., and Nebes, R. (1986). "Autobiographical memory across the adult lifespan," in *Autobiographical Memory*, ed. D. C. Rubin (Cambridge: Cambridge University Press), 202–221.

Rybash, J. M., and Monaghan, B. E. (1999). Episodic and semantic contributions to older adults' autobiographical recall. *J Gen Psychol* 126, 85–96. doi:10.1080/00221309909595353.

Ryle, G., and Dennett, D. C. (2000). *The Concept of Mind*. 1 edition. Chicago: University Of Chicago Press.

Schacter, D. L., and Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philos Trans R Soc Lond B Biol Sci* 362, 773–786. doi:10.1098/rstb.2007.2087.

Schaie, K. W. (1977). "Quasi-experimental designs in the psychology of aging," in *Handbook of the Psychology of Aging*, eds. J. E. Birren and K. W. Schaie (New York: Van Nostrand Reinhold), 39–58.

Schlagman, S., Kliegel, M., Schulz, J., and Kvavilashvili, L. (2009). Differential effects of age on involuntary and voluntary autobiographical memory. *Psychol Aging* 24, 397–411. doi:10.1037/a0015785.

Schlagman, S., and Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: how different are they from voluntary autobiographical memories? *Mem Cognit* 36, 920–932.

Schmidt, B., Papale, A., Redish, A. D., and Markus, E. J. (2013). Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors. *Learn. Mem.* 20, 130–138. doi:10.1101/lm.028753.112.

Schmidtke, K., Manner, H., Kaufmann, R., and Schmolck, H. (2002). Cognitive Procedural Learning in Patients With Fronto-Striatal Lesions. *Learn Mem* 9, 419–429. doi:10.1101/lm.47202.

Schmitzer-Torbert, N. (2007). Place and response learning in human virtual navigation: Behavioral measures and gender differences. *Behavioral neuroscience* 121, 277–290.

Schnitzspahn, K. M., Ihle, A., Henry, J. D., Rendell, P. G., and Kliegel, M. (2011). The age-prospective memory-paradox: an exploration of possible mechanisms. *Int Psychogeriatr* 23, 583–592. doi:10.1017/S1041610210001651.

Schooler, J. W., Reichle, E. D., and Halpern, D. V. (2004). "Zoning Out while Reading: Evidence for Dissociations between Experience and Metaconsciousness," in *Thinking and seeing: Visual metacognition in adults and children* (Cambridge, MA, US: MIT Press), 203–226.

Schrauf, R. W., and Rubin, D. C. (2000). Internal languages of retrieval: the bilingual encoding of memories for the personal past. *Mem Cognit* 28, 616–623.

Schwabe, L., Schächinger, H., de Kloet, E. R., and Oitzl, M. S. (2010). Corticosteroids operate as a switch between memory systems. *J Cogn Neurosci* 22, 1362–1372. doi:10.1162/jocn.2009.21278.

Scoville, W. B., and Milner, B. (2000). Loss of recent memory after bilateral hippocampal lesions. 1957. *J Neuropsychiatry Clin Neurosci* 12, 103–113.

Sheskin, D. J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures, Fifth Edition*. 5th ed. Chapman and Hall/CRC.

Smallwood, J., Nind, L., and O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Conscious Cogn* 18, 118–125. doi:10.1016/j.concog.2008.11.004.

Smallwood, J., and Schooler, J. W. (2006). The restless mind. *Psychol Bull* 132, 946–958. doi:10.1037/0033-2909.132.6.946.

Smith, K. S., Virkud, A., Deisseroth, K., and Graybiel, A. M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18932–18937. doi:10.1073/pnas.1216264109.

Spencer, W. D., and Raz, N. (1995). Differential effects of aging on memory for content and context: a meta-analysis. *Psychol Aging* 10, 527–539.

Spreng, R. N., and Levine, B. (2006). The temporal distribution of past and future autobiographical events across the lifespan. *Mem Cognit* 34, 1644–1651.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of learning and memory* 82, 171–177.

Squire, L. R., Knowlton, B., and Musen, G. (1993). The structure and organization of memory. *Annu Rev Psychol* 44, 453–495. doi:10.1146/annurev.ps.44.020193.002321.

Stawarczyk, D., Majerus, S., Maj, M., Van der Linden, M., and D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica* 136, 370–381. doi:10.1016/j.actpsy.2011.01.002.

St. Jacques, P. L., and Levine, B. (2007). Ageing and autobiographical memory for emotional and neutral events. *Memory* 15, 129–144. doi:10.1080/09658210601119762.

Sullivan, B. T., Johnson, L., Rothkopf, C. A., Ballard, D., and Hayhoe, M. (2012). The role of uncertainty and reward on eye movements in a virtual driving task. *J Vis* 12. doi:10.1167/12.13.19.

Sutin, A. R., and Robins, R. W. (2007). Phenomenology of autobiographical memories: the memory experiences questionnaire. *Memory* 15, 390–411. doi:10.1080/09658210701256654.

Svoboda, E., and Levine, B. (2009). The effects of rehearsal on the functional neuroanatomy of episodic autobiographical and semantic remembering: a functional magnetic resonance imaging study. *J. Neurosci.* 29, 3073–3082. doi:10.1523/JNEUROSCI.3452-08.2009.

Thompson, L. T., and Best, P. J. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. *Brain Res.* 509, 299–308.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review* 55, 189–208. doi:10.1037/h0061626.

Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychological Review* 45, 1–41. doi:37/h0062733.

Tolman, E. C., Ritchie, B. F., and Kalish, D. (1946). Studies in spatial learning; place learning versus response learning. *J Exp Psychol* 36, 221–229.

Tranel, D., Damasio, A. R., Damasio, H., and Brandt, J. P. (1994). Sensorimotor skill learning in amnesia: additional evidence for the neural basis of nondeclarative memory. *Learn. Mem.* 1, 165–179.

Trull, T. J., and Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. *Psychol Assess* 21, 457–462. doi:10.1037/a0017653.

Tulving, E. (1972). “Episodic and semantic memory,” in *Organization of memory*, eds. E. Tulving and W. Donaldson (New York: Academic Press), 381–403.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne* 26, 1–12. doi:10.1037/h0080017.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270. doi:10.1016/j.cognition.2003.10.008.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., VanPaesschen, W., and Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *SCIENCE* 277, 376–380.

Vazdarjanova, A., Ramirez-Amaya, V., Insel, N., Plummer, T. K., Rosi, S., Chowdhury, S., Mikhael, D., Worley, P. F., Guzowski, J. F., and Barnes, C. A. (2006). Spatial exploration induces ARC, a plasticity-related immediate-early gene, only in calcium/calmodulin-dependent protein kinase II-positive principal excitatory and inhibitory neurons of the rat forebrain. *J. Comp. Neurol.* 498, 317–329. doi:10.1002/cne.21003.

Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology* 18, 225–252. doi:10.1016/0010-0285(86)90013-7.

Walker, W. R., Skowronski, J. J., Gibbons, J. A., Vogl, R. J., and Ritchie, T. D. (2009). Why people rehearse their memories: frequency of use and relations to the intensity of emotions associated with autobiographical memories. *Memory* 17, 760–773. doi:10.1080/09658210903107846.

Waters, T. E. A. (2014). Relations between the functions of autobiographical memory and psychological wellbeing. *Memory* 22, 265–275. doi:10.1080/09658211.2013.778293.

Wehr, M., Hostick, U., Kyweriga, M., Tan, A., Weible, A. P., Wu, H., Wu, W., Callaway, E. M., and Kentros, C. (2009). Transgenic silencing of neurons in the mammalian brain by expression of the allatostatin receptor (AlstR). *J. Neurophysiol.* 102, 2554–2562. doi:10.1152/jn.00480.2009.

White, R. (1982). Memory for personal events. *Human Learning*, 171–183.

Wit, S. de, Watson, P., Harsay, H. A., Cohen, M. X., Vijver, I. van de, and Ridderinkhof, K. R. (2012). Corticostriatal Connectivity Underlies Individual Differences in the Balance between Habitual and Goal-Directed Action Control. *J. Neurosci.* 32, 12066–12075. doi:10.1523/JNEUROSCI.1088-12.2012.

Yin, H. H., and Knowlton, B. J. (2004). Contributions of striatal subregions to place and response learning. *Learn. Mem.* 11, 459–463. doi:10.1101/lm.81004.

Youmans, R. J., and Ohlsson, S. (2008). How practice produces suboptimal heuristics that render backup instruments ineffective. *Ergonomics* 51, 441–475. doi:10.1080/00140130701685659.

Zola-Morgan, S., and Squire, L. R. (1985). Medial temporal lesions in monkeys impair memory on a variety of tasks sensitive to human amnesia. *Behav. Neurosci.* 99, 22–34.

BIOGRAPHY

Robert Gardner received a Bachelor of Science degree in psychology from The Pennsylvania State University in 2003.