Spring 2015

Behavioral and neurophysiological investigations of short-term memory in primates

James Bigelow
University of Iowa

BEHAVIORAL AND NEUROPHYSIOLOGICAL INVESTIGATIONS OF SHORT-TERM MEMORY IN PRIMATES

by

James Bigelow

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology in the
Graduate College of The University of
Iowa

May 2015

Thesis Supervisor: Associate Professor Amy Poremba

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

James Bigelow

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Psychology at the May 2015 graduation.

Thesis Committee:

Amy Poremba, Thesis Supervisor

John Freeman

A. Kim Johnson

Ryan LaLumiere

Ed Wasserman

Dedicated to mom (Vickie), who inspired independence, creativity, and lifelong learning

Unfortunately, nature seems unaware of our intellectual need for convenience and unity, and very often takes delight in complication and diversity.

Santiago Ramón y Cajal
Nobel Lecture, December 12, 1906

ACKNOWLEDGMENTS

This dissertation is the result of collaboration, training, and support from more individuals and institutions than can be comprehensively acknowledged within the limited space here. I only briefly mention a select few whose contributions were especially important and/or essential to the final dissertation manuscript.

First, I am indebted to my thesis adviser, Dr. Amy Poremba, who encouraged me to do my best work, provided countless resources and opportunities for my development as a researcher and teacher, and absorbed steep costs in making the projects reported herein possible. I also express gratitude to the other faculty who were part of the PhD program, especially Dr. Ed Wasserman, Dr. John Freeman, Dr. Ryan Lalumiere, Dr. Kim Johnson, Dr. Jason Radley, and Dr. Mark Blumberg, for their key support and training through their roles as instructors, teaching supervisors, and committee members. I thank my labmates and fellow students, including Bree Rossi, Ryan Opheim, Dr. Damon Ng, Dr. Bethany Plakke, Iva Zdilar, and Jessie Bowden, who provided assistance, training, and importantly, comic relief from daily routine.

Several individuals need to be recognized for sharing their technical expertise, sometimes voluntarily, in designing and constructing hardware, software, and code. Alex Kirillov patiently responded to questions about programming, Mike Thornburg routinely created or repaired custom electronics, Keith Miller constructed numerous devices and structures, and Steven Bohm and Dr. Eric Mallet endured long hours of software design and testing.

My family members and friends were an enormous source of support to me. I thank them for their encouragement and for taking interest (or at least humoring me by feigning interest) in my work. Dad (Bill), Mel, Shane, Amelia, Norah, Dan, Brit, Blithe, Poet, Ben, Joe, Jessie, Steph, Joey, Ryn: thank you for regular calls, emails, photos, videos, and visits. Dad deserves extra credit for financial support, not to mention providing an enriched upbringing. Among this group, I especially thank and profoundly appreciate Mouna, my best friend, for her truly indispensable support, inspiring and uplifting presence, and for sharing (enduring) many late nights, weekends, and missed holidays with me in the lab.

Last but not least, I recognize financial support from the National Institutes of Health and The University of Iowa, including the Ballard and Seashore Dissertation Fellowship fund sponsored by the Graduate College. This work would not have been possible without these funding sources.

ABSTRACT

Detecting and interpreting sensory events, and remembering those events in in the service of future actions, forms the foundation of all behavior. Each of these pillars of the so-called "perception-action cycle" have been topics of extensive inquiry throughout recorded history, with philosophical foundations provided by early BCE and CE periods (especially during the Classic and Renaissance eras) leading to intensive empirical study in the twentieth and twenty-first centuries. Such experiments have described detailed (but incomplete) behavioral functions reflecting perception and memory, and have begun to unravel the extraordinarily complex substrates of these functions in the nervous system. The current dissertation was motivated by these findings, with the goal of meaningfully extending our understanding of such processes through a multi-experiment approach spanning the behavioral and neurophysiological levels. The focus of these experiments is on short-term memory (STM), though as we shall see, STM is ultimately inseparable from sensory perception and is directly or indirectly associated with guidance of motor responses. It thus provides a nexus between the sensory inputs and motor outputs that describe interactions between the organism and environment.

In Chapter 2, previous findings from nonhuman primate literature describing relatively poor performance for auditory compared to visual or tactile STM inspired similar comparisons among modalities in humans. In both STM and recognition memory paradigms, accuracy is shown to be lowest for the auditory modality, suggesting commonalities among primate species. Chapters 3–5

examined STM processing in nonhuman primates at the behavioral and neurophysiological levels. In Chapter 3, a systematic investigation of memory errors produced by recycling memoranda across trials (proactive interference) is provided for the understudied auditory modality in monkeys. Such errors were ameliorated (but not completely eliminated) by increasing the proportions of unique memoranda presented within a session, and by separating successive trials by greater time intervals. In Chapter 4, previous results revealing a human memory advantage for audiovisual events (compared to unimodal auditory or visual events) inspired a similar comparison in monkeys using a concurrent auditory, visual, and audiovisual STM task. Here, the primary results conformed to *a priori* expectations, with superior performance observed on audiovisual trials compared to either unimodal trial type. Surprisingly, two of three subjects exhibited superior unimodal performance on auditory trials. This result contrasts with previous results in nonhuman primates, but can be interpreted in light of these subjects' extensive prior experience with unimodal auditory STM tasks. In Chapter 5, the same subjects performed the concurrent audiovisual STM task while activity of single cells and local cell populations was recorded within prefrontal cortex (PFC), a region known to exhibit multisensory integrative and memory functions. The results indicate that both of these functions converge within PFC, down to the level of individual cells, as evidenced by audiovisual integrative responses within mnemonic processes such as delay-related changes in activity and detection of repeated versus different sensory cues. Further, a disproportionate number of the recorded units exhibited such mnemonic processes on audiovisual trials, a finding that corresponds to the superior behavioral performance on these trials. Taken

together, these findings reinforce the important role of PFC in STM and multisensory integration. They further strengthen the evidence that "memory" is not a unitary phenomenon, but can be seen as the outcome of processing within and among multiple subsystems, with substantial areas of overlap and separation across modalities. Finally, cross-species comparisons reveal substantial similarities in memory processing between humans and nonhuman primates, suggesting shared evolutionary heritage of systems underlying the perception-action cycle.

PUBLIC ABSTRACT

Evolution has endowed us with the remarkable ability to "remember" or store information that is not directly available to our senses. Though we may take it for granted, memory is fundamental to our daily lives, and more broadly, enables adaptive behavior and survival throughout the animal kingdom.

Given its general relevance to the human condition – and the debilitating conditions associated with memory loss – scientists have ambitiously pursued a detailed understanding of its functional properties and underlying biological mechanisms. Progress made largely within the past century has revealed memory as an extraordinarily complex phenomenon, orchestrated by networks of interacting brain regions, each comprising vast populations of individual cells with specialized functional roles.

Impressive though these discoveries may be, they amount to having assembled several thousand pieces of a billion-piece puzzle. The current work was inspired by these existing fragments, and brings several new pieces to the table.

One of the consistent findings in the current studies is that our ability to store information is dependent upon sensory modality. Thus, memory appears to be better for images than sounds, and even better for images and sounds presented together than for either alone. Physiological recordings within the frontal lobe revealed individual cells and local cell populations specialized for integrating such crossmodal information in memory, as well as other functions such as linking sensory events and behavioral choices separated by time. It is hoped that these findings will contribute to a more comprehensive portrait of memory against which pathologies can be meaningfully interpreted.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2-DG	2-[^{14}C]deoxyglucose
A+V	Audiovisual
ANOVA	Analysis of variance
Aud	Auditory
dlPFC	Dorsolateral prefrontal cortex
DMS	Delayed matching-to-sample
fMRI	Functional magnetic resonance imaging
FDR	False discovery rate
FR	Firing rate
ISI	Interstimulus interval
IT/ITC	Inferior temporal cortex
ITI	Intertrial interval
MD	Mediodorsal nucleus of the thalamus
M-NM	Match-nonmatch
MUA	Multi-unit activity
ODR	Oculomotor delayed response
PET	Positron emission tomography
PFC	Prefrontal cortex
PI	Proactive interference
SEM	Standard error of the mean
SER	Sensory-evoked response
STM	Short-term memory
STG	Superior temporal gyrus

STS	Superior temporal sulcus
SUA	Single-unit activity
U_{max}	Unimodal condition with maximum response
Vis	Visual
vlPFC	Ventrolateral prefrontal cortex

Chapter 1: General background and introduction

1.1 Short-term memory

Behaviorally relevant sensory information is often available for only a brief amount of time, and is usually encountered among a background of irrelevant information. Thus, adaptive behavior frequently depends on the ability to selectively retain relevant information that is no longer available in the sensory environment. The ability to maintain neural representations of information in the absence of direct stimulation, or short-term memory (STM), carries clear ecological advantages, and has therefore been considered the brain's "evolutionarily most significant achievement" (Goldman-Rakic, 1995) and the "heart of intelligent behavior" (Nęcka, 1992). In humans, STM capabilities have been associated with individual differences in attention, executive function, general intelligence, reading ability, and language comprehension (Baddeley, 2003), which are disrupted in many neurological disorders (Becker, 1988; Lewis et al., 2003; Litvan et al., 1988; Park and Holzman, 1992). For these reasons, understanding STM and its underlying neural circuitry has been a major research focus throughout the history of psychology and neuroscience.

One of the most commonly employed tests in behavioral and neural studies of STM is the delayed matching-to-sample (DMS) task, or one of its many variations (D'Amato, 1973; Medin et al., 1976; van Hest and Steckler, 1996). The earliest documented version of matching-to-sample (without a memory delay) was used by Nadie Kohts, a Russian primate researcher, to study visual object discrimination in the chimpanzee (see Yerkes & Petrunkevitch, 1925). Other

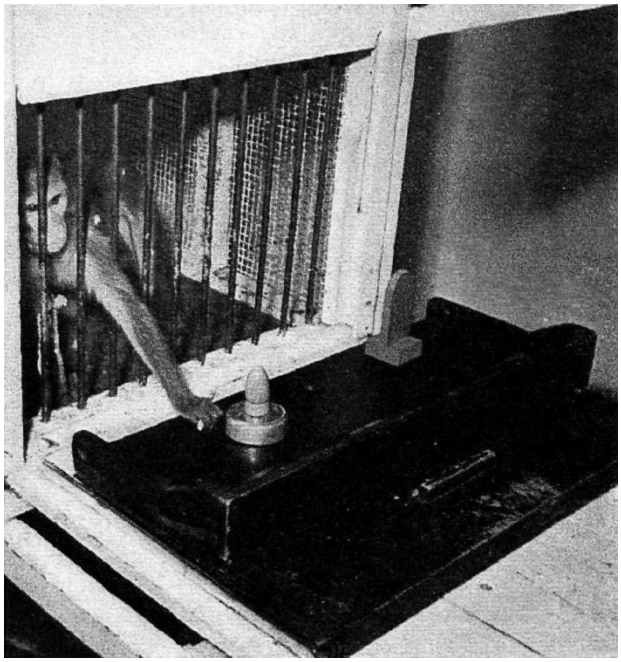

Figure 1. Photograph depicting a monkey subject performing a nonspatial delayed matching-to-sample task. The monkey observed as an experimenter placed a food reward in a well associated with one of two objects. A screen was then lowered in front of the cage to occlude the subject's view for the duration of a retention interval, during which time the spatial location (left or right) of the rewarded object was randomly switched. The screen was then raised and the monkey was rewarded only if the correct object was chosen (regardless of spatial location). In spatial versions of the task, the subject is rewarded for choosing the food well in the correct location. Adapted from Harlow and Dagnon (1943).

pioneering primate researchers such as Harry Harlow, Karl Pribram, and Mortimer Mishkin adopted similar or modified paradigms shortly thereafter to study various forms of visual discrimination and STM and its neural substrates in apes and monkeys (Harlow and Bromer, 1938; Harlow and Dagnon, 1943; Mishkin & Delacour, 1975; Mishkin & Pribram, 1956; Pribram and Mishkin, 1955). In a typical DMS task, subjects observed an experimenter placing a food reward in one of two wells in front of the cage (Figure 1). A screen was then lowered in front of the cage, occluding the subject's view for the duration of a retention interval, after which it was again raised. For spatial DMS tests, the subject was rewarded for choosing the location of the well in which the reward was initially placed. For nonspatial DMS tests, the subject was

rewarded for choosing the correct object covering the well (the location of which was randomly switched during the retention interval). Although modern DMS tasks use computer programs and hardware to precisely control stimulus presentations and measure response latencies, they rely on the same fundamental task structure in which memoranda are followed by a retention interval, after which a behavioral choice must be made based on STM. Even the more complex tasks used to study human STM, which may present multiple memoranda simultaneously or across time, can be seen as a more sophisticated variation of the classic DMS paradigm (e.g., Baddeley, 1992; Luck and Vogel, 1997; Sternberg, 1966; Wright et al., 1985). Thus, although each STM test differs in terms of complexity and specific task contingencies, subjects are invariably required to retain sensory information in order to guide future actions.

Inherent in the structure of most DMS tasks is the requirement that subjects continuously update representations of the pertinent memoranda on a trial-to-trial basis. In other words, not only must subjects respond on the basis of memoranda presented on the current trial, they also must *not* respond on the basis of memoranda from previous, irrelevant trials (D'Amato, 1973; Wright, 2006; Wright et al., 1986). The cost of failing to ignore irrelevant stimuli from previous trials was demonstrated in some of the earliest studies of STM in the chimpanzee. Yerkes and Yerkes (1928) first reported that chimpanzees had great difficulty learning a nonspatial DMS task in which visual characteristics such as color served as cues, with only one of four subjects successfully learning the task. However, Hayes and Thompson (1953) noted that these experiments recycled the same two stimuli as memoranda throughout the task, and suggested that the poor performance may have been attributable to the subjects confusing the correct choice between trials. They tested this idea using a similar nonspatial DMS task that presented new stimuli for each trial, and found that each of three subjects easily learned the task. Subsequent

studies of visual STM in humans and other animals have similarly reported a negative relationship between DMS performance and the degree to which stimuli are recycled across trials (Grant, 1975; Hartshorne, 2008; Mishkin and Delacour, 1975; Overman and Doty, 1980; Wright et al., 1986).

The finding that stimuli presented on previous trials can bias responses on the current trial is referred to as proactive interference (PI), inasmuch as memory processing at a given time interferes with memory processing at a subsequent time. PI may not have been deliberately included by the experimenters in many of the initial tests of STM (Nissen et al., 1938; Yerkes and Yerkes, 1928), but it was eventually recognized as a predominant cause of mnemonic errors in both human and animal experiments (Hayes and Thompson, 1953; Keppel and Underwood, 1962; Mishkin and Delacour, 1975; Underwood, 1957; Wright et al., 1986). Indeed, more recent studies have shown that individual variation in STM capabilities can be predicted by measuring susceptibility to PI (May et al., 1999; Whitney et al., 2001). In other words, performance on most STM tasks is as much a reflection of subjects' ability to filter irrelevant information as it is their ability to retain relevant information. Far from being an undesirable contaminant of an otherwise pure index of mnemonic failure, PI can be viewed as an ecologically relevant aspect of STM tasks, since real-world STM demands require that relevant cues be extracted and retained from a stream of irrelevant information, and moreover, because stimuli which are temporarily important may quickly become insignificant in a dynamic contextual environment. Even tasks in which trial-unique memoranda are presented may still produce some PI on the basis of perceptual similarity among nonidentical stimuli ("item-nonspecific PI"; Craig et al., 2013; Jitsumori et al., 1989; Reynolds & Medin, 1981; Visscher et al., 2009; Wickens, 1970).

In summary, STM is central to many of the flexible, adaptive behaviors observed in

humans and many other animals. During the past century, STM has been studied intensively using the DMS paradigm and its derivatives. Investigating mnemonic failure as a function of retention interval continues to be an important focus of STM research. In addition, STM errors can be frequently accounted for by PI, which occurs when responses are maladaptively biased by irrelevant stimuli presented on previous trials. Thus, two central requirements inherent in most DMS tasks are retention of relevant information and resolution of PI from previous trials. Both requirements are thought to reflect real-world STM demands.

1.2 The central role of the lateral prefrontal cortex in short-term memory

The prefrontal cortex (PFC) attracted the interest of scientists who first sought to identify brain areas that were important for STM (Jacobsen, 1935; Campbell and Harlow, 1945; Meyer et al., 1951; Pribram et al., 1952; Spaet and Harlow, 1945). Earlier experiments by Franz (1907) had shown that bilateral lesions of the PFC produced deficits in attention, perceptual association, learned motor habits, and problem solving in animals that were learning sequences of behaviors for food rewards. Jacobsen (1935) and later Harlow (Campbell and Harlow, 1945; Meyer et al., 1951; Spaet and Harlow, 1945) provided the first direct evidence that the PFC was involved in STM. In these experiments, bilateral lesions of the PFC in monkeys produced severe deficits in DMS task performance. It was subsequently shown that lesions of the lateral, but not orbitomedial divisions of the PFC, were necessary and sufficient for STM deficits (Meyer et al., 1951; Pribram et al., 1952).

Studies of anatomical connectivity have borne out the distinction between lateral and

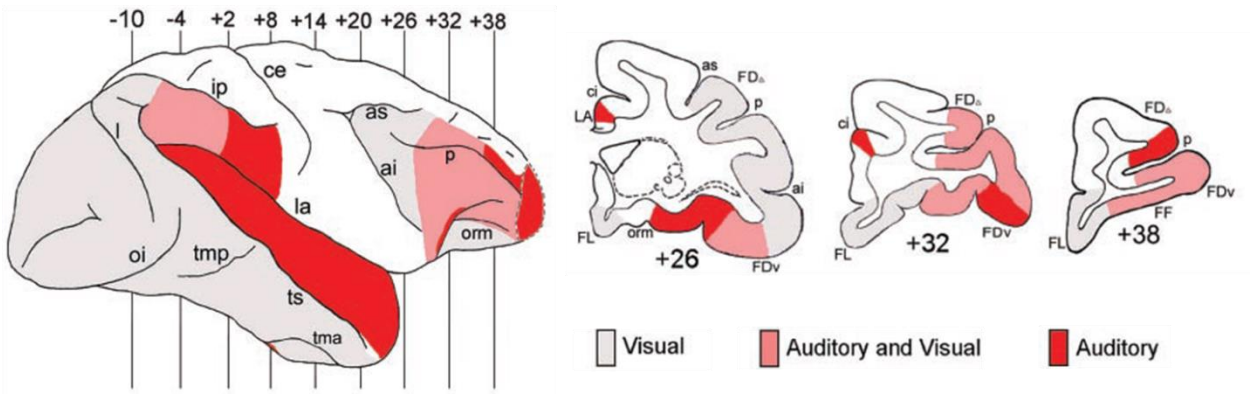

Figure 2. Comparison of visual and auditory cortical areas in the macaque brain revealed by 2-DG imaging. The colored areas represent cortical activation by passive exposure to sounds or images in an intact hemisphere compared to a deafferented hemisphere. A broad section of visual and auditory overlap in the lateral PFC is evident in the lateral surface view (left) and coronal sections (right). Numbers indicate distance in millimeters from the interaural plane.

orbitomedial PFC (Barbas et al., 2002; Porrino et al., 1981). The orbitomedial PFC is predominantly connected with the limbic structures (Carmichael and Price, 1995a), including the amygdala (Barbas and De Olmos, 1990; Porrino et al., 1981), cingulate cortex (Vogt and Pandya, 1987), hypothalamus (Ongür et al., 1998; Rempel-Clower and Barbas, 1998), hippocampus, entorhinal, perirhinal, and parahippocampal cortices (Barbas and Blatt, 1995; Cavada et al., 2000; Martin-Elkins and Horel, 1992; Rosene and Van Hoesen, 1977; Van Hoesen et al., 1975), as well as olfactory and gustatory cortex (Cavada et al., 2000; Rolls, 1989). Thus, functional specializations of the orbitomedial PFC include processing the emotional significance of stimuli, behavioral inhibition, and regulation of social behavior (Barbas et al., 2002; Zald and Andreotti, 2010). On the other hand, the lateral PFC is primarily connected with motor structures as well as visual (Jones and Powell, 1970; Petrides and Pandya, 2002a), auditory (Hackett et al., 1999; Romanski et al., 1999a; Romanski et al., 1999b), and somatosensory cortical areas (Preuss and Goldman-Rakic, 1989). A neuroimaging study by Poremba et al. (2003) revealed extensive overlap in the area of the lateral PFC responsive to visual and auditory stimulation (Figure 2),

and neuronal recording studies have demonstrated that many individual cells respond to both auditory and visual stimuli (Ito, 1982; Nelson and Bingall, 1973; Schechter and Murphy, 1975; Sugihara et al., 2006; Vaadia et al., 1986; Wollberg and Sela, 1980). Moreover, a subset of these neurons exhibit integrative responses (additive or superadditive enhancement or suppression) to combined audiovisual stimuli (Figure 3), which are thought to be important for audiovisual communication (Sugihara et al., 2006; Romanski and Averbeck, 2009). Prominent motor connections of the lateral PFC include the supplementary and premotor cortices (Goldman-Rakic, 1987; Lu et al., 1994), the striatum (Cavada et al., 2000; McGuire et al., 1991b; Yeterian and Pandya, 1991), and the cerebellum (Dum and Strick, 2003; Kelly and Strick, 2003). The lateral PFC is thus situated as an intermediary among multiple sensory inputs and motor output (Figure 4), and is therefore specialized for sensory integration and providing a temporal link between perception and action (Miller and Cohen, 2001). These functional specializations between lateral and orbitomedial PFC should be considered dominant but not exclusive, inasmuch as their anatomical distinctions are relative rather than absolute (Barbas et al., 2002), and because they are heavily interconnected (Barbas and Pandya, 1989).

The functions of the PFC have been revealed in part from human neuropsychological cases, and more recently, using neuroimaging technologies such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) (Fuster, 2008b, 2008c; Miller and Cummings, 1999; Shimamura, 2000; Wood and Grafman, 2000). However, because cellular recordings, precise lesions, and pharmacological manipulations are not feasible with human subjects, animal models are still widely used. Old-world monkeys (e.g., Macaca mulatta) have been popular animal models of PFC function ever since Jacobsen (1935) first reported STM deficits following ablation of the frontal lobes. The limited homology between human lateral

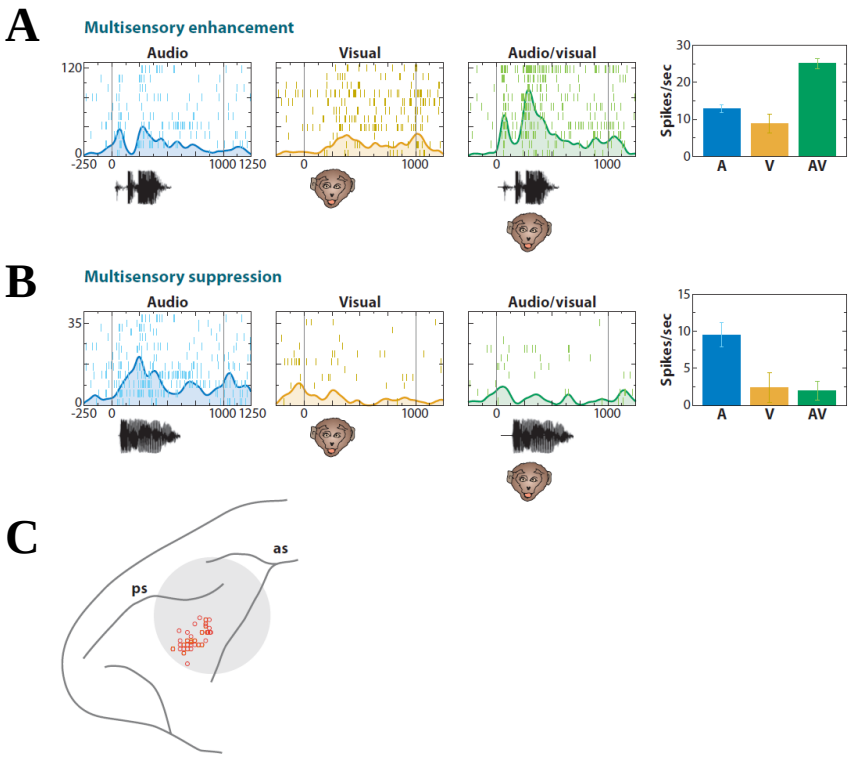

Figure 3. Audiovisual integration in single cells in ventrolateral prefrontal cortex (vlPFC). Raster and spike density plots are aligned to stimulus onset for a monkey vocalization alone (Audio, A), a monkey face alone (Visual, V), and both together (Audio/visual, AV). The cell in (A) exhibited a significantly enhanced response to the audiovisual stimulus, whereas the cell in (B) showed mild suppression. Summary bar graphs are shown on the right. The locations of multisensory neurons from this study are shown in (C). Adapted from Romanski and Averbeck (2009).

PFC and rodent PFC (Preuss, 1995; Uylings and Van Eden, 1990), and the extensive homology between human and nonhuman primate PFC (Figure 5) (Petrides, 2005; Petrides and Pandya, 1994, 1999) have been a major factors in the continued use of monkeys as model species of lateral PFC function.

1.2.1 Studies of visual short-term memory in the lateral prefrontal cortex

Jacobsen's (1935) discovery that bilateral frontal lesions produced severe deficits in delayed-response performance brought into focus the role of the PFC in visual STM. Harlow's lab extended these findings through a series of experiments showing that monkeys with frontal ablations were impaired on a variety tasks including spatial and nonspatial DMS and contingency reversals (Campbell and Harlow, 1945; Harlow and Dagnon, 1943; Spaet and Harlow, 1945). Subsequent research has shown that, regardless of additional task contingencies, the presence of a retention interval, or intratrial delay period, invariably predicts deficient performance in subjects with PFC lesions (Goldman and Rosvold, 1970; Goldman et al., 1971; Meyer et al., 1951; Treichler et al., 1971). Many additional experiments have further explored the nature of these deficits, attempting to detail the capabilities that are lost with frontal lesions and more precisely identify the functions carried out by specific regions of the PFC.

A follow-up study by Harlow and colleagues (Meyer et al., 1951) revealed that lesions restricted to the lateral surface of the prefrontal region in one hemisphere combined with extensive damage to the prefrontal region in the contralateral hemisphere were sufficient to produce impairment on a variety of STM tasks. This finding was elaborated by Pribram et al. (1952), who were among the first to directly examine the regional specificity of the frontal lesion deficit. Comparing baboons with dorsolateral and ventromedial resections to an intact control group, they reported that dorsolateral lesions produced far greater deficits in delayed-response performance than ventromedial lesions. Additional experiments using more selective lesions have indicated that the cortex surrounding the principal sulcus is particularly important for the delayed response and delayed alternation variations of the DMS task (Blum, 1952; Butters and

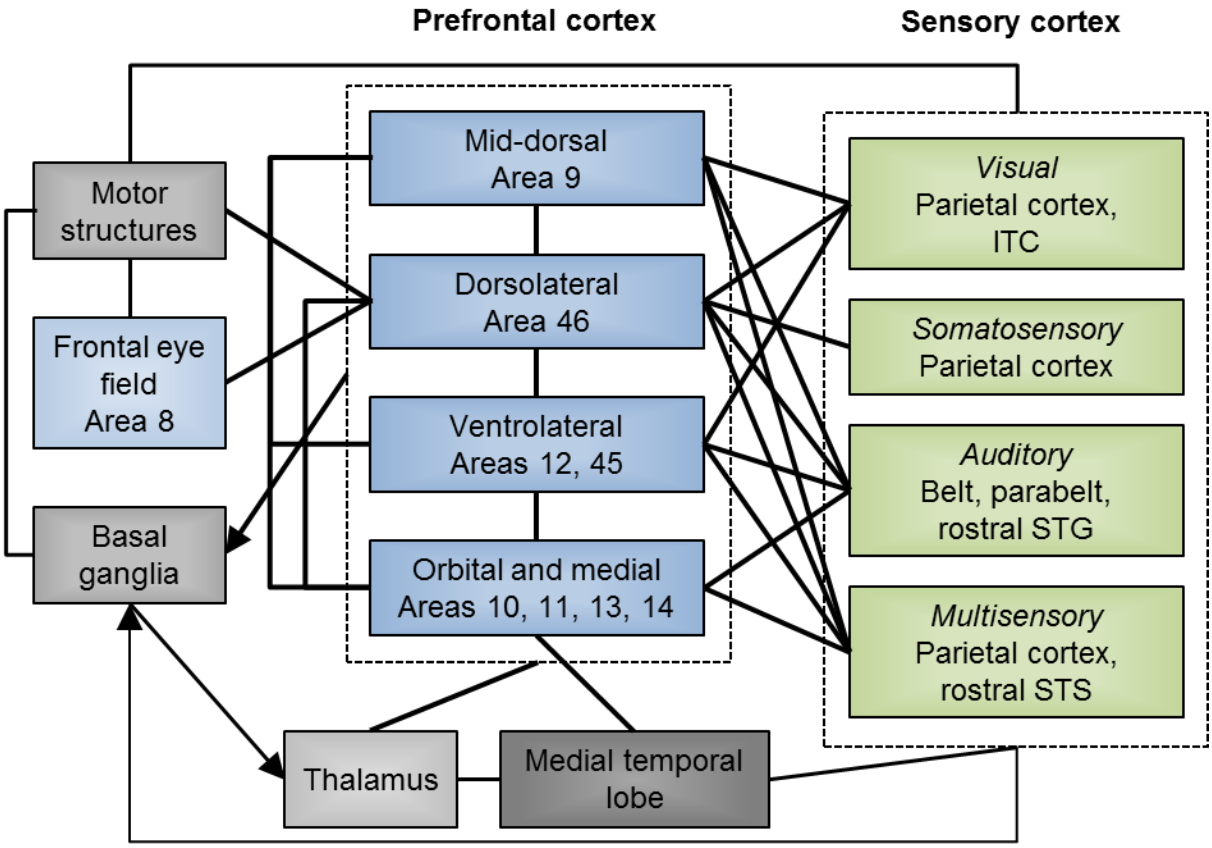

Figure 4. Schematic diagram of some of the prominent intrinsic and extrinsic connections of the primate PFC. Connections are reciprocal unless indicated by an arrow. Adapted from Miller and Cohen (2001).

Pandya, 1969; Goldman and Rosvold, 1970; Mishkin, 1957).

Although the first prefrontal lesion experiments reported that only bilateral lesions resulted in STM deficiencies (Jacobsen, 1935; Meyer et al., 1951), several later studies have demonstrated that unilateral lesions are sufficient to produce significant, if less extensive STM impairments (Warren et al., 1969; Warren and Nonneman, 1976). In these experiments, monkeys with unilateral lesions of the dorsolateral PFC were consistently impaired on a delayed response task compared to controls at short delays (0–5 s), and at long delays (20–40 s) performed at chance levels. These deficits were observed up to 11 months post operation, which was the latest point tested. Further, control subjects that had overlearned the task and later received unilateral

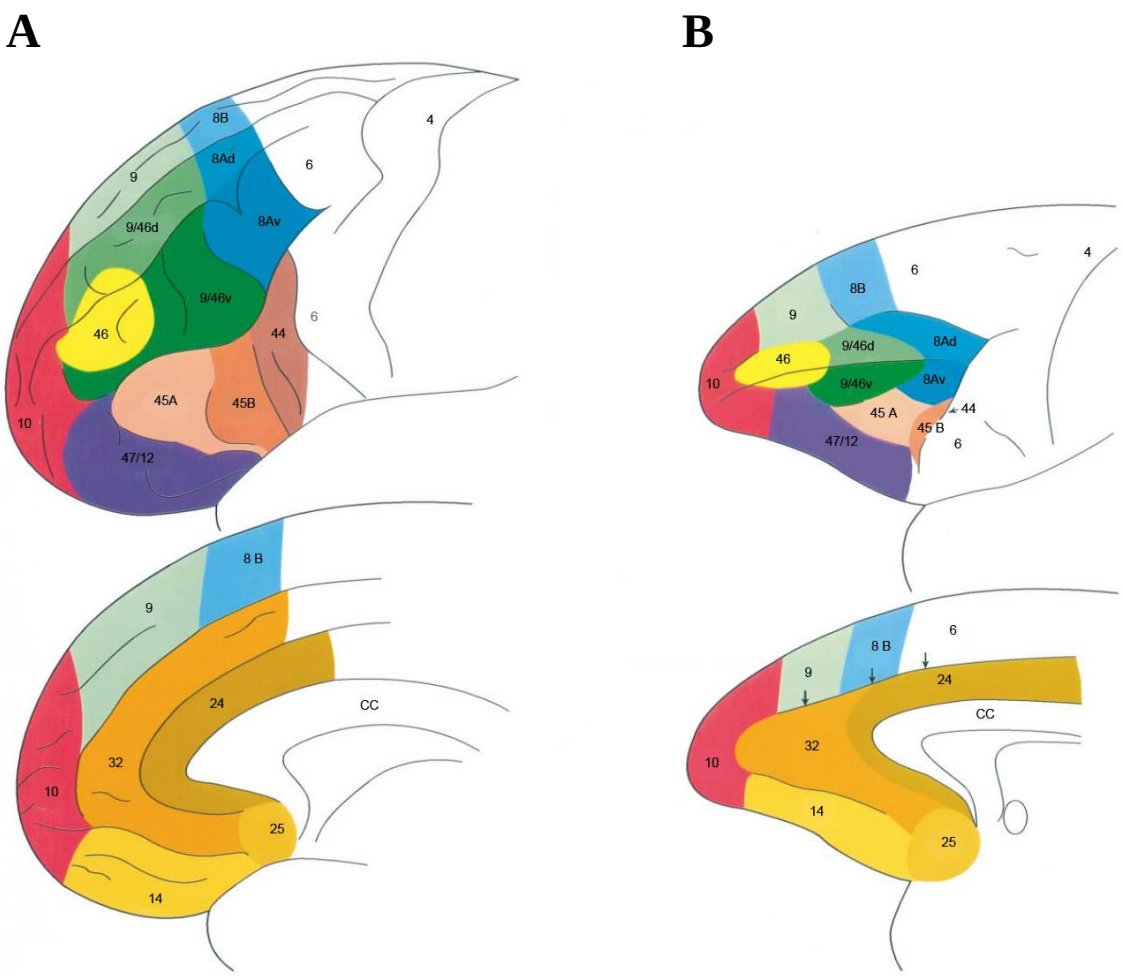

Figure 5. Comparison of the prefrontal cortex in humans (A) and macaques (B). The numbered colored areas designate cytoarchitectonically defined subdivisions of the cortex. The lateral surface is shown in the upper panels and the medial surface is shown in the lower panels. Adapted from Petrides and Pandya (1999).

lesions showed similar levels of impairment. The authors also noted that relatively large unilateral lesions produced greater STM deficiencies than relatively small lesions. This finding recalls Lashley's (1950) principle of mass action, which predicts that behavioral deficits will be correlated with the extent cerebral tissue damage. Further evidence for unilateral disruption of visual STM was provided by Weiskrantz et al. (1960), who showed that electrical stimulation applied to the principal sulcus region of either hemisphere was sufficient impair delayed alternation performance. Similarly, Fuster and Bauer (Bauer and Fuster, 1976; Fuster and Bauer,

1974) observed significant deficits in delayed response and DMS performance following unilateral cooling of the dorsolateral PFC in either hemisphere. As with permanent lesions, bilateral cooling or stimulation was more disruptive to performance than either hemisphere alone. Finally, Funahashi et al. (1993a) reported that unilateral lesions of the dorsolateral PFC significantly reduced monkeys' ability to remember visuospatial cues in the contralateral visual hemifield. Collectively, these experiments indicate that mnemonic processes that are disrupted in one hemisphere cannot be fully compensated for by the remaining hemisphere.

Visual STM deficits have also been observed following lesions of the mediodorsal (MD) nucleus of the thalamus (Isseroff et al., 1982). These studies are relevant to understanding PFC function because the MD nucleus is the primary thalamic structure with which the lateral PFC is reciprocally connected (Barbas et al., 1991; Nauta, 1972; Siwek and Panya, 1991). Schulman (1964) created bilateral lesions of the MD thalamus in monkeys using small sources of radiation. Subjects with complete or near complete destruction were severely impaired on a visual DMS task compared to preoperative levels. Incomplete lesions also resulted in some impairment, although these animals improved following prolonged training. Isseroff et al. (1982) have also shown that monkeys with MD thalamic lesions exhibit spatial STM deficiencies, even though their accuracy for a visual pattern discrimination task without a memory delay was unaffected. These observations suggest that the feedback loop between the lateral PFC and the MD thalamus is important for maintaining representations of visually encoded information.

Fuster and Alexander (1971) provided the first neurophysiological evidence for the role of the PFC in visual STM by recording the activity of single cells near the principal sulcus in monkeys performing a DMS task. The most noteworthy finding of this experiment was that many cells showed an elevated firing rate following the initial cue period which was sustained

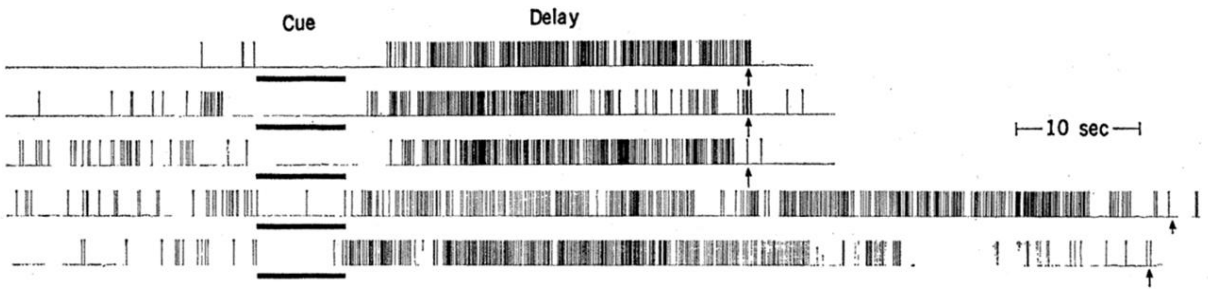

Figure 6. "Memory cells" in the lateral PFC show increased activity during the maintenance phase of a visual STM task. The vertical lines represent action potentials of a single cell recorded during five consecutive trials of a delayed response test. The period underscored in black indicates the initial cue presentation period. The first three trials have a retention interval of 32 s, which are followed by two trials with 67-s and 65-s retention intervals. The arrows indicate end of the retention interval, at which point the monkey made a behavioral response. Adapted from Fuster and Alexander (1971).

throughout the retention interval until the animal made a behavioral choice (Figure 6). This increased, tonic firing was sustained for even the longest delays tested, in excess of one minute. Fuster and colleagues have used the term "memory cells" to describe neurons that exhibit delay-related changes in firing rate because this activity is correlated with cue retention across time. This notion is further supported by several studies showing a direct relationship between behavioral performance and the level of activity in cells that show delay-related firing (Batuev et al., 1979; Fuster, 1973; Watanabe, 1986). Moreover, sustained firing during the retention interval does not occur in untrained animals (Fuster, 1973), suggesting that it does not simply reflect activity related to the offset of the stimulus.

Many subsequent studies have further characterized the nature of delay-related activity in a variety of STM tasks including nonspatial and spatial delayed response (Funahashi et al., 1993b; Kubota et al., 1974; Niki, 1974a), delayed alternation (Kubota and Niki, 1971; Niki, 1974b, 1974c), and other DMS tasks (Cromer et al., 2011; Miller et al., 1996). Cells that show increased activity during the retention period of STM tasks have been found in all areas of the lateral PFC, but are most concentrated in the area surrounding the principal sulcus. A variety of

patterns of delay-related changes in neuronal activity have been reported, including increases or decreases in firing rate that remain stable throughout the delay, steadily increasing or decreasing firing rates, as well as activity that is only elevated or inhibited for a fraction of the delay period (Fuster, 1973; Shafi et al., 2007). Some cells have shown activity increases during the delay period in excess of 500% of the baseline firing rate, and others become virtually inactive. The distribution of cells exhibiting memory related activity and the variability of this activity have been interpreted as evidence that stimulus retention is achieved by distributed, multifunctional networks (Shafi et al., 2007).

Neurophysiological recordings outside the PFC suggest that STM may depend on communication among the lateral PFC and other areas of the brain. Shortly after the initial discovery of "memory cells" in the lateral PFC, similar delay-related activity was recorded in the MD nucleus of the thalamus (Fuster and Alexander, 1973). These results, combined with anatomical (Barbas et al., 1991) and lesion studies (Isseroff et al., 1982; Schulman, 1964), emphasize the importance of mutual feedback between the lateral PFC and MD thalamus during STM. Neurons in the inferior temporal cortex (IT) have also been shown to exhibit delay-related increases in firing rate during visual STM (Fuster et al., 1985; Fuster and Jervey, 1981, 1982). Combining cooling inactivations of the IT with neurophysiological recordings in the PFC, Fuster et al. (1985) reported a decrease in delay-related activity in the PFC as well as impaired DMS performance. The same results were found in IT neurons when the PFC was inactivated, indicating that mutual influences between the two cortical areas are important for STM (see also Constantinidis and Procyk, 2004). Miller et al. (1996) have compared single-cell activity between the lateral PFC and the IT during a visual STM task which included a sample stimulus, followed by a varying number of nonmatching distracters, followed by a matching test stimulus.

Whereas delay-related activity in the IT returned to baseline after the presentation of a distracter, neurons in the PFC maintained a sustained increase or decrease in firing rate throughout each delay period in spite of the distracter stimuli. Neurons in both the IT and PFC often show differential responses—usually enhancement, but occasionally suppression—to matching versus nonmatching test stimuli (Cromer et al., 2011; Miller and Desimone, 1994; Miller et al., 1996; Miller et al., 1991, 1993; Rainer et al., 1999). Match enhancement and suppression have been interpreted as an additional neural mechanism that may underlie the recognition phase in STM tasks. The fact that these responses are observed in IT neurons despite interruption of delay-related activity by distracter stimuli has led to the suggestion that these responses may be influenced by the lateral PFC (Miller and Cohen, 2001). These results are consistent with a distributed network model of STM wherein the lateral PFC takes a central role in biasing representations of behaviorally relevant stimuli in sensory cortical areas.

In general, the findings from neuropsychological and neurophysiological studies of visual STM in nonhuman primates are compatible with the pattern of results described in neuropsychological and neuroimaging studies using human subjects. Thus, Bechara et al. (1998) reported that patients with lesions of the lateral PFC showed STM impairments, but did not differ from controls in a gambling task. In visual STM studies using fMRI, sustained increases in blood flow are observed in the lateral PFC during the retention interval (Curtis and D'Esposito, 2003), similar to the persistent changes in delay-related neuronal activity that have been well described in the monkey lateral PFC (Figure 7). As with neurophysiological recordings, these changes in activity are correlated with performance as well as memory load (Braver et al., 1997).

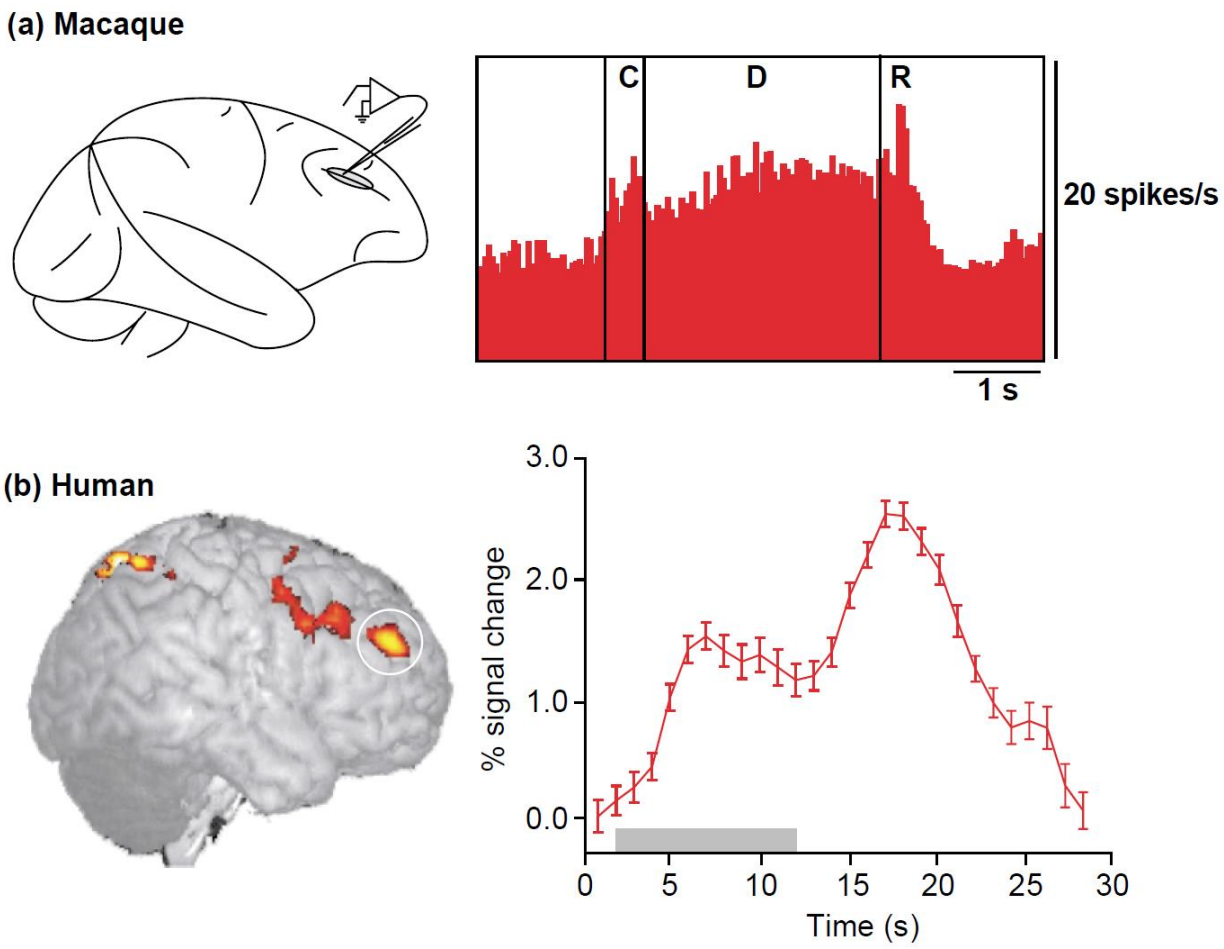

Figure 7. Activity in the monkey and human dlPFC during the retention interval of a spatial oculomotor delayed response (ODR) task. (a) Average of single-unit recordings from 46 neurons with delay period activity from the monkey dlPFC (area 46). C = cue; D = delay; R = response. (b) Significant maintenance-related activity (left) and average (±SE) fMRI signal (right) from right dlFPC (area 46; circled) in a human performing the ODR task depicted in Box 2, Fig. Ia. The grey bar represents the length of the delay interval. Notice how in both cases the level of dlPFC activity persists throughout the delay, seconds after the stimulus cue has disappeared. Figure and caption replicated from Curtis and D'Esposito (2003).

1.2.2 Studies of auditory short-term memory in the lateral prefrontal cortex

Comparatively few studies have investigated the role of the lateral PFC in auditory STM. This may be attributable, in part, to the long training times required for nonhuman primates to learn auditory DMS tasks (Cohen et al., 2005). Whereas monkeys require ~500 training trials to learn visual DMS tasks, they require ~15,000 trials to learn comparable auditory tasks (Fritz et al., 2005). Moreover, upon learning the auditory DMS rule, they are only capable of performing above chance at relatively short retention intervals (Fritz et al., 2005; Kojima, 1985), and are more susceptible to intratrial distracters (Scott et al., 2012). Whether a parallel deficit exists in human auditory STM remains an outstanding question (Munoz-Lopez et al., 2010).

In spite of these training challenges, several laboratories have begun to make progress in understanding the role of the lateral PFC in auditory STM. A few of these experiments have focused exclusively on auditory STM, and several others have used combined visual and auditory STM tasks. The available data indicate that the lateral PFC is crucial for auditory STM and that many of the same neurophysiological processes observed during visual STM are also seen during auditory STM. Further, a few studies have provided evidence that some individual cells in the lateral PFC are involved in STM for both sensory modalities.

The first experiment to address the role of the PFC in auditory STM compared performance on an auditory DMS task between groups of monkeys receiving lesions of the dorsolateral, midlateral, and ventrolateral PFC (Blum, 1952). The task presented either a buzzer or a bell, followed by a delay ranging from 5–30 s, after which the monkeys chose from one of two boxes. The buzzer signaled that a reward was located in the right-hand box and the bell

signaled reward on the left. All three lesion groups were severely impaired on this task, and monkeys with lesions of the midlateral PFC were most impaired. More recently, Sierra-Paredes and Fuster (2002) have demonstrated that inactivation of the dorsolateral PFC using cortical cooling disrupts short-term retention of auditory stimuli. The task in this experiment required monkeys to match a tone sample stimulus with a light test stimulus following a 0-s, 5-s, or 10-s retention interval. The impairment was correlated with the length of the retention interval in the lateral PFC, whereas cooling of a control area in the parietal cortex had no effect. Further, the monkeys did not show changes in reaction time between conditions, indicating that the cooling did not adversely affect general motor activity. These experiments provided the important first steps of showing that the lateral PFC is essential for normal auditory STM functioning.

The first neurophysiological investigation of auditory STM recorded from neurons in the dorsolateral PFC of monkeys performing a spatial auditory STM task (Joseph and Barone, 1987). The subjects fixated on a central point while an auditory stimulus consisting of 20 clicks per second was presented either on the right or the left side for one second. Following a delay, left and right key lights were illuminated and the monkeys were rewarded for making a saccade to the light position cued by the sound. As with similar visuospatial tasks, many cells showed an increased firing rate during the delay period. For some neurons, this delay-related activity was selective for either the right or left cue position, but others did not exhibit spatial preference. Together with results from visuospatial studies, these findings suggest that the dorsolateral PFC is instrumental in retaining spatial information cued by either visual or auditory stimuli.

Fuster and colleagues (Bodner et al., 1996; Fuster et al., 2000) subsequently showed that neurophysiological activity in the lateral PFC is associated with retention during nonspatial auditory STM tasks. In these experiments, subjects were presented auditory sample cues

followed by a choice of visual test stimuli. The monkeys were trained to choose a green light if a low tone had been presented and a red light if a high tone had been presented. As with previous visual experiments, some cells were observed to increase in firing rate during the memory delay following the auditory sample stimuli. For many cells, the choice of a "matching" visual cue elicited a significantly enhanced response, similar to what has been observed in unimodal DMS tasks. The authors interpreted this result as evidence for cross-modal and cross-temporal associations at the single-cell level in the lateral PFC. Further, this finding implies that match enhancement may reflect the behavioral significance of the test stimulus rather than its repeated presentation.

An important caveat regarding the auditory "memory cells" described in these studies (Bodner et al., 1996; Fuster et al., 2000) is that, owing to the nature of the DMS task, many cells apparently acquired and exhibited audiovisual associative responses. It has been shown that delay-related activity in PFC neurons can reflect the properties of the anticipated test stimulus as well as the sample stimulus (Rainer et al., 1999). Also, Gibson and Maunsell (1997) found that, in IT, an auditory sample stimulus could evoke delay-period responses when the animals expected to respond to a visual test stimulus. Thus, the implications of auditory-to-visual DMS studies for auditory STM are limited, inasmuch as the observed responses might have been partially driven by the learned auditory/visual associations.

Several more recent studies have examined PFC neuronal activity during purely auditory STM tasks. Plakke et al. (2013) examined single-cell activity in the lateral PFC during a nonspatial DMS task that presented a variety of auditory stimuli, ranging from pure tones to complex vocalizations. These cues were pseudorandomly ordered in the sample and test positions, and were separated by a 5-s retention interval. Many cells exhibited delay-related

activity, although sustained changes in firing rate throughout the entire delay period were observed less frequently than has typically been reported in the visual STM literature. Comparable to findings in visual DMS tasks, population analyses revealed that matching stimuli elicited enhanced responses compared to nonmatching stimuli. Match enhancement was also observed by Cohen and colleagues (Lee et al., 2009; Russ et al., 2008) in lateral PFC neurons during an auditory *same*/*different* task. In these experiments, a trial began with four repetitions of either the human-spoken word "bad" or "dad", separated by an interstimulus interval of approximately 1600 ms. An intermediate morph was then presented and the monkey had to decide whether it matched the sample stimuli (60–100% morphs) or not (0–40% morphs). Enhanced neuronal responses elicited by the test stimuli tended to reflect the monkeys' perceptual choices rather than the actual acoustic properties of the stimuli, implying decision-related processes in addition to basic stimulus analysis.

A few additional studies have examined the activity of single PFC neurons during STM for both visual and auditory stimuli. Kikuchi-Yorioka and Sawaguchi (2000) provided the first evidence that single neurons in the lateral PFC support STM for both visual and auditory stimuli. Single cells were recorded around the principal sulcus and prearcuate cortex during an oculomotor delayed response task that presented either a light or a tone in one of four locations. Consistent with previous visuospatial and audiospatial STM experiments, many cells exhibited spatially selective changes in activity during the 3-s delay period. The most noteworthy finding of this study was that, although the some cells exhibited modality-specific activity, 57% showed delay-related activity for specific cue locations regardless of the sensory modality. These results have been corroborated by a recent study that compared the activity of single cells near the principal sulcus during a concurrent visual and auditory spatial DMS task (Artchakov et al.,

2007). Each trial presented either a sample tone or light on the left or right side of a central fixation point. Following a 2- to 3-s delay, a test tone or light was presented and the subjects were rewarded for correctly identifying whether or not it had occurred on the same side as the sample. Each trial was unimodal (i.e., both cues were either auditory or visual) and the trial types occurred in random order. The results revealed that some neurons showed an increased firing rate following only visual or auditory cues, but others showed delay-related changes in activity during both trial types. All of the cells that were non-selective for stimulus modality were also non-selective for spatial location. These results suggest that some cells in the lateral PFC mediate the retention of behaviorally relevant stimuli regardless of their specific sensory or spatial characteristics.

Artchakov et al. (2009) reported the results of an additional experiment investigating the impact of irrelevant distracters on neuronal activity during visual and auditory spatial STM. They used the same spatial DMS task that was used in their earlier experiment (Artchakov et al., 2007) except that distracter stimuli interrupted the retention interval for a randomized portion of the trials. The auditory distracter was a tone that briefly alternated between the left and right speakers, and the visual distracter consisted of flashing lights surrounding the central fixation point. Although the distracters interrupted delay-related activity in some cells, other cells maintained spatially tuned delay-period activity in spite of distracters. This outcome echoes the results of Miller et al. (1996), who showed that the persistent delay-related activity in PFC neurons is undeterred by irrelevant stimuli during nonspatial visual STM. Artchakov and colleagues further pointed out that some cells showed spatially tuned delay activity only during trials that presented a distracter. They suggested that these cells may be latent during relatively easy conditions, but are recruited to perform more demanding cognitive tasks.

It is worth noting that the studies that have compared visual and auditory STM in the lateral PFC have generally reported that fewer cells show selective task-related responses for auditory stimuli than visual stimuli. In their sample of 190 neurons, Kikuchi-Yorioka and Sawaguchi (2000) reported that 17% responded only to auditory-cued locations, 53% were visual specific, and 57% responded to either stimulus type. Artchakov et al. (2007) recorded a total of 360 neurons from the lateral PFC and found that the majority of task-related responses were elicited by visual cues only (67%). A smaller portion of neurons responded only to tones (19%) and another subset was responsive to either cue modality (14%). Likewise, Watanabe (1992) compared neuronal activity between visual and auditory cues during a delayed association task. Of 289 neurons recorded near the principal sulcus and prearcuate cortex, only 4.5% responded selectively to auditory stimuli, 47% were visual selective, and 48.5% were bimodal. In all three of these experiments, it was also reported that the monkeys' performance was better for the visual than the auditory trials. These results raise the possibility of a relationship between the relatively poor performance of the monkeys during auditory tasks and the underrepresentation of auditory task-related responses in the PFC. However, the auditory cues in all three of these studies were pure tones, which may have negatively biased the responses (Gifford et al., 2005; Romanski et al., 2005; Romanski and Goldman-Rakic, 2002; Wollberg and Sela, 1980). Also, many of the cells reported by Kikuchi-Yorioka and Sawaguchi (2000) and Watanabe (1992) were recorded in the prearcuate cortex, where visual anatomical inputs and physiological responses are dominant. With these caveats in mind, future experiments comparing neuronal responses in PFC areas of substantial visual and auditory overlap (Poremba et al., 2003) during STM for complex visual and auditory cues may be helpful in clarifying this issue.

Collectively, the foregoing results indicate that lateral PFC has a pivotal role in

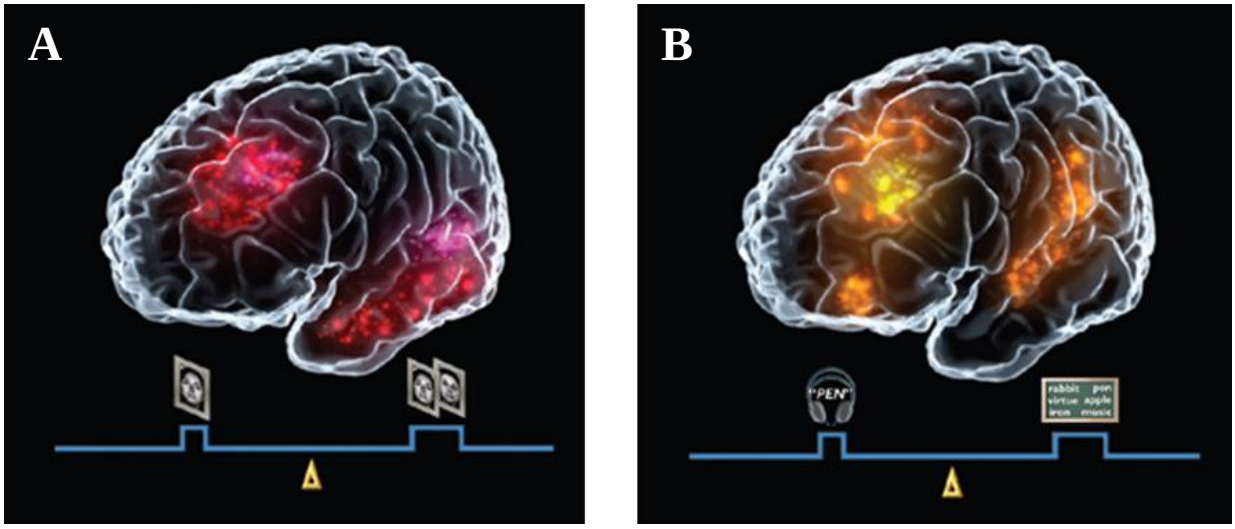

Figure 8. The lateral prefrontal cortex is active during visual and auditory working memory in humans. The brain images summarize the results of a meta-analysis of human imaging studies using visual (A) and auditory (B) working memory tasks. The trial progression is depicted below the activation maps: a sample cue is briefly presented, followed by a memory delay, after which the subject must select the cue from multiple test stimuli. The yellow triangle indicates the time during the task at which the imaging data were analyzed. Above-baseline activations of the lateral surface and gyri are shown in red (A) and orange (B), and activations of the medial surface and sulci are shown in pink (A) and yellow (B). Significant activation of the lateral prefrontal cortex is evident for both tasks. Adapted from Fuster (2008c).

maintaining representations of behaviorally relevant auditory as well as visual information. Similar conclusions can be drawn from human neuroimaging studies of STM in which memoranda were presented in the auditory or visual modalities, or both. For example, in collaboration with the UCLA Laboratory of NeuroImaging, Fuster (2008c) conducted a meta-analysis of several dozen neuroimaging studies of visual and auditory STM, and graphically overlaid the areas that were commonly activated between experiments (Figure 8). Although the overlap is not absolute, the results clearly demonstrate above baseline activation in the lateral PFC during the retention phase for both visual and auditory STM. Activations of visual and auditory areas in the temporal lobe corroborate results from neurophysiological studies suggesting interaction between the PFC and sensory cortex during STM (Constantinidis and

Procyk, 2004; Fuster et al., 1985). Thus, although neuroimaging measures are not directly comparable to neurophysiological recordings (Logothetis, 2003), experiments using both techniques have provided complementary results which together strongly implicate the lateral PFC in visual and auditory STM.

1.3 Remaining questions

In light of the foregoing results, it can safely be concluded that the lateral PFC occupies a pivotal position in the neural network that collectively enables STM for visual and auditory stimuli. Beyond this basic conclusion, however, many questions remain. In particular, the existing literature has only "scratched the surface" of the pending questions surrounding the neural substrates of auditory STM. As such, many opportunities exist for replication of experiments using auditory stimuli that have been conducted in the visual literature, thus enabling comparison of the basic processes that support STM for each modality. Some of the possible experiments that fall into this category include focal lesions and temporary inactivations of the PFC, studies of functional interaction between the lateral PFC and sensory cortices using lesion and recording techniques, recording and lesion studies of the involvement of the MD thalamus, and comparisons among various DMS tasks. Among other things, these comparisons could potentially shed light on the behavioral dissociations between visual and auditory STM that have been reported in nonhuman primates.

Experiments combining lateral PFC recordings with concurrent visual and auditory STM tasks may be especially informative because they allow comparisons at the level of single

neurons. Yet only a small number of such experiments has been conducted. Although it is known that some cells in the lateral PFC support STM for simple visuospatial and audiospatial stimuli, very little is known about similar processes underlying STM for complex, naturalistic stimuli. Moreover, although it is known that some neurons in the lateral PFC exhibit audiovisual integration during passive exposure (Sugihara et al., 2006), virtually nothing is known about how the lateral PFC might be involved in retaining representations of audiovisual stimuli during STM.

Finally, in addition to the numerous questions surrounding the neural substrates of STM, there are still many behavioral aspects of STM which are understood in only very little detail. Due in part to the training challenges associated with studying auditory STM in nonhuman primates, many of the basic phenomena reported in the visual STM literature remain to be explored in the auditory modality. For example, experiments characterizing PI in auditory DMS tasks and comparing PI in visual and auditory STM have not been conducted. In addition, although the deficit in auditory STM has been well established in nonhuman primates, whether human auditory STM might be similarly deficient has not been resolved. Addressing these behavioral questions will provide a more complete foundation upon which studies investigating neural substrates of STM can build.

Chapter 2: Inferior auditory short-term and recognition memory in humans

2.1 Introduction

It is well established that monkeys' auditory memory capabilities fall substantially short of their visual and tactile memory capabilities (Colombo and D'Amato, 1986; Cohen et al., 2005; Kojima, 1985; Munoz-Lopez et al., 2010; Wegener, 1964). Many studies have reported that monkeys require extensive training to learn auditory memory tasks (Colombo and D'Amato, 1986; Fritz et al., 2005; Scott et al., 2012). Indeed, some of the earliest attempts to train monkeys on auditory memory tasks reported that subjects learned only "after years of failure", and others failed to learn at all (Colombo and D'Amato, 1986; D'Amato and Colombo, 1985). Moreover, upon learning the task, they appear capable of retaining auditory information for only a brief period of time. Thus, several experiments have reported that monkeys' accuracy falls below 75% correct at retention intervals of 40 seconds or less (Colombo and D'Amato, 1986; Fritz et al., 2005; Kojima, 1985). In contrast, monkeys learn visual and tactile memory tasks relatively quickly and are capable of approximately 75% accuracy at retention intervals of 10 minutes or greater (Buffalo et al., 1999; Murray and Mishkin, 1998; Overman and Doty, 1980). Inferior memory performance in auditory tasks has been observed in both Old World (Fritz et al., 2005; Kojima, 1985; Scott et al., 2012) and New World monkeys (Colombo and D'Amato, 1986; D'Amato and Colombo, 1985), as well as in a chimpanzee (Hashiya and Kojima, 2001), raising the possibility that auditory memory may be deficient in nonhuman primates in general.

Neuropsychological studies in monkeys suggest that the auditory retention deficit may

result, at least in part, from a difference in the degree to which auditory memory is enabled by the perirhinal and entorhinal cortices (Fritz et al., 2005; Munoz-Lopez et al., 2010). Although the perirhinal cortex receives substantial projections from visual and tactile cortex, auditory projections are very sparse (Brown and Aggleton, 2001; Mohedano-Moriano et al., 2007; Munoz-Lopez et al., 2010). Consistent with this anatomical distinction, combined lesions of the rhinal cortices severely disrupt visual and tactile memory (Buffalo et al., 1999; Meunier et al., 1993; Murray and Mishkin, 1998), but do not significantly impair auditory memory (Fritz et al., 2005). Moreover, as reported by Fritz et al. (2005), visual memory performance of monkeys with combined rhinal lesions is comparable to auditory memory performance of intact monkeys. Thus, auditory memory may not be substantially supported by the rhinal cortices. Additional factors that have been suggested to contribute to the discrepancy between auditory and visual/tactile memory in monkeys include relatively small proportions of auditory-responsive cells in lateral PFC (Poremba and Bigelow, 2013), limited involvement of the PFC in the development of sound-initiated motor behavior (Gemba and Sasaki, 1988), limitations in auditory sensation and perception (Colombo and D'Amato, 1986; Wegener, 1964), the temporally dynamic and transient nature of auditory stimuli (Kojima, 1985), and the ethological likelihood of aversive events being signaled in the auditory modality (Kojima, 1985).

Although it is clear that auditory memory differs from visual and tactile memory in nonhuman primates, a similar pattern of results has not been clearly established in humans (Fritz et al., 2005; Munoz-Lopez et al., 2010; Scott et al., 2012). Many studies conducted over the past century have investigated differences in auditory and visual memory, and some results indicate that humans may be relatively limited in retaining auditory information. For instance, Münsterberg (1894) reported that subjects were able to recall the serial order of digits and colors

with greater accuracy when they were presented visually compared to when they were spoken by the experimenter, also noting that even greater accuracy resulted from combined audiovisual presentation. Similarly, Kirkpatrick (1894) found that subjects' recall for lists of objects was substantially better when they viewed the physical objects themselves compared to when they heard the names of the objects pronounced by the experimenter. This outcome was consistent when subjects' recall was tested immediately, as well as after a 3-day delay.

Most of the subsequent experiments investigating modality differences have largely concentrated on recall for lists of verbal information such as digits or letters presented in the auditory or visual modalities (see reviews by Greene, 1992; Penney, 1975, 1989). Superior accuracy for the visual presentation modality has been observed only when a retention interval follows the list presentation (Jensen, 1971). On the other hand, if subjects are allowed to recall the items from the list immediately after the final item is presented, accuracy is typically higher for the auditory modality (Corballis, 1966), primarily due to superior recall of the final items presented in the list (i.e., a greater recency effect).

Beyond recall for lists of verbal cues, Cohen et al. (2009) have recently tested subjects' ability to recognize complex, naturalistic sound clips or images that had been previously presented during a study phase. Recognition accuracy was substantially lower for sound clips than for visual objects, even when additional cues, such as descriptions of the sounds were provided. A subsequent study by Cohen et al. (2011) similarly reported inferior auditory recognition memory even in subjects with considerable auditory expertise (professional musicians).

In summary, several experiments using delayed recall and recognition memory

paradigms have suggested that humans may have difficulty retaining auditory compared to visual stimuli. However, it is not clear from these studies whether this difference reflects a deficit in auditory compared to both visual and tactile memory (as in nonhuman primates), or whether there might be an advantage for retaining visual over auditory and tactile stimuli. It is also possible that memory might differ for each of these modalities. One study by Larsson and Bäckman (1998) provides some evidence that auditory retention may be inferior to both visual and tactile retention. In their study, subjects were briefly exposed to 40 common objects, which were presented in either the auditory, visual, tactile, or olfactory modality. Subjects were then instructed to identify the objects from a list of correct names mixed with distractors. The results indicated that auditory recognition was significantly lower than both tactile and visual recognition, which did not differ from each other. Olfactory recognition was intermediate between auditory and visual/tactile recognition. However, the results of this study were seriously compromised by the fact that the names of the objects were pronounced by the experimenter during the visual, tactile, and olfactory phases (i.e., bimodal presentation), whereas only the name of the object was given during the auditory phase (unimodal presentation). Moreover, subjects were given 6 s to study the objects during the visual, tactile, and olfactory phases, whereas pronouncing the name of the object during the auditory phase was likely accomplished in a shorter amount of time. Thus, it is likely that the bimodal presentation format and longer stimulus exposure time provided as significant advantage for visual, tactile, and olfactory phases compared to the auditory phase.

In addition to these ambiguities, several recent experiments have questioned whether differences reported in human auditory and visual memory tasks reflect inherent mnemonic differences between these sensory modalities (Ward et al., 2005; Visscher et al., 2007). Instead,

they have suggested that significant differences in memory functions may result from nonequivalent stimuli or task requirements. For example, Visscher et al. (2007) examined auditory and visual STM using artificial, nonverbal stimuli that had been equated in terms of discriminability, stimulus exposure time, and temporal dynamics. Under these conditions, the decrease in accuracy associated with larger memory sets and longer retention intervals was approximately equal for auditory and visual stimuli. Thus, prior experiments reporting differences in auditory and visual memory might have been biased by differences in discriminability among the stimuli, or perhaps by the verbal nature of the auditory stimuli. It is worth noting, however, that some results reported by Visscher et al. (2007) suggested a trend toward a greater recency advantage for auditory stimuli. Because the maximum retention interval used in this study was less than 10 s, it is possible that this trend could become more substantial under more taxing retention demands.

The current experiments were designed to address two primary questions. First, if comparable stimuli are used, are there significant differences in auditory and visual retention capabilities that might emerge at relatively long delays? Second, how might these results compare to tactile memory? Specifically, is there a deficit in auditory memory similar to that reported in nonhuman primate studies? Two experiments tested human subjects' memory for auditory, visual, and tactile stimuli using short-term and recognition memory paradigms. In general, we find support for the hypothesis that auditory memory is inferior to visual and tactile memory.

2.2 Experiment 1: Short-term memory

2.2.1 Experiment 1: Methods

Subjects

A total of 54 undergraduate students (37 female) with normal or corrected-to-normal vision and hearing participated in this experiment for course credit. Subjects gave their consent to participate in the study, and all procedures were approved by the Human Subjects Office at the University of Iowa.

Stimuli

The memoranda were simple, nonverbal stimuli that were matched in terms of stimulus exposure time (1 s), temporal dynamics (the stimuli did not vary over time), and discriminability at short retention intervals (described below). Auditory stimuli consisted of pure tones presented binaurally through headphones (Sennheiser HD-280), visual stimuli consisted of red squares (14 cm) presented on an LCD monitor positioned approximately 20 cm in front of the subject at eye level (~38° viewing angle), and tactile stimuli consisted of vibrations presented through a vertical aluminum bar which the subjects gripped with their left hand. The vibrotactile stimuli were produced by passing a digitally generated sine wave through a tactile transducer (TST209, Clark Synthesis, Inc., Highlands Ranch, CO). The vibrations were generated at a very low intensity to ensure that they were not audible to the subjects (acceleration values measured from the surface of the bar: 0.8 ± 0.1; VM-6360 digital vibration meter, Landtek Instruments, Guangzhou, China). Inaudibility was confirmed with a sound level meter (407740, Extech

Instruments Corporation, Nashua, NH), which did not detect change in sound pressure level produced by the vibration stimuli above the ambient noise in the room (35-36 dB).

Short-term memory task

Subjects' STM was tested using the *same*/*different* variation of the delayed matching-to-sample (DMS) task, which is frequently used in testing memory in nonhuman primates (Medin et al., 1976; Wright, 2006). Each trial began with a sample stimulus, which was followed by a variable retention interval of 1, 2, 4, 8, 16, or 32 s, after which a test stimulus that was either identical (same or match trials) or nonidentical (different or nonmatch trials) to the sample. An equal number of match and nonmatch trials using each of the six retention intervals were presented in random order. Upon termination of the test stimulus, the words "Same or different?" appeared on the screen. For match trials, subjects were instructed to click the left button of a mouse held with the right hand, whereas for nonmatch trials they were instructed to click the right button. Following each response, feedback was given by displaying the words "Correct" or "Incorrect" on the monitor for 250 ms, or "No response" if a response did not occur within 1.5 s. "No response" trials were discarded from further analysis (2.0% of total trials). Following feedback, the next trial began after a 1-s intertrial interval (ITI). The experiment was divided into three blocks, each consisting of 12 trials for each retention interval (total = 72 trials per block). Each block was identical except that the modality of the memoranda was either auditory, visual, or tactile. The order in which the sensory modality blocks occurred was fully counterbalanced across subjects, such that nine subjects were randomly assigned to participate in each of the six possible block sequences. All task events were controlled and recorded using E-prime 2.0 (Psychological Software Tools, Inc., Pittsburgh, PA).

Pilot experiments were used to identify two stimulus values for each sensory modality that yielded approximately 90% discrimination accuracy when the stimuli were separated by 1 s. The resulting values were tone frequencies of 1000 and 1018 Hz, red squares with RGB values of 224/0/0 and 255/0/0, and vibration frequencies of 60 and 205 Hz. Within each block of the experiment, the two stimulus values appeared as the sample and test stimuli on an equal number of trials in random order.

2.2.2 Experiment 1: Results

As seen in Figure 9, accuracy was very similar for each stimulus modality at the 1-s retention interval (auditory: 90.3%; visual: 91.5%; tactile: 89.7%). However, accuracy declined at longer retention intervals to a greater degree for auditory stimuli, such that accuracy at the 32-s retention interval was 61.8%, whereas for visual and tactile stimuli it was 78.3% and 78.8%, respectively. These differences were confirmed by repeated measures analysis of variance (ANOVA) with sensory modality (auditory, visual, tactile) and retention interval (1, 2, 4, 8, 16, 32 s) as factors, which revealed main effects of both retention interval ($F[5,265] = 57.88$, $p < .05$) and sensory modality ($F[2,106] = 11.61$, $p < .05$), as well as a significant interaction of these factors ($F[10,530] = 7.78$, $p < .05$). Of particular significance, post hoc tests ($p < .05$; Bonferroni correction for multiple comparisons) revealed that accuracy did not differ among sensory modalities at the 1-4s retention intervals, suggesting that lower accuracy observed at the longer retention intervals in the auditory block was not attributable to differences in stimulus discriminability.

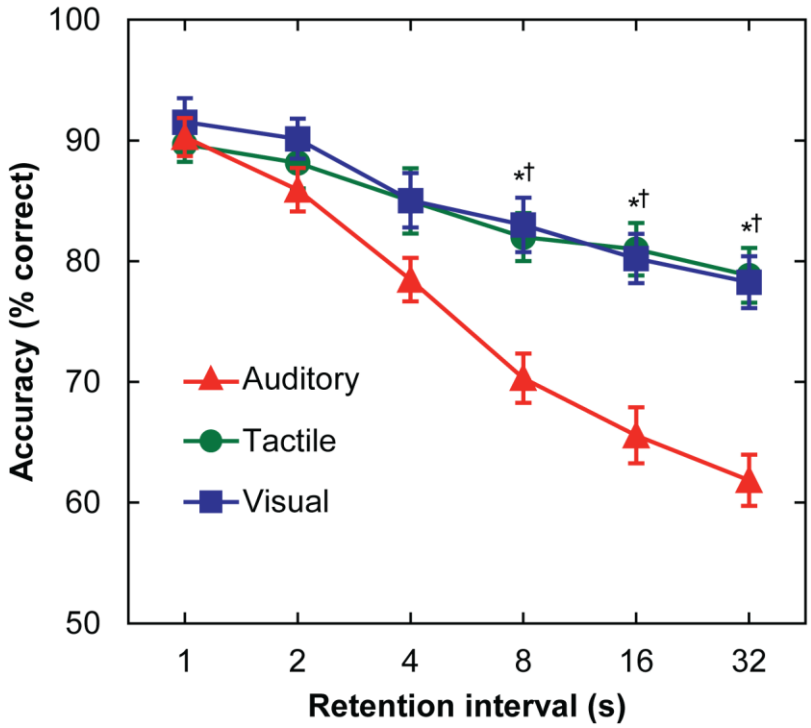

Figure 9. Experiment 1: Mean (± SEM) short-term memory accuracy among sensory modalities for simple, artificial stimuli (see Methods). Short-term retention of auditory stimuli declines at a greater rate than retention of visual or tactile stimuli. There were no differences in accuracy among the sensory modalities for trials with brief retention intervals (1–4 s), indicating that the initial discriminability of the stimuli was approximately equal. However, at longer retention intervals (8–32 s), accuracy for auditory trials was significantly lower than visual and tactile trials. Post hoc tests ($p < .05$, Bonferroni correction for multiple comparisons): *Accuracy in the auditory block significantly lower than the tactile block. †Accuracy in the auditory block significantly lower than the visual block.

Two additional analyses were conducted to address the possibility that these results might be attributable to factors other than a deficit in auditory retention capability. First, we investigated whether our results might have been biased by differential practice effects within different sensory modality blocks, similar to those observed in some previous experiments (e.g., Visscher et al., 2007). In other words, it is conceivable that lower mean accuracy in the auditory block could have resulted if the subjects took longer to become familiar with the auditory stimuli than the visual or tactile stimuli. To test this possibility, each modality block of the experiment

was subdivided into six successive sub-blocks of 12 trials. Repeated measures ANOVA with modality and trial sub-block as factors reconfirmed the significant effect of sensory modality block (F[2,106] = 11.07, $p < .05$), and indicated that there were significant practice effects (F[5,265] = 12.05, $p < .05$). Post hoc comparisons indicated that subjects improved during the first two sub-blocks of 12 trials, reaching asymptotic performance by the third sub-block. However, there was no significant interaction of sensory modality and trial sub-block (F[10,530] = 0.56, $p > .05$), disconfirming the likelihood that the lower mean accuracy observed in the auditory block resulted from slower familiarization with the stimuli.

The second additional analysis was concerned with the potential influence of proactive interference (PI), which may occur if a minimal number of stimuli are recycled as memoranda from trial to trial. Specifically, studies of both human and animal memory show that subjects are more likely to commit an incorrect "match" response on a nonmatch trial if the test stimulus had been presented on the previous trial (Chapter 3; Bigelow and Poremba, 2013a, 2013b; Hartshorne, 2008; Wright, 2006). In our study, the lower mean accuracy in the auditory block might have been partially influenced by increased susceptibility to PI for auditory stimuli. This possibility was addressed by comparing accuracy on nonmatch trials for which the test stimulus had occurred (PI) or had not occurred (no PI) as the sample stimulus on the previous trial. Repeated measures ANOVA with modality and PI (PI, no PI) as factors again revealed the significant effect of sensory modality (F[2,106] = 3.72, $p < .05$). Contrary to our expectations, however, there was neither a significant effect of PI (F[1,53] = 0.01, $p > .05$), nor a significant interaction of PI and modality block (F[2,106] = 1.15, $p > .05$). In light of these results, it can be safely concluded that PI did not contribute to the observed performance deficit in auditory trials.

In summary, Experiment 1 revealed that retention was limited for auditory stimuli

compared to visual or tactile stimuli, even though these stimuli did not differ in terms of discriminability at very short retention intervals. Further analyses revealed that these results were not influenced by differential practice effects or susceptibility to PI among sensory modalities. These results support the hypothesis that, as in nonhuman primates, auditory retention capabilities in humans may be relatively limited.

2.3 Experiment 2: Recognition memory

2.3.1 Experiment 2: Methods

Subjects

A total of 82 undergraduate students (42 female) with normal or corrected-to-normal vision and hearing participated in this experiment for course credit. Subjects gave their consent to participate in the study, and all procedures were approved by the Human Subjects Office at the University of Iowa.

Stimuli

Using simple, artificial stimuli with carefully controlled stimulus properties in Experiment 1, we observed relatively poor retention of acoustic information compared to visual or tactile information. The primary goal of Experiment 2 was to investigate the real-world applicability of this finding, i.e., whether this pattern of results generalizes to complex, naturalistic stimuli likely to be encountered in everyday life. Thus, the auditory stimuli used in

this experiment were sound recordings of easily recognizable, everyday events (e.g., dog barking), presented binaurally through headphones (Sennheiser HD-280). Similarly, visual stimuli comprised silent videos of scenes and events (e.g., scuba diver; dimensions: 6" × 3.5", or 15.24 cm × 8.89 cm) presented on an LCD monitor positioned approximately 20 cm in front of the subject at eye level (~42° viewing angle). For tactile stimuli, common physical objects (e.g., coffee mug) were presented to subjects, which they were allowed to touch and manipulate but not see or hear. During the tactile block, a research assistant sat facing the subject on the opposite side of the desk. The tactile objects were stored on a bookshelf next to the desk, facing away from the subject so that they were not visible. For each trial, the research assistant placed one object inside of an opaque box (48 cm × 55 cm × 33 cm) that was sitting on the desk through an opening in the back of the box (20 cm × 48 cm) that was not visible to the subject. In order to reach the object, the subjects put their arms through two small openings (13 cm × 13 cm) in the front of the box. Heavy tassels hung from the inside of the arm openings to prevent the subjects from seeing the object in the box. Several steps were taken to minimize the possibility that the tactile objects could produce perceptible auditory cues. First, tactile stimuli were initially selected for the experiment on the condition that they did not produce salient or characteristic auditory cues that might reveal the object independent from its physical structure. Second, the box in which the objects were placed was lined with foam to minimize percussive sounds that could be produced when the object was placed inside the box. Finally, the headphones worn by the subjects during the tactile block provided 32 dB of external sound attenuation.

In contrast to the artificial stimuli used in Experiment 1, which can be easily manipulated along a relevant dimension, naturalistic stimuli are much more difficult to control in terms of

discriminability and other stimulus attributes. Nevertheless, several measures were taken to ensure that the stimulus sets for each modality were as comparable as possible. First, the stimuli chosen for each sensory modality were temporally dynamic. Thus, videos were chosen as visual stimuli instead of images, because like the naturalistic sound recordings, the stimulus information unfolds over time. Similarly, different parts of the hand and fingers are stimulated over time as subjects touch and manipulate the three dimensional physical objects, and only partial stimulus information is available to the sensory receptors at a given time.

Second, stimulus exposure time was roughly equated for each modality block. The sound recordings and video clips were each trimmed to 5 s in duration. To ensure that the tactile stimulus exposure time was approximately equal to that of the auditory and visual blocks, cues were presented on the LCD monitor instructing the subjects when to begin and cease touching the objects. During the ITI, a gray screen displayed the words "Put hands in box, but don't touch object yet" above a countdown starting 5 s before the stimulus presentation period. The screen then turned red and displayed the words "Touch object" above a 5-s countdown indicating the duration of the stimulus presentation period. In addition, subjects wore headphones (Sennheiser HD-280) through which a tone (880 Hz, 500 ms) was presented to signal the beginning of the stimulus presentation period. At the end of the stimulus presentation period, the screen returned to gray for the subsequent ITI countdown or response window depending on whether the stimuli were presented during the study phase or recognition phase (see below).

Finally, before conducting the recognition experiment, 10 subjects (6 female) with native English fluency participated in an object identification task. This was used as a rough index of the discriminability or recognizability of the stimuli for each sensory modality (for a similar approach, see Cohen et al., 2009). Each subject was exposed to 100 stimuli for each sensory

modality. A single stimulus was presented on each trial, after which subjects were instructed to identify the name of the stimulus from a list of ten options that remained on the screen until a choice was made (chance accuracy = 10%). The nine incorrect object names were randomly selected from the remaining 99 stimuli within the same sensory modality. Following each response, feedback was given by displaying the words "Correct" or "Incorrect" on the monitor along with cumulative accuracy for the session. The feedback display terminated when the subject pressed either of two foot pedals located beneath the desk. Following a 5-s ITI, the next stimulus was presented. Each subject achieved greater than 97% object identification accuracy. For the recognition task, 90 stimuli were selected for each sensory modality block that were correctly identified by all ten subjects (i.e., 100% accuracy).

Recognition memory task

The recognition task consisted of a study phase followed by a recognition phase. The study and recognition phases each had separate auditory, visual, and tactile blocks. For each block during the study phase, subjects were exposed to 60 stimuli and instructed that their recognition of these items would be tested during the subsequent recognition phase. After each stimulus was presented, subjects were instructed to press either foot pedal to advance to the next stimulus. Stimulus presentations for all blocks were separated by a 5-s ITI to ensure equal temporal spacing of the study items. The recognition phase was similar to the study phase except that 30 of the stimuli for each block were repeated from the study phase (*old* trials) and 30 were presented for the first time (*new* trials). An equal number of *old* and *new* trials were presented in random order, and the stimuli selected for the study and recognition phases were randomized across subjects. Upon termination of each stimulus, the words "Old or new?" appeared on the screen, and subjects were instructed to press the left foot pedal for *new* stimuli and the right foot

pedal for *old* stimuli. Following each response, feedback was given by displaying the words "Correct" or "Incorrect" on the screen until a press of either foot pedal initiated the next trial. All task events were controlled and recorded using E-prime 2.0 (Psychological Software Tools, Inc., Pittsburgh, PA).

The time between the study and recognition phases differed for three groups. The *same-day recognition* group (n = 24, 11 female) began the recognition phase immediately after the study phase (the study phase lasted approximately 45-60 minutes depending on how quickly the subjects responded and advanced through the directions). The *next-day recognition* group (n = 24, 10 female) and *next-week recognition* group (n = 24, 15 female) began the recognition phase 24 hours and 7 days after the study phase, respectively. The order in which the sensory modality blocks occurred was fully counterbalanced across subjects, such that four subjects per group were randomly assigned to participate in each of the six possible block sequences.

2.3.2 Experiment 2: Results

For the *same-day recognition* group, repeated measures ANOVA revealed a significant effect of modality block ($F[2,46] = 29.69$, $p < .05$). Consistent with the pattern of results observed in Experiment 1, post hoc analyses ($p < .05$; Bonferroni correction for multiple comparisons) indicated that mean accuracy for the auditory block (88.61%) was significantly lower than both the visual (96.74%) and tactile (97.99%) blocks, which did not significantly differ from each other (Figure 10A). Thus both STM for simple, artificial stimuli and recognition memory for complex, naturalistic stimuli appear to be inferior in the auditory modality.

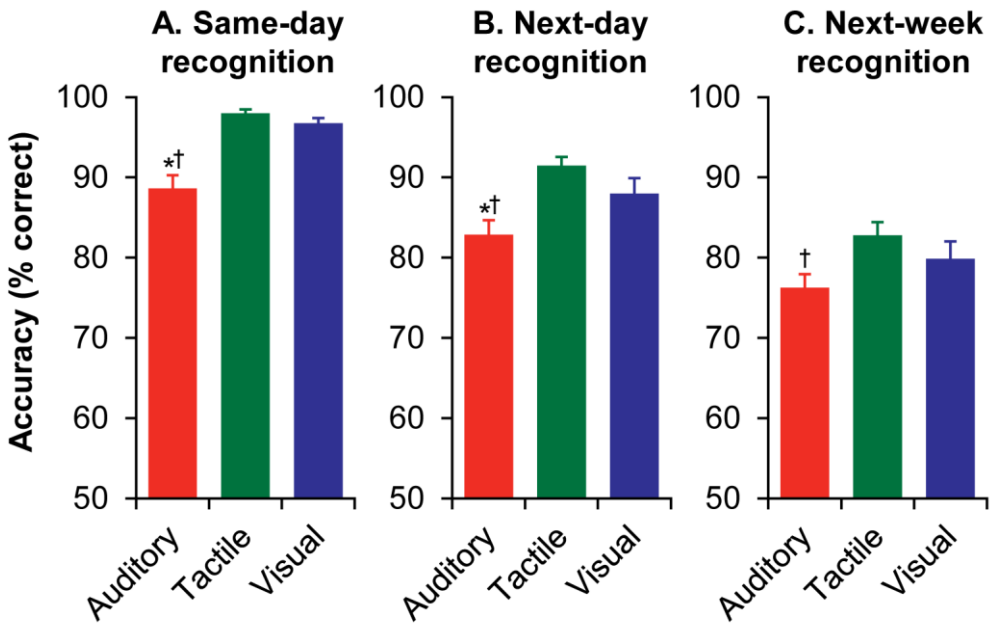

Figure 10. Experiment 2: Mean (+ SEM) recognition accuracy among sensory modalities for complex, naturalistic stimuli (see Methods). (A) When tested immediately after the study phase, recognition accuracy was lower for auditory stimuli than visual or tactile stimuli. (B) Similarly, recognition was lower for auditory stimuli when tested 24 hours after the study phase. (C) When tested one week after the study phase, recognition accuracy was significantly lower for auditory stimuli than tactile stimuli, but the difference between auditory and visual recognition was not significant. Post hoc tests ($p < .05$; Bonferroni correction): *Accuracy in the auditory block significantly lower than the tactile block. †Accuracy in the auditory block significantly lower than the visual block.

For the *same-day recognition* group, accuracy for both the visual and tactile recognition blocks was near ceiling, which might have concealed differences in recognition memory between these two modalities. For this reason, the *next-day recognition* and *next-week recognition* groups were added to the experiment so that visual and tactile recognition could be compared under conditions in which accuracy was unlikely to reach ceiling. As expected, mean overall accuracy declined at each successively longer delay (One way ANOVA: $F[2,69] = 38.61$, $p < .05$; all pairwise comparisons significant). Repeated measures ANOVAs again revealed significant effects of modality block for both the *next-day recognition* group ($F[2,46] = 9.51$, $p < .05$) and

the *next-week recognition* group (F[2,46] = 5.38, $p < .05$). For the *next-day recognition* group, post hoc comparisons indicated that mean accuracy during the auditory block (82.85%) was again significantly lower than both the visual (87.99%) and tactile (91.46%) blocks, which did not significantly differ from each other (Figure 10B). For the *next-week recognition* group, mean accuracy during the auditory block (76.25%) was significantly lower than the tactile (82.78%) block (Figure 10C). However, although accuracy was lower in the auditory block than in the visual block (79.86%), this difference was not significant. Again, the difference between visual and tactile recognition accuracy was not significant.

Although accuracy predictably decreased with increasing time between the study and recognition phases, as indicated by mean accuracy scores, the magnitude of the deficit in auditory recognition compared to visual and tactile recognition diminished at the longer delays. This outcome contradicted our *a priori* expectation that, since auditory recognition accuracy was relatively poor after a short delay period, this difference would become more pronounced with time. It is also unexpected in light of the sharper decline in accuracy with increasing retention intervals observed during auditory blocks in Experiment 1. Although this trend is somewhat paradoxical, a mixed factors ANOVA with modality as a within subjects factor and delay (same day, next day, next week) as a between subjects factor indicated that the interaction of these variables was not significant (F[4,138] = 0.91, $p > .05$). Nevertheless, future studies should be conducted to determine whether a significant trend might emerge using longer delays (and perhaps a fully within subjects design).

In summary, recognition accuracy was lowest for the auditory stimuli in the *same-day recognition*, *next-day recognition*, and *next-week recognition* groups. These differences were statistically significant in nearly all cases, with the exception that auditory accuracy did not differ

from visual accuracy in the *next-week recognition* group. Visual and tactile accuracy, on the other hand, did not differ significantly for any of the groups. Together with the results of Experiment 1, these outcomes suggest that, like nonhuman primates, humans are relatively limited in retaining acoustic information.

2.4 Discussion

In general, we observed that retention was inferior for acoustic stimuli compared to visual and tactile stimuli, whereas retention for visual and tactile stimuli was approximately equal. Similar outcomes were observed in tests of STM for simple, artificial stimuli as well as recognition memory for complex, naturalistic stimuli. The deficit in auditory retention was not attributable to differences in the discriminability, exposure time, or temporal dynamics of the stimuli. Further, the results were neither biased by differential practice effects nor increased susceptibility to PI in the auditory modality.

The findings that human STM and recognition memory are inferior for auditory stimuli have several significant implications. In the first place, our results are qualitatively similar to the pattern of results that has been established in the nonhuman primate literature over the past several decades (Figure 11). The findings thus add to the homologies observed between humans and nonhuman primates in numerous other aspects of cognition (Matsuzawa, 2001; Wasserman and Zentall, 2006), and importantly, lend increased validity to primate models of human cognitive deficits including amnesia. In addition to these comparative questions, our data strengthen the evidence that memory capabilities are at least in part modality dependent, and thus

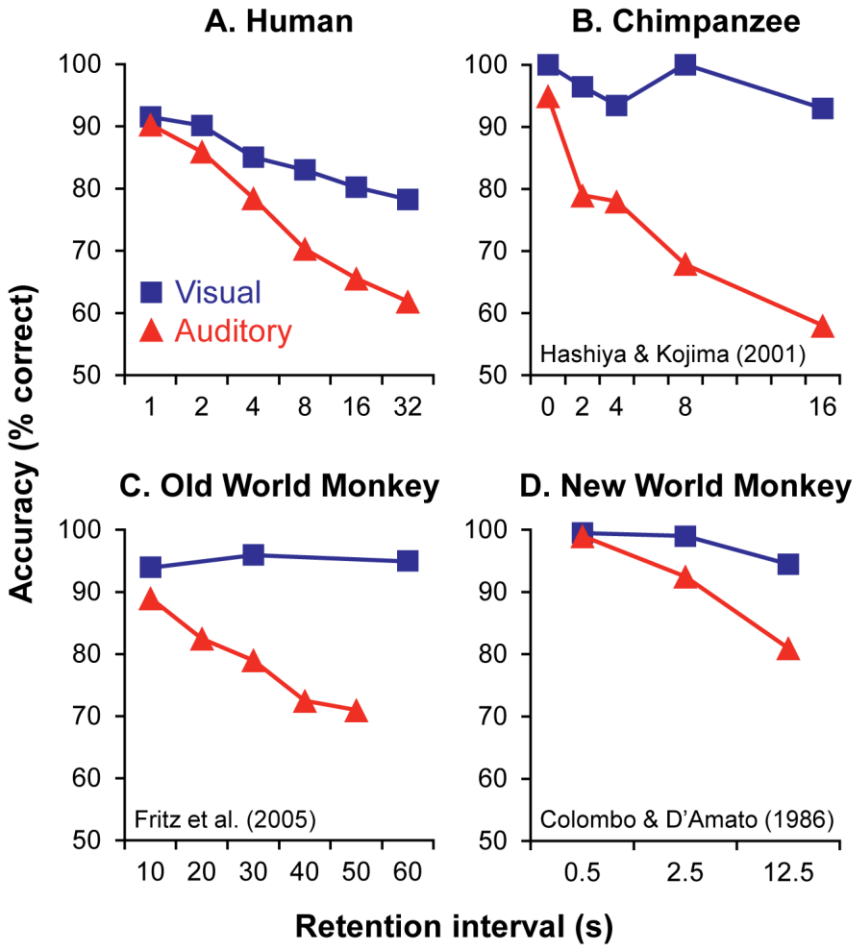

Figure 11. Comparison of visual and auditory STM among primates. In the present experiment (A), inferior retention was observed for auditory compared to visual stimuli in human subjects. This pattern of results is qualitatively similar to that which has been observed in the chimpanzee (B), as well as both old-world (C) and new world monkeys (D). (B) adapted from Hashiya and Kojima (2001); C adapted from Fritz et al. (2005); (D) adapted from Colombo and D'Amato (1986).

provide support for theories of memory that account for differences in sensory processing pathways (Pasternak and Greenlee, 2005).

In nonhuman primates, neuropsychological experiments have suggested that the perirhinal and entorhinal cortices are less involved in auditory memory than visual and tactile memory (Buffalo et al., 1999; Munoz-Lopez et al., 2010; Murray and Mishkin, 1998; Fritz et al.,

2005). Very few studies have addressed whether a similar dissociation might exist in humans. Patients with extensive lesions of the medial temporal lobe, including noted patient H. M., exhibit deficits in both visual and auditory recognition memory (Milner, 1972; Squire et al., 2001). Yet in each of these cases, lesions encompassed not only the perirhinal and entorhinal cortices, but also at least parts of the hippocampus and parahippocampal gyrus. In contrast to the rhinal cortices, the parahippocampal gyrus in nonhuman primates receives significant input from auditory cortices in the superior temporal gyrus (Suzuki and Amaral, 1994). Thus the deficit in auditory recognition may have been caused primarily by damage to the parahippocampal cortex. This suggestion is supported by a human neuroimaging study of auditory and visual recognition memory by Peters et al. (2007). During an encoding session, subjects saw images of common objects presented against a background of either 'lawn' or 'clouds', and heard names of common objects spoken by either a male or female voice. In the recognition session, visual stimuli were presented on a neutral background and auditory stimuli were spoken by a gender-neutral 'robot voice'. Subjects were instructed to indicate whether each stimulus was old or new, and for the old items, to report the context in which the item had initially been presented (lawn or cloud background, male or female voice). For auditory but not visual trials, activity in the left and right parahippocampal cortices discriminated between correct and incorrect judgments of the context in which the stimuli had been encoded. On the other hand, overall activation of the right perirhinal cortex was greater during visual encoding, and activity in the left perirhinal cortex discriminated between correct and incorrect context judgments for visual but not auditory trials. The latter observations correspond roughly to the engagement of the nonhuman primate perirhinal cortex in visual but not auditory recognition memory.

It is possible then, that the deficits in auditory retention observed in our experiments as

well as in previous studies (Cohen et al., 2009, 2011) may reflect a difference in the degree to which memory is supported by the rhinal cortices. If this were true, it would contribute to a growing body of literature suggesting a specialized role for the rhinal cortices in familiarity-based recognition (Brown and Aggleton, 2001; Eichenbaum et al., 2007; Peters et al., 2007). Indeed, in many of the human and nonhuman primate studies that have reported relatively poor auditory performance (including our own), the tasks are such that successful performance could be accomplished by relying on a familiarity-based recognition strategy. However, additional experiments are needed before this view can be fully validated. For example, human neuroimaging studies using additional stimulus modalities could reveal whether activation of the rhinal cortices is greater during tactile and perhaps olfactory memory compared to auditory memory. In ideal circumstances, studies of patients with lesions restricted to the rhinal cortices could be used to determine whether recognition memory deficits were observed for auditory stimuli. Nonhuman primate studies may also be useful for determining whether parahippocampal lesions might disrupt memory for auditory stimuli, as the studies in humans suggest (Peters et al., 2007).

Although our findings are consistent with a number of previous human and nonhuman primate studies showing limited retention of auditory information (Cohen et al., 2009, 2011; Colombo and D'Amato, 1986; Fritz et al., 2005; Hashiya and Kojima, 2001; Jensen, 1971; Kojima, 1985; Scott et al., 2012), these results do not necessarily imply that memory is inferior in the auditory modality for every taxonomic class of memory. On the contrary, many studies have demonstrated that immediate recall for lists of verbal materials is superior when presented in the auditory modality (Corballis, 1966; Greene, 1992; Jensen, 1971; Penney, 1975, 1989). Further, lesions that impair familiarity-based forms of recognition memory do not affect other

forms of memory such as priming (Squire and Zola-Morgan, 1991). Thus, future comparisons of memory across sensory modalities should be mindful of specific memory processes likely to be engaged by a given task.

In conclusion, our results suggest that primates may have inferior retention capabilities for auditory events. Further, they imply that memory is to some extent modality dependent, which is likely a consequence of differences among neural pathways in which memoranda are processed. These views are not new; indeed, they have been held by memory researchers for over a century (e.g., Münsterberg, 1894; Kirkpatrick, 1894), and can be found in folk wisdom dating much earlier. For example, a common English translation of an old Chinese proverb states "I hear, and I forget... I see, and I remember." In light of the current experimental data, this adage might be amended to include "touch" as an additional mode of superior memory.

Chapter 3: A behavioral investigation of auditory proactive interference in nonhuman primates

3.1 Introduction

Proactive interference (PI) occurs when memory processing at a given point in time disrupts memory processing at a future time (Baddeley, 1990; Wright, Urcuioli, & Sands, 1986). For example, processing a particular stimulus on a given trial of a memory task can interfere with processing on a subsequent trial that uses the same stimulus. Early studies in human verbal memory led investigators to conclude that PI is a predominant cause of mnemonic failure in laboratory experiments as well as everyday memory usage (Keppel & Underwood, 1962; Underwood, 1957).

PI appears to be pervasive in human memory, having been observed in a variety of visual (Badre & Wagner, 2005; Hartshorne, 2008; Makovski & Jiang, 2008; Mecklinger et al., 2003), motor (Burwitz, 1974; Cothros et al., 2006; Herman & Bailey, 1970), verbal (Feredoes, Tononi, & Postle, 2006; Kane & Engle, 2000; Keppel & Underwood, 1962; Postle et al., 2001), and auditory memory tasks (Ruusuvirta, 2000; Ruusuvirta, Astikainen, & Wikgren, 2002; Ruusuvirta, Wikgren, & Astikainen, 2008; Visscher, Kahana, & Sekuler, 2009). Similarly, the influence of PI has been widely reported in studies of memory processing in animals, including pigeons (Grant, 1975; Hogan, Edwards, & Zentall, 1981; Wright, Katz, & Ma, 2012), rats (De Rosa & Hasselmo, 2000; Dunnett & Martel, 1990; Gleitman & Jung, 1963; Grant, 1981), monkeys (Mishkin & Delacour, 1975; Overman & Doty, 1980; Wright, 2006, 2007), chimpanzees (Hayes & Thompson, 1953), chickadees (Hampton, Shettleworth, & Westwood,

1998), and dolphins (Herman, 1975; Thompson & Herman, 1981). With the exception of studies in dolphins, nearly all studies of PI in animals have focused on visual STM (see further discussion below).

One of the traditional approaches to assessing PI in nonhuman primates has been varying the degree to which stimuli are reused from trial to trial in a memory task. The typical finding is that memory capabilities improve when new stimuli are used for each trial because confusion arising from stimulus repetitions between trials is reduced. For instance, Hayes and Thompson (1953) found that chimpanzees committed fewer errors on a delayed response task if new stimuli were used for each trial than if a single pair of stimuli were alternately used as the sample and comparison throughout the experiment. Similarly, Mishkin and Delacour (1975) observed that monkeys require relatively few sessions to learn visual memory tasks if trial-unique stimuli are used, whereas they require significantly more sessions or fail to learn if only two memoranda are repeatedly presented throughout a session. Using a similar visual task, Overman and Doty (1980) found that the maximum retention interval in monkeys increased from under 30 s when two stimuli were repeatedly reused to over 24 h when trial-unique stimuli were used. The positive relationship between performance and the number of stimuli used in a session (i.e., stimulus set size) has been consistently reported in several additional studies of visual memory in monkeys (Mason & Wilson, 1974; Medin, 1980; Sands & Wright, 1980; Worsham, 1975).

An additional, more direct method for examining the influence of PI is to evaluate performance on trial *n* as a function of the stimuli presented on previous trials (intertrial PI). A number of studies have demonstrated that performance on trial *n* can be significantly altered by memory processing on the immediately preceding trial, trial *n* – 1 (Edhouse & White, 1988; Grant, 1975; Hogan et al., 1981; Makovski & Jiang, 2008; Moise, 1976; Reynolds & Medin,

1981; Thompson & Herman, 1981; Worsham, 1975). These studies are consistent with decreased overall performance associated with smaller stimulus set sizes because the frequency with which the same stimuli are used for both trial *n* and trial *n* – 1 increases when smaller stimulus sets are used.

Relatively few studies have examined the effect of stimulus repetition beyond nonadjacent trials, i.e., when the same stimuli are used for trials *n* and *n* – 2 or trials *n* and *n* – *k*. One such experiment reported that visuospatial memory in rats was influenced by PI produced by trial *n* – 1, but not by more distant, nonadjacent trials (Dunnett & Martel, 1990). By contrast, Hartshorne (2008) reported that human visual memory was susceptible to PI caused by stimulus repetitions across at least four trials (see also Monsell, 1978). Similarly, Wright, Katz, and Ma (2012) recently reported that significant PI in pigeon visual memory was produced by stimulus repetitions separated by as many as 16 trials, which was the longest distance tested. Finally, recent data from our laboratory indicate that stimulus repetitions separated by up to 10 trials can produce significant PI in auditory memory in monkeys (Bigelow & Poremba, 2013a). These studies highlight the potentially perseverative nature of PI by showing that repeating a small number of stimuli throughout a session can negatively impact performance beyond the immediately subsequent trial.

In addition to reducing the number of intertrial stimulus repetitions, an additional means whereby PI can be reduced is increasing the time that elapses between each trial, or intertrial interval (ITI). Increases in overall accuracy resulting from increasing the ITI have been demonstrated in a variety of memory tasks in humans and nonhuman animals (Cermak, 1970; Cohen, Reid, & Chew, 1994; Herman, 1975; Maki, Moe, & Bierley, 1977; Mason & Wilson, 1974; Nelson & Wasserman, 1978; Roberts, 1980; Roberts & Kraemer, 1982). In some cases, the

benefits of increasing the ITI have been relatively modest. For instance, Jarrard and Moise (1971) reported that monkeys' visual STM accuracy improved by approximately 5 to 10% (depending on the retention interval) after increasing the ITI from 5 to 15 s, but no significant advantage was gained by further extending the ITI to 30 or 60 s. In other cases, increasing the ITI has led to more substantial benefits, to the extent that the influence of intertrial PI has been reduced to zero. In Dunnett and Martel's (1990) study of rat visuospatial memory, PI effects from trial $n - 1$ were eliminated by increasing the ITI from 5 s to 15 s. Similarly, two studies of pigeon visual memory reported significant intertrial PI effects when the ITI was 2 s, but not when it was 10 s or greater (Grant, 1975; Hogan et al., 1981). These results suggest that increasing the time between trials allows greater decay of irrelevant memory traces, which might otherwise compete with memory demands of the current trial.

With few exceptions, studies of PI in animals including monkeys have been concerned with the visual sensory modality. Because auditory perception and memory are crucial for key aspects of nonhuman primate ethology such as predator evasion and conspecific communication (e.g., Ghazanfar & Hauser, 2001), the relatively sparse auditory memory literature, including PI, constitutes a significant deficit in scientific understanding. One likely reason for the lack of experimental data in the auditory modality is that, unlike visual memory tasks, monkeys require extensive training to learn auditory memory tasks (Cohen, Russ, & Gifford, 2005; Colombo & D'Amato, 1986; D'Amato & Colombo, 1985; Fritz, Mishkin, & Saunders, 2005; Kojima, 1985; Scott, Mishkin, & Yin, 2012). For instance, Fritz et al. (2005) reported that monkeys required ~15,000 trials to learn an auditory memory task, whereas Mishkin and Delacour (1975) reported that only ~500 trials were needed to learn a comparable visual memory task. A related finding is that the maximum reported retention interval for auditory memory in monkeys ranges from 16 to

35 s (Colombo & D'Amato, 1986; Fritz et al., 2005; Kojima, 1985), whereas for visual memory it ranges from minutes to hours or more (Murray & Mishkin, 1998; Overman & Doty, 1980). Monkeys are also capable of retaining tactile information for at least several minutes (Bauer & Steele, 1985; Buffalo et al., 1999; Suzuki et al., 1993), suggesting that auditory memory tasks may be uniquely challenging.

Because an investigation of the basic parameters of auditory PI in monkeys is lacking, we conducted two experiments to address the effects of stimulus set size (Experiment 1) and the duration of the ITI (Experiment 2) in an auditory STM task in monkeys. The experiments were designed in such a way that the results would be roughly comparable to previous studies of memory in monkeys using the visual modality. As with visual PI, we hypothesized that the influence of auditory PI would diminish as a function of the stimulus set size and the duration of the ITI.

3.2 Experiment 1: The role of stimulus set size

3.2.1 Experiment 1: Methods

Subjects

Two adult male macaque monkeys (*Macaca mulatta*) served as subjects for this experiment (Monkeys F and S). The monkeys had been trained to perform the auditory STM task prior to the experiment. The animals were housed in individual cages with ad libitum access to water and controlled feeding schedules, under a 12:12-h light:dark cycle. Experimental sessions

were conducted 3-5 days per week. The majority of food was given after the experimental session each day (Harlan monkey diet plus fresh fruit, vegetables, and treats) and each animal was maintained above 85% of their weight during the use of a controlled feeding schedule. All procedures were carried out with approval from The Institutional Animal Care and Use Committee at the University of Iowa.

Apparatus

Subjects were placed in a sound attenuated chamber for the duration of each experimental session. The animal was held in a custom made primate chair that allowed free arm movement for behavioral responses. Acoustic stimuli were delivered through a single speaker located 15 cm in front of the primate chair at eye level. Behavioral responses were made via a single acrylic button positioned 3 cm below the speaker. Small food rewards were dispensed from a pellet dispenser (Med Associates, Georgia, VT) through a copper tube into a dish located 3 cm below the response button. A dimmed 40-watt house light provided illumination throughout the duration of the experiment, and a second light provided additional illumination during the ITI. Custom designed software (LabVIEW, National Instruments, Dallas, TX) controlled and recorded the stimulus presentations and other task events.

Procedure

Task

The STM task used for this experiment was a variation of the delayed matching-to-sample (DMS) task, which is suitable for use with auditory stimuli. The typical DMS task begins with the presentation of a sample stimulus, which is followed by a retention interval, after which subjects are rewarded for identifying the sample from two test stimuli. In the *same*/*different*

variation of the DMS task (Wright, 2006), a single test stimulus is presented following the retention interval and the subject must indicate whether it is identical to (match trials) or different from the sample (nonmatch trials). The traditional two choice and *same*/*different* versions of the DMS task produce very similar outcomes in visual STM performance in monkeys (D'Amato & Worsham, 1974). Thus visual memory experiments using the traditional DMS task can be reasonably compared to the current experiment as well as previous studies, which use the auditory *same*/*different* DMS task (Colombo & D'Amato, 1986; Kojima, 1985; Konorski, 1959; Wright, Shyan, & Jitsumori, 1990).

The task for Experiment 1 used a fixed retention interval, or interstimulus interval (ISI), of 5 s, and a variable ITI averaging 10 s (range: 8–12 s). Each session consisted of a total of 128 trials with an equal number of match and nonmatch trials presented in pseudorandom order. Following the presentation of the test stimulus, the response button was illuminated for 1 s to indicate that a response could be made. If a button press was made outside of the 1-s response window, the current trial was aborted and replaced with a new trial. For match trials, correct responses were defined by the presence of a button press ("go" response) following the test stimulus, whereas for nonmatch trials, correct responses were defined by the absence of a button press ("no-go" response). The task used an asymmetric reinforcement contingency in which correct "go" responses on match trials were rewarded with a small food pellet and incorrect button presses (false alarms) on nonmatch trials were occasionally punished by a brief, mild air puff presented indirectly from a distance of approximately 15 cm from the animal (approximately 1/10 of incorrect nonmatch trials were punished on a variable schedule). Similar asymmetric reinforcement contingencies have been used in previous studies of auditory STM in monkeys because they facilitate learning the match vs. nonmatch rule (Colombo & D'Amato,

1986; Kojima, 1985; Stepien & Cordeau, 1960).

Stimulus sets

Experiment 1 consisted of a systematic manipulation of the number of sounds that were recycled as the sample and test stimuli throughout each 128-trial session. For a given experimental session, one of the following stimulus set sizes was randomly selected: 2, 4, 8, 16, 32, 64, or trial unique. For the trial-unique condition, 192 stimuli were used (64 sounds used for 64 match trials, 128 sounds used for 64 nonmatch trials). After a stimulus set size was used for a session, it was not used again until the remaining six set sizes had been used. Each animal completed a total of 20 sessions with each stimulus set size.

The stimuli used for each session were randomly selected from a collection of 192 sounds consisting of 32 exemplars from each of the following six sound classes: conspecific monkey vocalizations, human vocalizations, animal vocalizations, natural and environmental sounds, music clips, and synthetic/abstract sounds. The monkey vocalizations were recorded at a natural monkey reserve in South Carolina, USA (by Amy Poremba), and included a variety of coos, grunts, screams, shrill barks, and harmonic arches. The human vocalizations consisted of a variety speech and nonspeech vocalizations from a variety of speakers, including members of each gender. Animal vocalizations came from a wide variety of birds and mammals other than monkeys and humans. Natural and environmental sounds included natural phenomena, such as thunder and breaking tree branches, as well as sounds that the animals might have been exposed to in the laboratory, such as a door closing or a broom falling on the floor. Music clips comprised multi-note sequences extracted from variety of sources such as solo instrument performances, popular music recordings, and TV commercials. Synthetic and abstract sounds were artificial

sounds generated by electronic synthesizers or downloaded from abstract sound categories (e.g., "science fiction") of a commercially available sound effects database (www.soundsnap.com). All sounds were trimmed to 500 ms, volume normalized, and presented at 75 ± 5 decibels. Within a session, each sound had an equal chance to be presented on a given trial as the sample and/or test stimulus, depending on whether it was a match or nonmatch trial, with the constraint that each sound was presented an equal number of times throughout the session.

We found no evidence that the effects of PI differed among the sound types used in our study, thus, the results presented below are collapsed across sounds. It should be noted, however, that our experiment was not specifically designed to test for differences in PI effects among sound types. For instance, by randomly selecting the sounds to be used for each session, there were an unequal number of sessions using a given sound type. Thus, the question of whether PI interacts with sound type remains to be addressed by future studies.

Analysis

Although PI literature has traditionally focused on accuracy (percent correct) as the dependent measure, a relatively small number of publications have reported modulation of response latency by PI (Hendrikx, 1986; Monsell, 1978; Wixted & Rohrer, 1993). For this reason, both accuracy and response latency were evaluated as dependent variables in our study.

The animals occasionally stopped responding before the final trial of the experiment session. For sessions in which the subject did not make a single response during the last 20 trials, we considered the final response as the end of the session. The remaining trials were rejected from accuracy and response latency analyses to ensure that any observed effects could be attributed to mnemonic rather than attentional or motivational factors. The statistical test used to

evaluate all effects of PI was repeated measures ANOVA with an alpha level of .05, using session means as individual data points. There were small but significant differences in mean overall accuracy between the monkeys in both experiments. In Experiment 1, Monkey F averaged 79.7% correct, and Monkey S averaged 83.5% (F[1,19] = 7.57, $p < .05$, $\eta_p^2 = .29$). In Experiment 2, Monkey F averaged 68.0% correct, and Monkey S averaged 74.3% (F[1,19] = 27.13, $p < .05$, $\eta_p^2 = .59$). However, accuracy for both animals was affected by PI in similar ways: we replicated each of the analyses below with *subject* (Monkey F, Monkey S) as an additional factor and found no significant interactions. Thus, the results below are given as the combined average of both animals. Any data points that were missing for a given analysis were substituted with the series mean (Bigelow & Poremba, 2013a; Roth, 1994). For example, in the unusual case that no incorrect button presses were made on nonmatch trials for a given session, and therefore no incorrect response latency data were available, the missing data point was estimated for the repeated ANOVA as the mean incorrect nonmatch response latency of the remaining sessions within the same stimulus set size condition in which such responses occurred.

3.2.2 Experiment 1: Results

An unanticipated but interesting initial observation in our data set was that the animals were more likely to quit responding before the final trial of experimental sessions using the smallest stimulus sets. Thus, the mean number of incomplete trials per session for the two-stimulus condition was 25.7, whereas for trial-unique sessions it was 12.7. (Figure 12). A repeated measures ANOVA confirmed that the effect of stimulus set size on early quitting was significant: $F(6, 234) = 3.82$, $p < .05$, $\eta_p^2 = .09$. Post hoc tests (Fisher's LSD, alpha level: $p < .05$) indicated

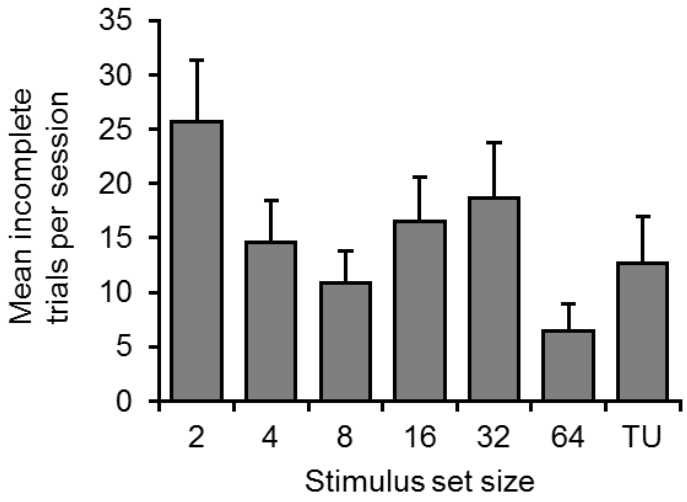

Figure 12. Average number of trials not completed (out of 128) as a function of stimulus set size. Early quitting was more frequently observed during sessions using the smallest stimulus set size (two). TU = trial unique. Error bars indicate the standard error of the mean.

that the number of incomplete trials was significantly greater for the two-stimulus condition than the 8, 64, or trial-unique conditions. This finding perhaps reflects the tendency of the animals to quit when the task becomes particularly difficult (see below).

As expected, overall accuracy was significantly affected by stimulus set size: $F(6, 234) = 37.38$, $p < .05$, $\eta_p^2 = .49$. There was also a significant linear trend indicating that accuracy increased as a function of stimulus set size: $F(1, 39) = 141.82$, $p < .05$, $\eta_p^2 = .78$. Accuracy was poorest (72%) during sessions for which only two stimuli were used and increased for the larger stimulus set sizes, reaching its maximum (88%) during sessions with trial-unique stimuli (Figure 13). By comparison, Mishkin and Delacour (1975) reported that visual DMS accuracy fell from 90% when trial-unique stimuli were used to 65% when only two stimuli were used. The differences in accuracy between sessions using only two stimuli versus trial-unique stimuli for the visual (25%) and auditory (16%) DMS tasks suggest that the influence of PI is roughly comparable, if perhaps somewhat less severe for auditory STM. However, because this comparison is limited by differences in subjects and experimental design, future research is needed to substantiate any difference in PI among sensory modalities.

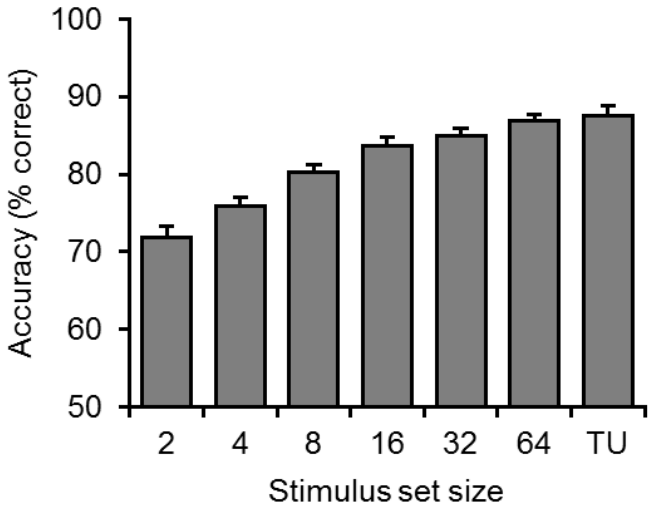

Figure 13. Overall accuracy improves as a function of stimulus set size. Accuracy for trial-unique sessions was significantly greater than sessions using stimulus set sizes of 32 or smaller. TU = trial unique. Error bars indicate the standard error of the mean.

In order to examine the influence of stimulus set size on response latency, it was first necessary to separate the data by trial type, since button presses on match trials reflected correct responses whereas button presses on nonmatch trials reflected incorrect responses. A repeated measures ANOVA with trial type and stimulus set size as factors resulted in significant main effects (trial type: $F[1, 39] = 109.19$, $p < .05$, $\eta_p^2 = .74$; stimulus set size: $F[6, 234] = 5.38$, $p < .05$, $\eta_p^2 = .12$) as well as a significant interaction ($F[6, 234] = 5.73$, $p < .05$, $\eta_p^2 = .13$). We also observed significant trends in the nonmatch and match response latency data. For incorrect responses on nonmatch trials, there was a significant linear trend indicating that the latency of incorrect responses increased as a function of stimulus set size ($F[1, 39] = 29.02$, $p < .05$, $\eta_p^2 = .43$). For correct responses on match trials, there was a significant quadratic trend ($F[1, 39] = 7.65$, $p < .05$, $\eta_p^2 = .16$). This outcome suggests that correct responses are made more slowly under high compared to moderate PI conditions. However, when PI is very low or absent,

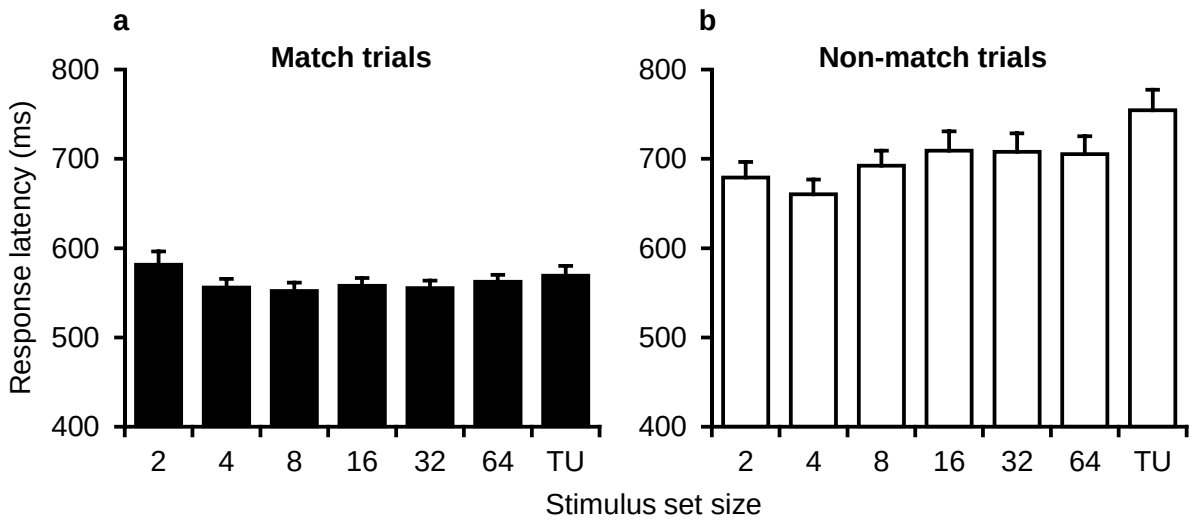

Figure 14. Response latency for match and nonmatch trials as a function of stimulus set size. **a** Response latency for match trials was significantly slower during sessions using the smallest set size (two), suggesting increased processing time for correct "match" responses under relatively high PI conditions. **b** By contrast, erroneous button presses on nonmatch trials were significantly slower for trial unique conditions, suggesting that errors are committed more quickly under high PI conditions. TU = trial unique. Error bars indicate the standard error of the mean.

response time is greater than under moderate amount of PI is present. Correct match responses were faster on average (562 ms) than incorrect responses (701 ms). Post hoc analyses (Fisher's LSD, alpha level: $p < .05$) revealed that nonmatch errors were slower for the trial-unique condition than all other stimulus set sizes (Figure 14). By contrast, correct match responses were slower for the two-stimulus set than several of the larger stimulus sets (4, 8, and 32). These results suggest that PI negatively impacts performance by increasing the speed with which nonmatch errors are made and by slowing correct match responses.

Evaluating accuracy for match and nonmatch trials separately revealed that the majority of the deficit in overall accuracy associated with the smaller set sizes (Figure 15) is due to increased errors or false alarms on nonmatch trials (Figure 16). A two factor ANOVA testing these differences resulted in significant main effects of trial type ($F[1, 234] = 19.13$, $p < .05$, η_p^2

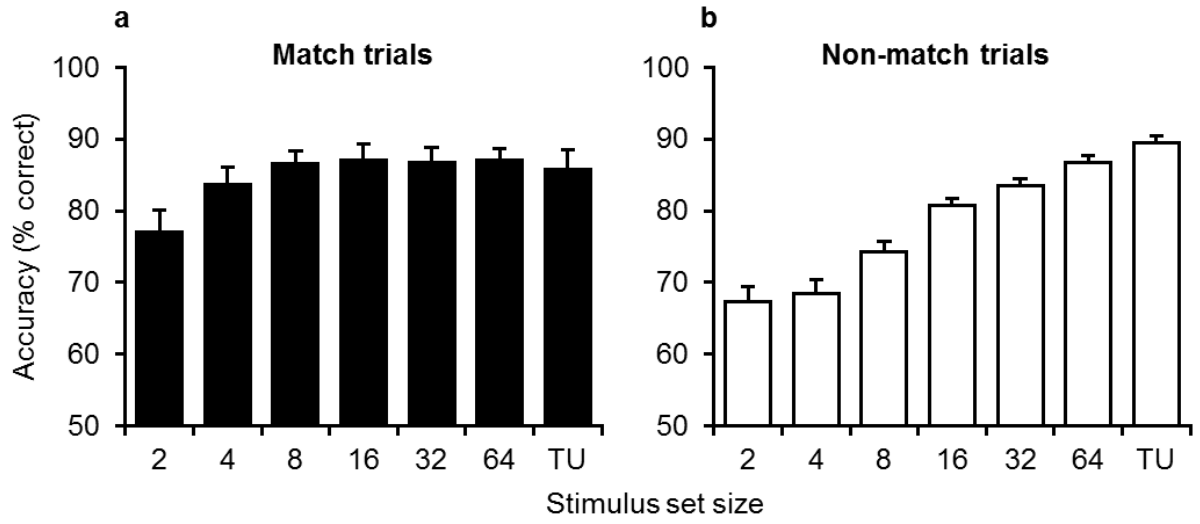

Figure 15. Accuracy for match and non-match trials as a function of stimulus set size. **a** Accuracy for match trials was relatively stable across stimulus set sizes, except that fewer correct match responses were made during sessions using the smallest sets (two and four). **b** PI associated with the smaller stimulus sets had a much larger impact on non-match accuracy, rising steadily from 67% in the two-stimulus set condition to 89% in the trial-unique condition. TU = trial unique. Error bars indicate the standard error of the mean.

= .33) and set size (F[6, 234] = 35.65, $p < .05$, $\eta_p^2 = .48$), as well as a significant interaction of these factors (F[6, 234] = 4.91, $p < .05$, $\eta_p^2 = .11$). Post hoc analyses (Fisher's LSD, alpha level: $p < .05$) indicated that nonmatch accuracy declined steadily from 89% for trial unique sessions to 67% for sessions repeatedly reusing the same two stimuli. This outcome is consistent with the view that increasing the frequency with which stimuli are presented from trial to trial leads subjects to commit more false positive errors because they might have heard a sound on a recent trial that "matches" the test stimulus on the current trial.

In contrast to the effects of PI on nonmatch trials, accuracy for match trials was only significantly reduced for the smallest set sizes (set size two: 77%; set size four: 84%), with no differences among set sizes eight through trial unique (range: 86 87%). One possible interpretation for why match accuracy might decrease under high PI conditions involves the

concept of feedback related changes in the criterion of familiarity with which *same*/*different* judgments are made as discussed by Wright (2006, 2007). According to Wright, subjects will make a "match" response only if the degree of familiarity evoked by a test stimulus exceeds the animal's familiarity criterion; otherwise a nonmatch response will be made. When trial unique stimuli are used, a test stimulus will only be familiar if it matches the sample stimulus from the same trial. In this context, adopting any criterion level of familiarity will suffice in making accurate *same*/*different* judgments. However, when stimuli are recycled from trial to trial, a test stimulus might evoke a certain level of familiarity by virtue of having been presented on a recent (and now irrelevant) trial, and not because it matches the sample stimulus of the current trial. Under these conditions, adopting a relatively lax criterion of familiarity will result in a high rate of false matches. Thus, one strategy for coping with a high degree of PI is to rely on a more rigid familiarity criterion such that only the most familiar test stimuli are accepted as matches. One of the predicted consequences of increasing the familiarity criterion is that, along with a decrease in false matches, the frequency with which true matches are rejected will also increase. Our data fit well with this prediction, inasmuch as more false negative errors were observed for the sessions with the greatest amount of PI.

To provide more direct evidence for shift in familiarity criterion resulting from PI, we investigated whether the rate of false matches (nonmatch errors) and false rejections (match errors) changed as the experimental session progressed for each stimulus set size. Accuracy data were separated by trial type and averaged for the first, second, third, and fourth quarters of the session (i.e., successive blocks of 32 trials). A three-way repeated ANOVA produced a significant interaction among trial type, stimulus set size, and trial block ($F[18, 702] = 3.12$, $p < .05$, $\eta_p^2 = .07$). As seen in Figure 17, the higher PI conditions led to a steady decrease in false

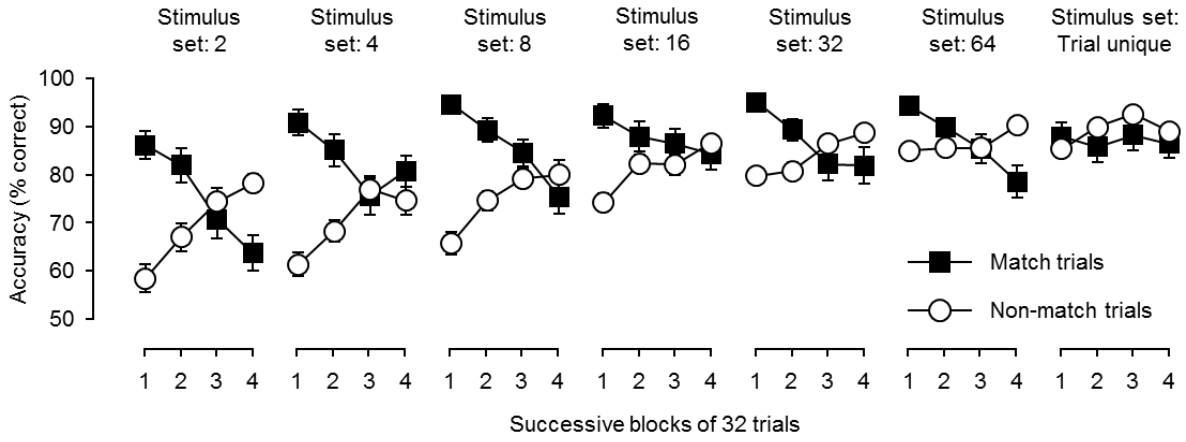

Figure 16. Progressive changes in accuracy by trial type for the first through the fourth quarters of the experimental session (successive blocks of 32 trials). Non-match errors became less frequent, whereas match errors became more common as the session progressed. The magnitude of this interaction diminished with increasing stimulus set size, such that no significant interaction was observed for trial-unique sessions. TU = trial unique. Error bars indicate the standard error of the mean.

matches throughout the session as well as a corresponding increase in false rejections. The magnitude of these reciprocal trends diminished as a function of stimulus set size, to the extent that there was no significant *trial type* × *trial block* interaction for the trial unique condition. The absence of a progressive change in accuracy during trial unique sessions is helpful in interpreting the *trial type* × *trial block* interactions observed for the remaining conditions because it argues against attentional or motivational explanations for the changes in error rate associated with the smaller set sizes. For example, it is unlikely that the observed decrease in match responses for the smaller set sizes reflects reward satiation because these changes were not observed when new stimuli were used for each trial. Rather, it seems likely that the shift toward fewer "match" responses (for both trial types) reflects the number of negative outcomes associated with false "match" responses. This interpretation lends support to Wright's (2006, 2007) suggestion that PI will gradually produce a proportional increase in the familiarity criterion that forms the basis for the match vs. nonmatch decision.

Curiously, the main effect of trial block and the interaction between trial block and stimulus set size were both nonsignificant (F[3, 117] = 1.36, $p > .05$, $\eta_p^2 = .03$, and F[18, 702] = 1.40, $p > .05$, $\eta_p^2 = .04$, respectively). This indicates that, although the rate of false match responses decreases throughout the session, the rate of true match responses decreases at an approximately equal rate. Thus, the benefits of reducing nonmatch errors are offset completely by the costs of increasing match errors, leading to zero net improvement during the session. We also examined overall accuracy for each stimulus set over the course of four successive blocks of five experimental sessions, but found no evidence of significant improvement: a repeated measures ANOVA returned neither a main effect of the block of five experimental sessions (F[3, 27] = 0.87, $p > .05$, $\eta_p^2 = .09$) nor an interaction between stimulus set size and experiment block (F[18, 162] = 1.22, $p > .05$, $\eta_p^2 = .12$). In several previous visual STM experiments, pigeons and monkeys have shown improvements in overall accuracy despite high PI conditions that gradually emerged with consistent, extended experience with small stimulus sets (D'Amato, 1973; Grant, 1975, 1976; Wright, 2007). It is possible that, since sessions with small and large stimulus sets were interleaved in our study, the animals lacked sufficient, consistent experience with high PI conditions to result in an adaptive adjustment of their familiarity criterion for match responses. A design presenting a block of multiple, consecutive sessions using a given stimulus set size, followed by a subsequent block using a different set size could be useful in determining if this is the case. Alternatively, it is possible that the lack of improvement over the course of the experiment is related to the difficulty monkeys have in establishing enduring memories in the auditory modality.

Two additional predictions are made by Wright's (2006, 2007) suggestion that animals' *same*/*different* decisions are under the influence of a familiarity criterion, which can be

modulated by error-related feedback. The first is that intertrial PI should have a graded effect on accuracy, such that stimulus repetitions separated by a large number of trials should have a smaller effect than stimulus repetitions separated by only a few trials or between adjacent trials. For example, fewer false match responses should occur if the test stimulus on trial *n* was most recently presented on trial *n* – 20 than if it was presented on trial *n* – 1 or *n* – 2. The second prediction is that the more rigid familiarity criterion that results from frequent exposure to PI should result in fewer nonmatch errors on trials for which the test stimulus has been presented on a relatively recent trial. Thus, even though training with smaller stimulus sets should yield relatively poor overall nonmatch accuracy, the frequency with which nonmatching test stimuli are erroneously accepted as matches by virtue of having appeared on a recent trial such as *n* – 1 or *n* – 2 should decrease (see Figure 9.5 in Wright, 2006 for hypothetical relationship between susceptibility to intertrial PI and familiarity criterion).

To test these predictions, we evaluated nonmatch accuracy on trials for which the test stimulus on trial *n* had most recently been presented on trials *n* – 1, *n* – 2, or *n* – 3. We only evaluated the effects of PI from trials *n* – 1 through *n* – 3 because, for the two-stimulus set size, the number of nonmatch trials for which the test stimulus most recently occurred on trial *n* – 4 or greater was insubstantial (0.1% of total trials). Similarly, we did not include sessions using stimulus set sizes of 16 or greater in this analysis because there were too few nonmatch trials for which the test stimulus had been presented on trials *n* – 1 through *n* – 3. Repeated measures ANOVA indicated significant main effects of stimulus set size ($F[2, 78] = 9.42$, $p < .05$, $\eta_p^2 = .20$) and PI location ($F[2, 78] = 70.72$, $p < .05$, $\eta_p^2 = .65$); the interaction was not significant ($F[4, 156] = 0.97$, $p > .05$, $\eta_p^2 = .02$). Consistent with similar analyses in several previous studies (Hartshorne, 2008; Wright et al., 2012), nonmatch accuracy increased steadily according to the

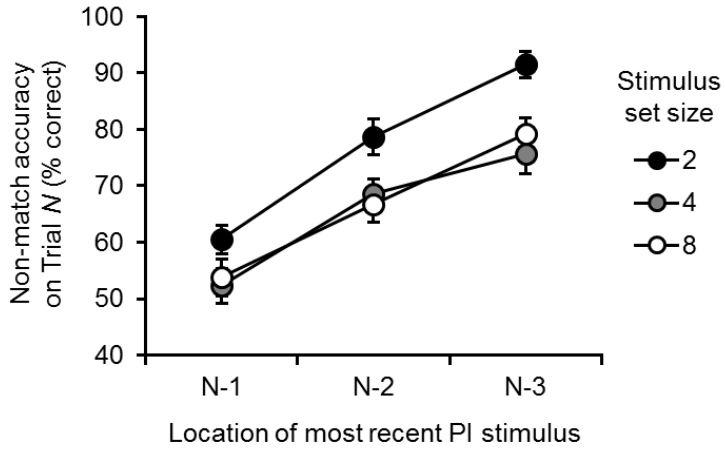

Figure 17. Non-match accuracy on trials for which the test stimulus had been presented on trials $n - 1$, $n - 2$, or $n - 3$. Although overall non-match accuracy was lowest for the two-stimulus set condition (see text), accuracy on the subset of trials with recent PI included in this analysis was greater for the two-stimulus set than for the four- or eight-stimulus sets. Intertrial PI had a graded effect on non-match accuracy for all three conditions. Error bars indicate the standard error of the mean.

distance of the most recent test stimulus repetition (Figure 17). Further, in confirmation of Wright's (2006, 2007) expectations, false match responses on trials with PI from trials $n - 1$, $n - 2$, and $n - 3$ were less likely for sessions that used only two stimuli compared to sessions that used four or eight stimuli. It should be noted that overall nonmatch accuracy was better for the larger set sizes (Figure 15) because of the larger number of trials in those sessions that did not have PI from trials $n - 1$, $n - 2$, or $n - 3$. Nevertheless, for the two-stimulus set size condition, accuracy on trial n when the test stimulus hadn't been presented since trial $n - 3$ reached a similar level (91%) to that observed for trial-unique sessions (89%).

To summarize the results of Experiment 1, PI produced by reusing a relatively small number of sounds from trial to trial decreases overall accuracy primarily by producing more false alarms on nonmatch trials. These nonmatch errors tend to be committed faster under high PI conditions than when trial-unique stimuli are used. The smallest set sizes, and therefore those that produced the most pervasive PI, also reduced the number of correct "same" decisions on match trials, and increased the amount of time before these decisions were made. These outcomes are consistent with Wright's (2006, 2007) prediction that subjects will adopt a more

stringent criterion of familiarity for "same" judgments when PI becomes highly saturated. Wright's notion of a familiarity criterion is further supported by several additional results from our study. First, as the experimental session progressed, subjects committed fewer false alarms on nonmatch trials, but fewer correct "same" responses on match trials. This effect was roughly proportional to the degree of PI caused by stimulus repetitions. Second, as with several previous studies (Bigelow & Poremba, 2013a; Hartshorne, 2008; Wright et al., 2012), PI originating from progressively more distant trials produced a graded effect on nonmatch accuracy. Finally, a greater degree of PI throughout the session (resulting from a smaller stimulus set size) resulted in fewer nonmatch errors on trials with PI originating from one of the three most recent trials.

3.3 Experiment 2: The role of intertrial interval

3.3.1 Experiment 2: Methods

Subjects

The subjects were the same as those used in Experiment 1.

Apparatus

The apparatus was the same one used in Experiment 1.

Procedure

Task

The task was similar to the one used in Experiment 1 except that the stimulus set size was

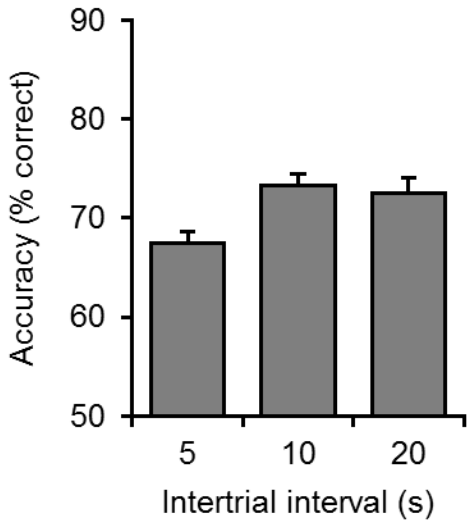

Figure 18. Overall accuracy as a function of the duration of the ITI. Accuracy improved significantly when the ITI was extended from 5s to 10 s, but no further advantage was gained by increasing the ITI to 20 s. Error bars indicate the standard error of the mean.

held constant and the ITI was manipulated between sessions. A set of four stimuli was used to produce an intermediate amount of PI. Stimuli for each session were randomly drawn from the same stimulus population that was used in Experiment 1.

Intertrial Intervals

Experiment 2 consisted of a parametric manipulation of the ITI duration. A fixed ITI of 5, 10, or 20 s was randomly selected for each session. Note that these values correspond to ISI:ITI ratios of 1:1, 1:2, and 1:4, respectively (see D'Amato, 1973). After a session was completed using one of the ITI values, it was not used again until the monkeys had completed sessions using the remaining two ITIs. As in Experiment 1, each animal completed a total of 20 sessions using each ITI.

Analysis

Analyses were similar to Experiment 1 except that the independent variable was the ITI.

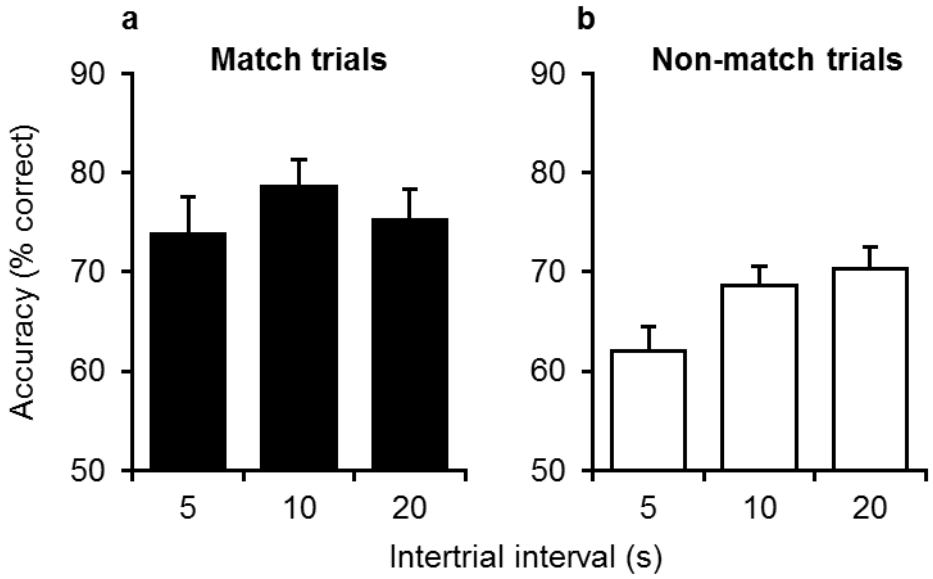

Figure 19. Accuracy for match and non-match trials as a function of the duration of the ITI. **a** There was no significant effect of ITI on match accuracy. **b** However, non-match accuracy improved significantly by increasing the ITI from 5 s to 10 or 20 s. Error bars indicate the standard error of the mean.

3.3.2 Experiment 2: Results

Repeated measures ANOVA indicated that ITI had a significant effect on overall accuracy (F[2, 78] = 7.41, $p < .05$, $\eta_p^2 = .16$). As revealed by post hoc tests (Fisher's LSD, alpha level: $p < .05$), the 5-s ITI resulted in lower overall accuracy (67%), whereas the 10-s and 20-s ITI conditions were equal (73%; Figure 18). Evaluating the differential effect of ITI on match and nonmatch accuracy again revealed that the decrease in overall accuracy was caused primarily by an increase in nonmatch errors at the shortest ITI (Figure 19). Repeated measures ANOVA revealed significant main effects of ITI (F[2, 78] = 6.50, $p < .05$, $\eta_p^2 = .14$) and trial type (F[1, 39] = 7.44, $p < .05$, $\eta_p^2 = .16$), but there was no significant interaction between the two factors (F[2, 78] = 0.60, $p > .05$, $\eta_p^2 = .02$). Post hoc analyses indicated that there were no significant differences among the three ITI conditions for match accuracy. As with overall accuracy, nonmatch

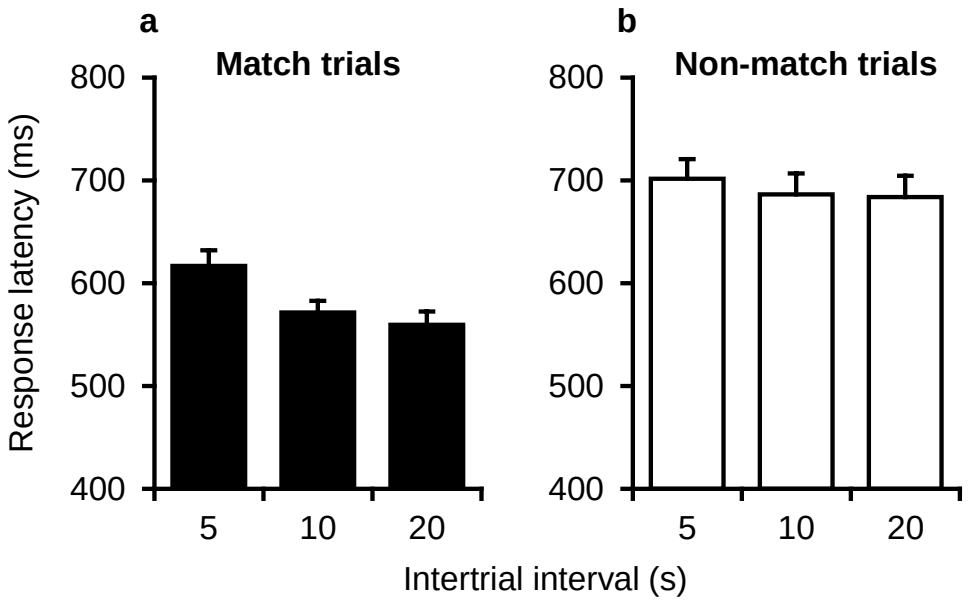

Figure 20. Response latency for match and non-match trials as a function of the duration of the ITI. **a** Correct match responses were significantly slower for sessions using the shortest ITI (five seconds). **b** No significant effect of ITI was found for erroneous responses on non-match trials. Error bars indicate the standard error of the mean.

accuracy was significantly reduced for the 5-s ITI condition, whereas the 10-s and 20-s ITI conditions did not differ from each other.

As in Experiment 1, we examined response latency by separating the data by trial type. Repeated measures ANOVA revealed significant effects of trial type (F[1, 39] = 95.04, $p < .05$, $\eta_p^2 = .71$) and ITI (F[2, 78] = 6.45, $p < .05$, $\eta_p^2 = .14$), as well as a significant interaction (F[2, 78] = 4.22, $p < .05$, $\eta_p^2 = .10$). As in Experiment 1, correct match responses were faster (582 ms) than erroneous responses on nonmatch trials (690 ms). Unlike Experiment 1, post hoc tests revealed no differences for nonmatch errors as a function of ITI (Figure 20). For correct match responses, slower response latency was observed for the 5-s ITI condition, but no differences were observed between sessions using a 10-s and 20-s ITI. The latter outcome is consistent with results from Experiment 1 in suggesting that the decision time for correct match trials increases when PI becomes saturated.

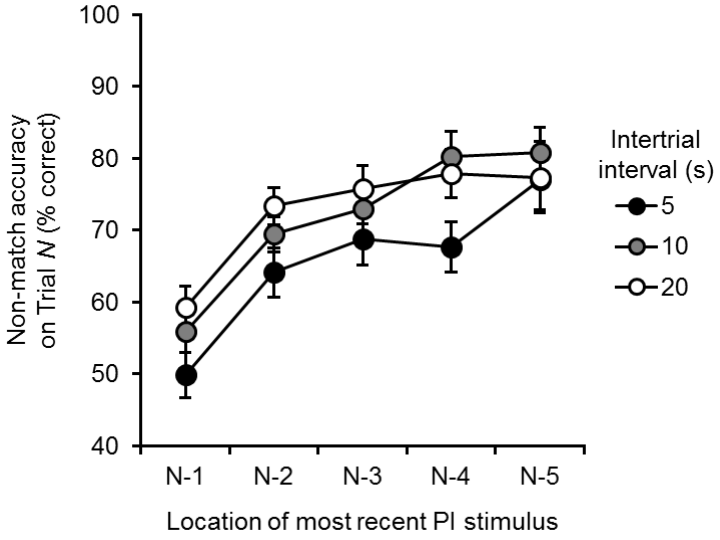

Figure 21. Intertrial PI as a function of the duration of the ITI. PI had the largest effect in the 5-s ITI condition, whereas the 10-s and 20-s ITI conditions were similar. PI had a significant influence spanning multiple trials for all three ITI conditions. Error bars indicate the standard error of the mean.

Although nonmatch accuracy increased significantly by extending the ITI from 5 s (62%) to 10 s (69%) or 20 s (70%), it was still well below nonmatch accuracy achieved in trial-unique sessions in Experiment 1 (89%). This outcome suggested that there might still be substantial PI caused by reusing stimuli even when trials have been separated by as much as 20 s. Thus, as in Experiment 1, we directly investigated the influence of intertrial PI by evaluating nonmatch accuracy on trial *n* as a function of the trial on which the test stimulus was most recently presented. Since a four-stimulus set was used in all conditions, it was possible to evaluate the influence of PI originating from trials $n - 1$ through $n - 5$ (the most recent PI stimulus was found on trial $n - 6$ or greater for only 9.0% of the nonmatch trials). As seen in Figure 21, accuracy improved for each ITI condition as the number of trials between stimulus repetitions increased. A repeated measures ANOVA resulted in a main effect of ITI that was of borderline significance ($F[2, 78] = 3.10$, $p = .05$, $\eta_p^2 = .07$) and a main effect of PI location ($F[4, 156] = 34.46$, $p < .05$, $\eta_p^2 = .47$), but the interaction of these factors was not significant ($F[8, 312] = 0.84$, $p > .05$, $\eta_p^2 = .02$). These results indicate that, unlike several experiments in pigeons and rats, which eliminated visual PI by increasing the ITI to 15 or 20 s (Dunnett & Martel, 1990; Grant, 1975; Hogan et al.,

1981, but see Wright et al., 2002), auditory PI in monkeys can influence subsequent trials even when separated by up to 20 s.

Post hoc tests indicated that, for each condition, nonmatch accuracy improved when the test stimulus had most recently occurred on trial $n - 2$ through $n - 5$ compared to when it had occurred on the previous trial, $n - 1$. For sessions using the 5-s ITI, accuracy on trials for which the test stimulus had most recently occurred on trial $n - 5$ was significantly higher than when it had occurred on either trial $n - 1$ or $n - 2$. Similarly, for the 10-s ITI condition, accuracy was greater when the test stimulus hadn't been presented since trial $n - 4$ or $n - 5$ than when it occurred on trial $n - 1$ or $n - 2$. However, for the 20-s ITI condition, there was no statistically significant difference in accuracy when the test stimulus had most recently occurred on trials $n - 2$ through $n - 5$. This suggests that increasing the ITI from 10 to 20 s may slightly reduce the extent of intertrial PI, although this difference was insufficient to result in any significant change in overall accuracy (Figure 18) or averaged nonmatch accuracy (Figure 19).

In a prior study, we examined the extent of intertrial PI in monkeys performing an auditory DMS task with a stimulus set size of eight and a retention interval of 5 s (Bigelow & Poremba, 2013a). A variable ITI averaging approximately 10 s was used, making it most similar to the 10-s ITI condition in Experiment 2. In that study, we found that the monkeys were more likely to commit errors on nonmatch trials when the test stimulus was repeated after as many as 10 trials. By contrast, in the current study (Experiment 2, 10-s ITI), the effects of PI appear to reach asymptote by about trial $n - 4$. This difference could plausibly result from an increase in the familiarity criterion for "match" responses resulting from the relatively high PI conditions produced by the smaller stimulus set size (four compared to eight). Specifically, if the monkeys were using a relatively lax criterion of familiarity for "match" decisions in the eight-stimulus set

condition, they would be more likely to erroneously accept a nonmatching test stimulus as a "match" if it had occurred on trial $n - 4$ or $n - 5$, or even earlier.

To summarize the results of Experiment 2, we observed that overall accuracy increased, if only to a small extent, by extending the ITI from 5 s to 10 s, but no additional improvement was gained by increasing the ITI to 20 s. These outcomes are comparable to a study of visual STM in monkeys reported by Jarrard and Moise (1971), in which a small increase in accuracy was produced by increasing the ITI from 5 to 15 s, but not by further extending the ITI to 30 or 60 s. We further found that the effects of intertrial PI are substantial even when trials are separated by 20 s. Although these results are generally consistent with previous literature, the decay of PI from previous trials produced by increasing the ITI was less than we had initially expected. Given the significant intertrial effects of PI in each condition, it is likely that future studies will require an ITI substantially longer than 20 s in order to completely eliminate auditory PI in monkeys.

3.4 Discussion

Like other forms of memory, our experiments show that auditory STM in monkeys is susceptible to PI caused by reusing the same stimuli for multiple trials within an experimental session. In Experiment 1, PI was most severe when only a very small number of stimuli (e.g., two or four) were used throughout the session. The effects of PI diminished steadily as the stimulus set size increased, and maximum accuracy was observed when new stimuli were used for each trial (Figure 13). In Experiment 2, PI was modestly attenuated by increasing the ITI from 5 s to 10 s,

but no additional advantage was gained by increasing the ITI to 20 s (Figures 18-21).

In both experiments, PI reduced overall accuracy primarily by increasing the number of erroneous "match" responses on nonmatch trials (Figures 15 and 19). This finding is similar to what has been reported in experiments using list memory tasks in which a small number of stimuli are reused from trial to trial (Wright, 1998). In these tasks, a list of several sample stimuli is presented, each separated by a brief ISI. The last item in the list is followed by a retention interval, after which a single probe stimulus is presented, and the subject must indicate whether the probe had been presented in the list ("same") or not ("different"). In both visual and auditory list memory tasks in monkeys, using a small stimulus set throughout the session, and thereby increasing item repetition among trials, increases the rate of erroneous "same" responses on trials in which the probe was different from the items presented in the list (Sands & Wright, 1980; Wright, 1999). This is likely because, although the probe did not match one of the list items from the current trial, it may have matched a list item from a recent trial. In other words, the error being committed by the subjects seems to be forgetting whether the probe occurred on the current trial, or some previous, now irrelevant trial (Wright, 2006). In both the current study as well as previous DMS and list memory experiments, errors on nonmatch trials are most likely if the probe had been presented on the immediately previous trial, and become less likely as the number of trials since the probe had been presented increases (Bigelow & Poremba, 2013a; Hartshorne, 2008; Wright et al., 1986, 2012).

Several results from Experiment 1 provide support for Wright's (2006, 2007) view that a criterion threshold of familiarity is used to make the same vs. different choice, and that this familiarity criterion can be increased as a result of committing frequent false match errors. Although PI consistently increased the false alarm rate on nonmatch trials, the extremely

saturated PI conditions also reduced the number of correct "same" judgments on match trials (Figure 15). Additional analyses revealed a steady decrease in "match" responses as the session progressed for both match and nonmatch trials that was most pronounced for the highest PI conditions (Figure 16). Further, although overall nonmatch accuracy was lowest for the two-stimulus set, nonmatch accuracy on trials with recent PI (originating from trials $n - 1$, $n - 2$, or $n - 3$) was higher for the two-stimulus set than the four- or eight-stimulus sets (Figure 17). Each of these results can be seen as a consequence of increasing the familiarity criterion for "match" responses that reflects the degree of PI within an experimental session. These observations fit well with several previous animal studies showing gradual improvement over time under consistently high PI conditions (D'Amato, 1973; Grant, 1975, 1976; Wright, 2007), and with data from humans showing that their familiarity criterion can be modified by changes in stimulus presentation frequency (Yonelinas, 2002).

In general, our results show that the effects of stimulus set size and ITI in auditory STM in monkeys are similar to that which has been previously reported in visual studies. In Experiment 1, subjects reached 88% overall accuracy for sessions that used trial-unique stimuli, which compares favorably with the 90% accuracy reported by Mishkin and Delacour (1975) for monkeys performing a trial-unique visual DMS task. When only two stimuli were repetitively used throughout the session, accuracy fell to 72% in our study and 65% in Mishkin and Delacour's visual study. In Experiment 2, overall accuracy increased from 67% when the ITI was 5 s to 73% when the ITI was increased to 10 or 20 s. This modest increase in accuracy is similar to that observed by Jarrard and Moise (1971), who reported that increasing the ITI from 5 to 15 s increased accuracy by 5 to 10% for various retention intervals in monkeys performing visual DMS. Moreover, like our study, in which no additional benefit was gained by extending

the ITI from 10 to 20 s, Jarrard and Moise also observed no significant increase in accuracy after increasing the ITI from 15 to 30 or 60 s. In both experiments, we observed significant effects of intertrial PI that extended beyond immediately adjacent trials. PI in visual STM in monkeys has similarly been observed to span multiple trials (Wright, 2007). In Wright's study as well as ours, the effects of PI diminished as a function of number of trials separating the current trial from the source of the PI. However, a direct comparison of the extent and impact of intertrial PI in these studies is complicated by differences in experimental parameters. For instance, in Wright's study, the influence of PI was far more severe when a relatively long (20 s) compared to a short (1 s) retention interval was used, whereas our study used an intermediate retention interval (5 s). These differences notwithstanding, the foregoing results imply that PI in auditory and visual STM in monkeys are at least qualitatively similar. Additional studies using similar task parameters, and ideally the same subjects, are needed to determine whether any quantitative differences exist between PI in auditory and visual STM.

In view of our experimental outcomes as well as previous studies, future attempts to maximize accuracy by decreasing PI should avoid recycling stimuli among trials as much as possible. Ideally, new stimuli should be used for each trial. However, for some experimental paradigms, presenting multiple trials using the same stimuli is unavoidable. For instance, neurophysiological investigations of visual and auditory STM require multiple repetitions of each stimulus in order to establish reliable stimulus-evoked neuronal responses (Bigelow & Poremba, 2013a). Under such circumstances, a relatively long ITI may help reduce PI by increasing the decay time for irrelevant memory traces from previous trials. In addition to these factors, previous studies have indicated that PI may be reduced by increasing the stimulus exposure time (Grant, 1975) and by reducing the retention interval (Meudell, 1977; Wright,

2007; Wright et al., 2012; Zentall & Hogan, 1974). Thus, at minimum, optimal performance on DMS and similar tasks depends on the stimulus set size, ITI, retention interval, and stimulus exposure time, as well as interactions among these variables (see also van Hest & Steckler, 1996).

In summary, auditory memory in monkeys is highly susceptible to PI, which can be minimized by increasing the number of new stimuli that are presented throughout the trial. To a lesser extent, PI may be reduced by allowing an adequate interval of time between trials for sessions that reuse stimuli from trial to trial. Whether or not the monkeys will make a "same" judgment (whether correct or incorrect) may depend on whether the test stimulus exceeds a threshold level of familiarity, which may result either from having been recently presented as the sample for the current trial or from a presentation on a prior, and currently irrelevant trial. Following error-related feedback from having incorrectly chosen a nonmatching test stimulus as a "match", this threshold of familiarity may become more stringent in order to minimize future nonmatch errors. These findings expand our understanding of the variables governing auditory STM in monkeys. Further studies directly comparing auditory, visual, and tactile STM are needed in order to reveal the extent to which these findings generalize across sensory modalities.

Chapter 4: A comparison of auditory, visual, and audiovisual short-term memory in nonhuman primates

4.1 Introduction

Studies of STM in nonhuman primates have traditionally focused on the visual sensory modality (Colombo & D'Amato, 1986). More recent studies have begun to characterize auditory STM in nonhuman primates, and have consistently revealed an asymmetry in their mnemonic abilities for auditory and visual stimuli. Thus, whereas monkeys readily learn visual STM tasks within a few dozen training sessions, they often require several hundred training sessions to learn comparable auditory STM tasks (Fritz et al., 2005; Wright, 2007). Monkeys are also capable of performing visual STM tasks at substantially longer retention intervals (hours to days; Overman & Doty, 1980) than auditory STM tasks (15-50 s; Fritz et al., 2005; Kojima, 1985). Finally, recent experiments have revealed that monkeys are more susceptible to intratrial distractor stimuli during auditory STM tasks (Scott et al., 2012, 2013).

In spite of the progress in understanding STM processing in the visual and auditory modalities, as of yet, there are no published studies comparing performance for each unimodal stimulus type to STM for bimodal, audiovisual stimuli. Such studies could make a significant contribution to the literature for at least two reasons. First, integrating and retaining audiovisual events is thought to be important for several aspects of primate ethology including conspecific communication and predator evasion (Romanski & Averbeck, 2009). Second, in the human memory literature, a performance advantage has been repeatedly found for audiovisual

memoranda compared to unimodal auditory or visual memoranda (e.g., Delogu et al., 2009; Mastroberardino et al., 2008; Shams & Seitz, 2008; Thompson & Paivio, 1994). Thus, a comparison of STM for auditory, visual, and audiovisual stimuli in nonhuman primates could potentially reveal a processing advantage for audiovisual stimuli, which could be relevant to other domains of study such as audiovisual communication.

An additional aspect of STM that has only recently been investigated in the auditory modality is PI, which occurs when memory processing at one time during the task interferes with subsequent memory processing. Our lab has recently conducted studies of auditory PI using the *same*/*different* variation of the DMS task (Chapter 3; Bigelow & Poremba, 2013a, 2013b), wherein subjects indicate whether sample and test stimuli separated by a retention interval are identical (match trials) or nonidentical (nonmatch trials). In several experiments, we observed that subjects were more likely to incorrectly accept a nonmatching test stimulus as a match if it had also been presented on a recent, previous trial. The effects of intertrial PI were graded, such that errors diminished as a function of the number of trials separating the current nonmatching test stimulus from its most recent presentation. Although these outcomes were qualitatively similar to previous reports of visual PI in monkeys (Wright et al., 1986), humans (Hartshorne, 2008), and other animals (Wright et al., 2012), differences in task procedures and contingencies precluded a direct comparison of the effects of PI in auditory and visual STM. One possibility is that auditory STM may be more susceptible to PI, which could be related to subjects' lower overall auditory STM performance. Alternatively, because monkeys appear to be able to retain visual information for longer amounts of time than auditory information (Fritz et al., 2005; Kojima, 1985; Overman & Doty, 1980; Scott et al., 2012), PI stemming from stimulus presentations on previous trials may be more perseverant for visual STM.

A comparison of STM for auditory, visual, and audiovisual memoranda would allow two additional analyses of PI which have not previously been reported in the literature. First, comparison of PI on audiovisual and unimodal trials could reveal potential audiovisual interactions within PI effects. Such interactions, if any, could result in more or less severe audiovisual PI (reflecting the strongest or weakest unimodal PI effects, respectively), or perhaps intermediate PI effects reflecting the average influence of auditory and visual PI. Second, analyses accounting for the modality of the PI source trial could reveal additional differences in susceptibility to PI produced by recent previous presentations of unimodal or audiovisual stimulus types (e.g., PI may depend on whether the nonmatch test of an auditory trial had recently occurred in the context of a previous auditory or audiovisual trial). Further, analyses of PI on unimodal trials with respect to recent presentations of corresponding memoranda of the opposite modality (auditory-visual, visual-auditory), could reveal whether subjects formed implicit crossmodal associations between auditory-visual stimulus components, which were consistently paired throughout the study. For instance, supposing sound *x* was consistently paired with image *y* whenever it occurred in the context of an audiovisual trial, and if sound *x* presented alone on auditory trial *k* then disrupted performance on a subsequent visual trial *n* in which the nonmatch test stimulus was image *y*, it could be inferred that the recent presentation of sound *x* resulted in the associative activation of image *y*, leading to the inappropriate acceptance of nonmatching image *y* as a match.

In the current study, we trained monkeys to perform a concurrent audiovisual DMS task in which the memoranda for each trial were auditory, visual, or audiovisual. The task was designed to address two primary questions. First, we compared accuracy between unimodal (auditory, visual) and audiovisual trials to test the hypothesis that there is a bimodal advantage in

STM for nonhuman primates as there is in humans. Second, we investigated whether the influence of PI was modulated by the stimulus modality of the current or PI source trial, which might help explain differences in overall performance by modality.

4.2 Methods

Subjects

Three rhesus macaques (*Macaca mulatta*), each 18 years old at the beginning of the experiment, served as subjects (Monkey V: female; Monkeys F and S: male). As part of ongoing studies, each subject had previously learned an auditory DMS task (similar to the ones described in Chapter 3 and depicted in Figure 22A), and were regularly tested with this paradigm over approximately 10 years (e.g., Ng et al., 2009). The male monkeys (Monkeys F and S) were the same subjects tested in Chapter 3. Subjects were individually housed with *ad libitum* access to water and controlled feeding schedules, under a 12:12 light:dark cycle. Animals were fed after daily training (Harlan monkey diet plus fruit, vegetables, and treats) and maintained above 85% free feeding weight. All procedures conformed to National Institutes of Health guidelines and were approved by the University of Iowa IACUC.

Apparatus

Experiments were conducted inside a sound attenuation chamber (Industrial Acoustics Company, Bronx, NY), where subjects sat in a custom primate chair allowing free arm movements. Because behavioral data were occasionally collected concurrently with neurophysiological recordings, the subjects' head position was fixed during the session by a

holder attached to the chair (see Chapter 5 Methods for additional details). Images were presented centrally on a monitor (MicroTouch C1500SS, 3M Touch Systems, Methuen, MA) located 32 ± 5 cm from subjects' eye position, and sounds were delivered through a central speaker directly above the monitor. An acrylic response button (8 × 8 cm) was positioned centrally below the monitor, which was equipped with LED backlights to signal the response window. Rewards were delivered by a pellet dispenser (Med Associates, Georgia, VT) into a dish immediately below the response button. An overhead LED provided low-level illumination throughout the session, and a second LED served as a trial segregation cue during the intertrial interval (ITI). Custom LabVIEW software (National Instruments, Dallas, TX) controlled all task events.

Short-term memory task

All training and testing conditions used the *same*/*different* version of the DMS task depicted in Figure 22. Following a 5-s ITI, each trial presented 0.5-s sample and test stimuli separated by a 1.5-s retention interval. Following a 0.5–1-s delay after the test stimulus, the response button was illuminated, signaling a 1.5-s response window. The delay separating the test stimulus and response window was included to separate stimulus-evoked and motor-related activity in the neurophysiological recordings. Previous studies suggest that a brief pre-response delay (1–2 s) is not detrimental to performance, but may provide modest facilitation (Lemus et al., 2007). Subjects were trained to press the button following identical test stimuli (match), but to withhold button presses following nonidentical test stimuli (nonmatch). Responses were subject to an asymmetric reinforcement contingency wherein correct match responses were rewarded with a small food pellet and incorrect button presses on nonmatch trials were punished by a 2-s dark timeout. Responses outside the response window aborted the trial. Memoranda

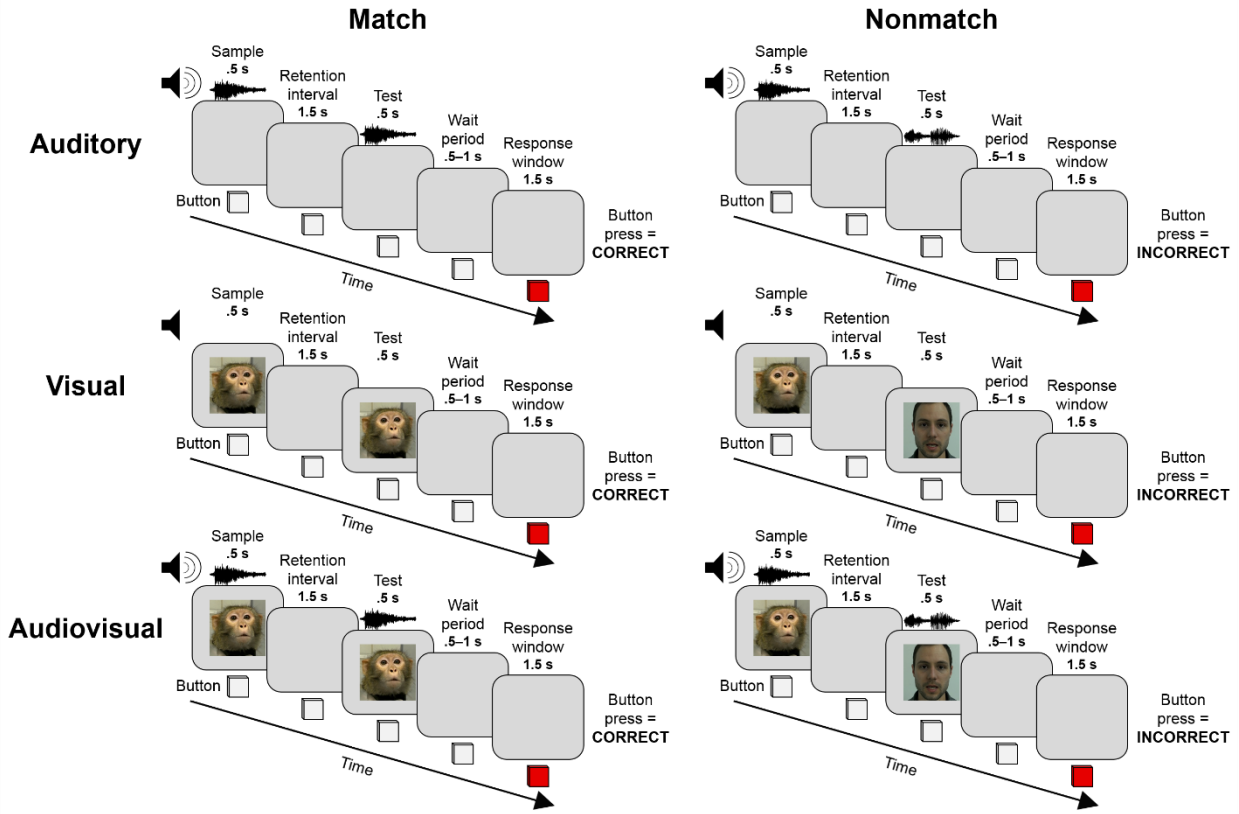

Figure 22. Diagram of the concurrent audiovisual DMS task. Each trial consisted of 0.5-s sample and probe stimuli separated by a 1.5-s retention interval. The response button was illuminated following the probe to signal a 1.5-s response window. Subjects were trained to press the button following identical probes (match trials), but to withhold button presses following nonidentical probes (nonmatch trials). Responses outside the response window aborted the trial. Memoranda comprised (A) sounds for auditory trials, (B) images for visual trials, and (C) sounds and images presented simultaneously for audiovisual trials. For audiovisual nonmatch trials, both the sound and image of the probe differed from the sample. Each of the six trial types were presented within experimental sessions equally often in random order. Subjects had previous experience with auditory DMS tasks similar to the one shown in (A), and were trained with the visual DMS task depicted in (B) before being tested with the concurrent audiovisual task comprising all three stimulus presentation formats.

comprised sounds for auditory trials (Figure 22A), images for visual trials (Figure 22B), and sounds and images presented simultaneously for audiovisual trials (Figure 22C). For audiovisual nonmatch trials, both the sound and image of the test stimulus differed from the sample. Within sessions, each trial type (match, nonmatch; auditory, visual, audiovisual) occurred equally often in random order. For occasional sessions in which subjects ceased responding 20 or more trials

before the end of the programmed session, the last response was retroactively assigned as the end of the session. Training sessions comprised 120–300 trials, and were retained for analysis only if mean overall accuracy exceeded chance (χ^2 test, $p < 0.05$).

Stimuli

Stimuli serving as memoranda varied from day to day and comprised 4–20 exemplars drawn from a collection of conspecific monkey faces and vocalizations, human faces and vocalizations, heterospecific animal faces and vocalizations, and abstract images and sounds. Each exemplar included image and sound components, presented separately on unimodal trials and together on audiovisual trials as described above. Exemplars were presented equally often as the sample and test (match and nonmatch). Conspecific monkey faces and vocalization stimuli were created from video recordings collected in the primate colony at the University of Iowa (acquired at 30 frames/s; frame size: 1920 × 1080 pixels), and from other labs. Vocalization sounds were extracted from the auditory track of the video recordings. Images that corresponded to the vocalizations were created by selecting one of the frames from the midpoint of the vocalization and cropping the frame around the face. Human and heterospecific vocalization stimuli were created in a similar manner from video recordings obtained in the lab and from open source videos available online. Unlike the face and vocalization pairings, the auditory and visual components of the synthetic/abstract stimuli were obtained from independent sources. The auditory components comprised complex, artificial sounds generated by electronic synthesizers or downloaded from abstract sound categories (e.g., "science fiction") of sound effects collections available online. The visual components comprised complex, abstract images (e.g., fractals) obtained from open source image collections available online. All image-sound pairings for audiovisual trials were kept constant within and across sessions. Images were normalized for

mean luminance (Photoshop, Adobe Systems, San Jose, CA), and presented centrally at eye level in full color at 20 ± 5° viewing angle. Sounds were normalized for root-mean-square (RMS) amplitude (Audition, Adobe Systems, San Jose, CA), and presented centrally at 75 ± 5 dB measured from subjects' ear level.

4.3 Results

Task acquisition

Before being tested with the concurrent audiovisual DMS task, several approaches were used to train the monkeys with the DMS rule using visual stimuli (Figure 22B). The first training approach tested whether the monkeys would exhibit spontaneous generalization from auditory to visual DMS. After establishing a baseline of five auditory DMS training sessions, subjects were tested with a visual DMS task that was identical to the auditory task except that memoranda were images instead of sounds. As seen in Figure 23 (i, iv, viii), each subject's performance fell to chance in the visual DMS test session. In an earlier study, African green monkeys learned an auditory DMS task in which memoranda consisted of clicks of varying frequencies, after which their performance immediately transferred to an analogous visual task using flashing lights of identical frequencies to the clicks (Stepien et al., 1960). Our subjects' failure to generalize may reflect increased difficulty of transfer testing with complex sounds and images, which were less physically similar to each other than the clicks and flashes.

Having failed to directly transfer from auditory to visual DMS, we next attempted a fading procedure in which sounds were initially presented at full volume with images on each

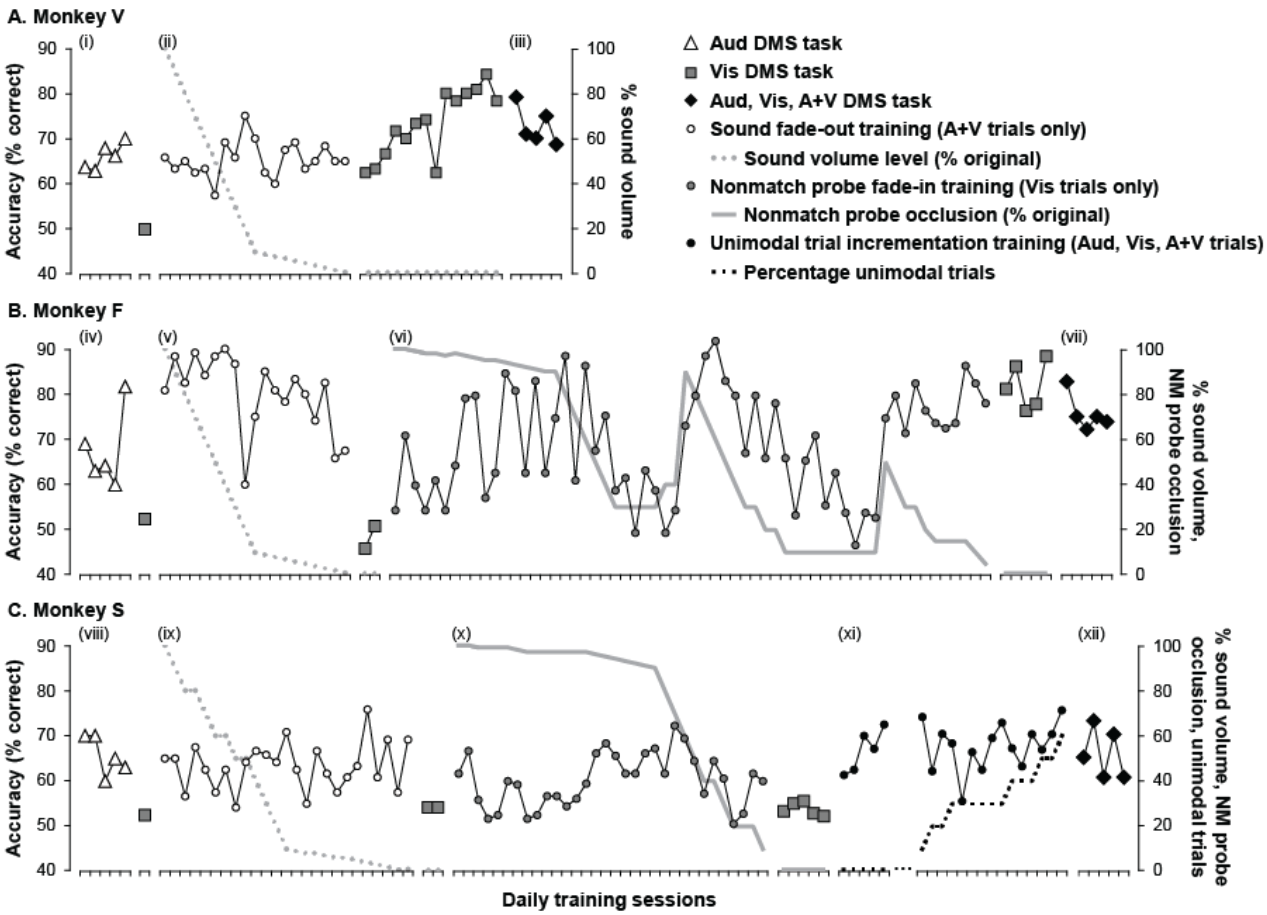

Figure 23. Training and acquisition of the audiovisual DMS task. Each subject had prior experience with auditory DMS, but had to learn the visual DMS rule before being tested with the full audiovisual DMS task. (A) After failing to transfer directly from auditory to visual DMS (i), Monkey V was trained with a DMS task in which compound auditory-visual memoranda were presented on each trial (ii). The volume of the sounds was reduced after each successive day of training where performance exceeded chance (χ^2 test, $p < 0.05$), eventually leaving only images as memoranda (sound fade-out training). Monkey V successfully learned the visual DMS rule using this approach, and required no additional training to perform the concurrent audiovisual DMS task (iii). (B) Monkey F failed to acquire the visual DMS rule through both the direct transfer test (iv) and sound fade-out training (v). The next training approach was a visual DMS task in which the probe image on nonmatch trials was initially occluded, and then gradually faded in after successive training sessions of above-chance performance (nonmatch probe fade-in training). Monkey F eventually acquired the visual DMS rule with this approach (vi), and was able to perform the concurrent audiovisual DMS task immediately thereafter (vii). (C) After failing to learn visual DMS through direct transfer (viii), sound fade-out training (ix), and nonmatch probe fade-in training (x), Monkey S was trained with a task in which all trials were initially audiovisual, after which increasing proportions of randomly-interleaved unimodal auditory and visual trials were introduced after successive training sessions of above-criterion (70%) performance (xi). The stricter performance criterion was adopted to stabilize performance before advancing to more challenging steps. Note that the ellipsis and extra space in (xi) represent discontiguity between 0% and 10% unimodal training sessions due to treatment and monitoring for illness by veterinary staff. Monkey S gradually learned the audiovisual DMS task with this method (xii), eventually surpassing performance of the other subjects (see Figure 24).

trial. Following sessions in which performance exceeded chance, the volume of the sounds was reduced by 10% of the original sound pressure level (SPL; GoldWave, GoldWave, Inc., St. John's, Newfoundland, Canada). Once the volume reached 10% of the original SPL, it was reduced for subsequent sessions in 1% steps until only the images remained ("sound fade-out training"). This approach worked well for Monkey V (Figure 23 ii). Indeed, Monkey V was granted additional visual DMS testing sessions until performance reached asymptotic levels that exceeded earlier auditory DMS performance. Having acquired the visual DMS task, Monkey V was able to perform the concurrent audiovisual DMS task without further training (Figure 23 iii).

By contrast, Monkeys F and S failed to learn visual DMS through this approach, despite progressing rapidly through the sound attenuation steps (Figure 23 v, ix). In a previous study, rhesus monkeys failed to transfer from visual to auditory DMS using a similar fading procedure in which images were gradually faded out, leaving only sounds as memoranda (Wright et al., 1990). The success of one of our subjects to transfer via a comparable fading procedure suggests that intermodal transfer may be easier from audition to vision than vice versa. Nevertheless, the other subjects' failure to transfer despite similar prior experience, age, and training methods highlights individual differences in intermodal transfer and the general difficulty thereof. Incidentally, it also indicates that monkeys are capable of performing auditory DMS above chance at extremely low sound levels (sounds 1% of their original SPL were measured at 37 ± 2 dB, just above the 34 ± 1 dB ambience).

After failing to learn visual DMS through direct transfer and sound fade-out training, Monkeys F and S were next trained with a visual DMS task in which the nonmatch test image was initially 100% occluded (LabVIEW brightness manipulation tools), and then gradually faded in after sessions in which performance exceeded chance ("nonmatch test fade-in training").

Occlusion was reduced in 1% steps until reaching 90% occlusion, after which it was further reduced in 10% steps until reaching full image brightness. The monkeys' original auditory DMS training followed a similar algorithm: volume of the nonmatch test sounds was initially set to zero, and then gradually faded in as subjects achieved criterion performance (A Poremba, unpublished data; cf. Scott et al., 2012; Wright et al., 1990). The increased visual contrast between the full-brightness sample and the faint or absent test was intended to enhance the perceptual difference between the nonmatching images and thereby facilitate learning the *same*/*different* rule in the visual modality. Both monkeys' performance exceeded chance by the second training session using 100% nonmatch test occlusion (Figure 23 vi, x). Monkey F gradually progressed through decreasing occlusion levels, but required several repeated sessions. Indeed, we opted to reset training twice at higher occlusion levels after performance fell precipitously following stretches of promising performance (>80%). After eventually achieving performance well above chance on the nonhandicapped visual DMS task, Monkey F was then able to perform the concurrent audiovisual DMS task without additional training (Figure 23 vii). Monkey S progressed reasonably through the decreasing occlusion steps, although like Monkey F, required several repeated sessions (Figure 23 x). Unexpectedly, although Monkey S was able to perform the task with only 10% nonmatch test occlusion, performance fell below significance when tested with the nonhandicapped visual DMS task.

Because Monkey S had not shown the same high levels of performance in earlier stages of nonmatch test fade-in training observed for Monkey F, we opted not to reset training at higher occlusion levels, but to instead attempt an alternative training method. In the first stage of this approach, memoranda comprised both sounds and images for every trial. Then as accuracy reached a criterion of 70% correct (the higher criterion was intended to stabilize performance

before progressing to more challenging steps), increasing proportions of unimodal trials (auditory and visual) were randomly shuffled into the session ("unimodal trial incrementation training"). Thus, the task could initially be performed by relying on the sounds, but as the proportion of unimodal trials increased by 10% with each successive step, the visual trials had to be solved to reach the nonhandicapped audiovisual DMS task with equal proportions of auditory, visual, and audiovisual trials. Using this training approach, Monkey S gradually progressed through the increasing unimodal trial steps (Figure 23 xi), and performed above chance when tested with the full concurrent audiovisual DMS task (Figure 23 xii). Ironically, although Monkey S was the last to learn the audiovisual DMS task, performance continued to improve and eventually exceeded that of the other subjects (see Figure 24).

Superior audiovisual memory performance

After learning the concurrent audiovisual DMS task, Monkey V completed 151 total sessions, Monkey F completed 223 sessions, and Monkey S completed 109 sessions. Mean accuracy and response latencies are shown for each subject in Figure 24A. Consistent with human studies reporting a bimodal memory advantage, on average, accuracy was highest on audiovisual trials for all subjects (Monkey V: 76.6%; Monkey F: 80.0%; Monkey S: 81.8%). Repeated ANOVA confirmed that accuracy was significantly affected by stimulus modality (Monkey V: $F[2,300] = 345.8$, $p < .001$; Monkey F: $F[2,444] = 148.2$, $p < .001$; Monkey S: $F[2,216] = 99.2$, $p < .001$), with post hoc tests revealing significant differences among all conditions ($p < .05$; Bonferroni correction). Comparing audiovisual accuracy within individual sessions to the unimodal trial type with the highest accuracy revealed that the superior audiovisual accuracy emerged on average against a background of substantial session-to-session variability (Figure 24B), in which unimodal accuracy was in some cases highest.

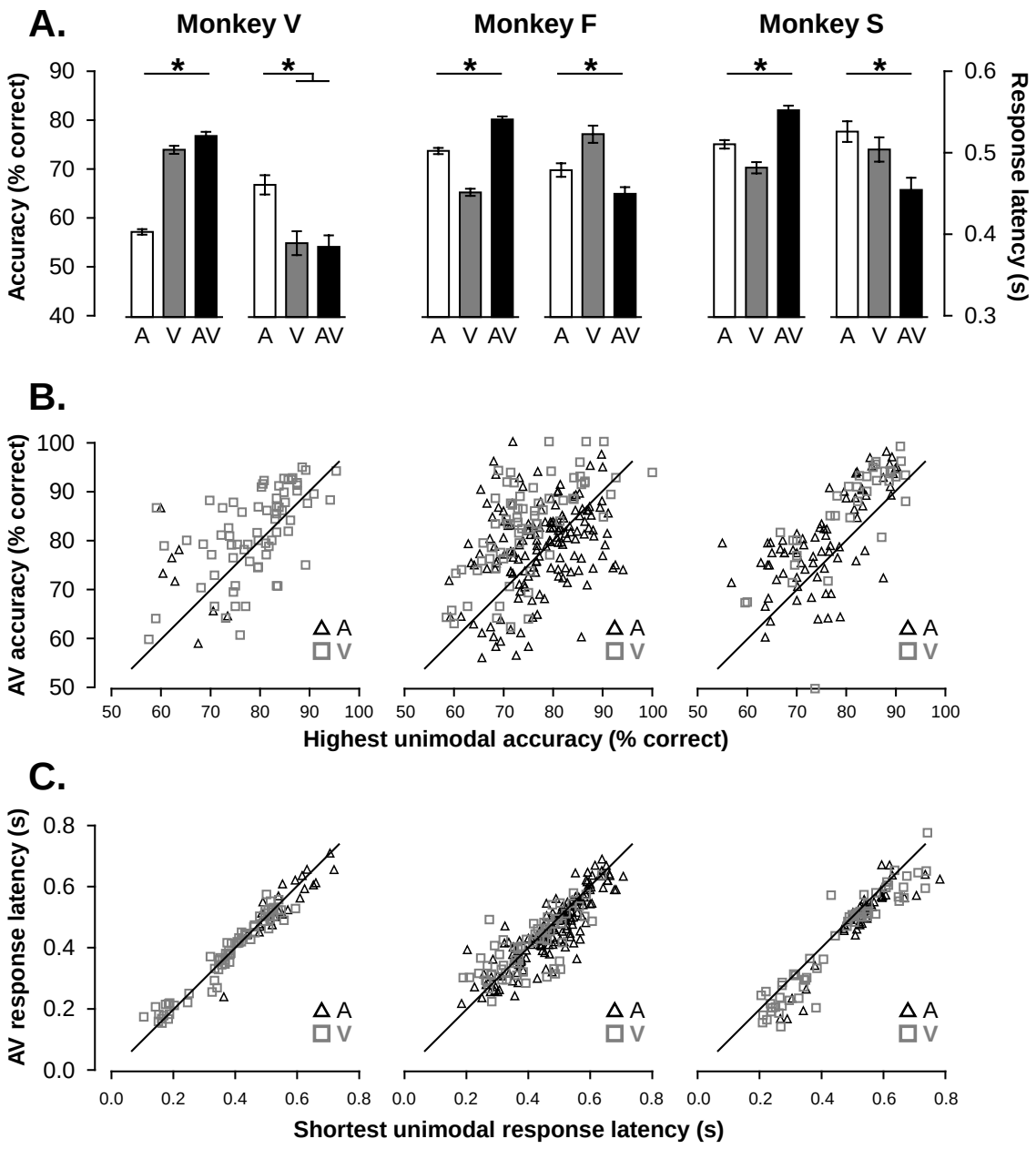

Figure 24. Audiovisual DMS performance. (A) Left graphs: mean accuracy (±SEM) was highest on audiovisual trials for all subjects. For unimodal trials, Monkey V exhibited superior mean visual accuracy, whereas Monkeys F and S exhibited superior mean auditory accuracy. Right graphs: mean latencies (±SEM) for correct match responses were shorter on audiovisual trials than either unimodal trial type for Monkeys F and S. Monkey F responded significantly faster on auditory unimodal trials, whereas Monkey S responded faster on visual unimodal trials. Monkey V exhibited mean audiovisual and visual responses latencies that were both faster than auditory trials, but not significantly different from each other. Pairwise comparisons: *$p < .05$ (Bonferroni correction for multiple comparisons). (B) Session mean accuracy values for audiovisual trials plotted against the unimodal trial type with the highest accuracy. Values above the diagonal line indicate that the highest mean accuracy of the session was observed on audiovisual trials. (C) Session mean response latency values for audiovisual trials plotted against the unimodal trial type with the shortest response latency. Values below the diagonal line indicate that the shortest mean response latencies of the session were observed on audiovisual trials. Note the y-axis scales in (B) and (C) are broadened to accommodate the wider ranges of individual session values.

For unimodal trials, Monkey V exhibited superior visual performance (visual = 73.8%, auditory = 57.2%), consistent with the superior visual DMS performance observed following sound fade-out training (Figure 23 ii). Surprisingly, Monkeys F and S both exhibited superior unimodal performance on auditory trials (Monkey F: visual = 65.2%, auditory = 73.6%; Monkey S: visual = 70.2%, auditory = 74.9%). Although this result is consistent with their longer visual DMS acquisition times, it is a substantial departure from previous studies of auditory and visual memory in monkeys (Wright et al., 1990; Fritz et al., 2005; Scott et al., 2012). We speculate that this outcome may reflect carryover or perseveration effects in these subjects owing to their long training history with auditory DMS. This interpretation is supported by studies reporting longer training times and impaired task switching in DMS and other tasks in older monkeys of similar age to our subjects (Herndon et al., 1997; Moore et al., 2003). These outcomes call for additional studies to work out systematic influences of age, training history, individual differences, and other factors in the emergence of unimodal preferences in memory.

In addition to increased accuracy, subjects were fastest to correctly identify matching test stimuli on audiovisual trials (Figure 24A). Repeated ANOVA confirmed the significant effects of stimulus modality on match response latency (Monkey V: $F[2,300] = 118.6$, $p < .001$; Monkey F: $F[2,444] = 87.2$, $p < .001$; Monkey S: $F[2,216] = 64.2$, $p < .001$). Post hoc tests ($p < .05$; Bonferroni correction) indicated audiovisual response latencies were shorter than either unimodal trial type for Monkeys F and S; Monkey F responded faster on auditory trials, whereas Monkey S responded faster on visual trials. Response latencies for Monkey V were similarly fastest on audiovisual trials, with significant differences observed between audiovisual and auditory trials as well as visual and auditory trials. As was the case for accuracy outcomes, analysis of individual sessions revealed that subjects exhibited shorter audiovisual response

latencies on average, even though this wasn't necessarily the case for every individual session (Figure 24C).

Proactive interference

PI was assessed by identifying the most recent previous trial (*k*) on which the nonmatching test stimulus of the current trial (*n*) had been presented. The resulting *n* – *k* trial separation values were then transformed into recency quartiles (calculated within session), producing a range of "high" to "low" PI conditions. The stimulus modality of both the current trial (*n*) and the PI source trial (*k*) were taken into account, resulting in nine possible intertrial PI combinations (auditory, visual, and audiovisual trials × auditory, visual, and audiovisual PI source trials). For three of these combinations, both the auditory and visual components, or the lack thereof, were identical for trials *n* and *k* (auditory-auditory, visual-visual, audiovisual-audiovisual). Of the six remaining combinations, two shared only the auditory component (auditory-audiovisual, audiovisual-auditory), and two shared only the visual component (visual-audiovisual, audiovisual-visual). The final two combinations shared neither the auditory nor the visual component (auditory-visual, visual-auditory), and here, PI would only be expected if the monkeys had formed crossmodal associations – without explicitly being trained to do so – between the auditory-visual stimulus component pairings. Effects of PI within each of the nine conditions were considered significant if ANOVA comparing accuracy among recency quartiles was significant at the $p < .05$ level.

To facilitate comparison of PI effects across modalities and subjects, nonmatch accuracy values in Figure 25 are presented as percent correct for each recency quartile minus the mean. With several minor exceptions, accuracy outcomes among modalities evaluated separately for

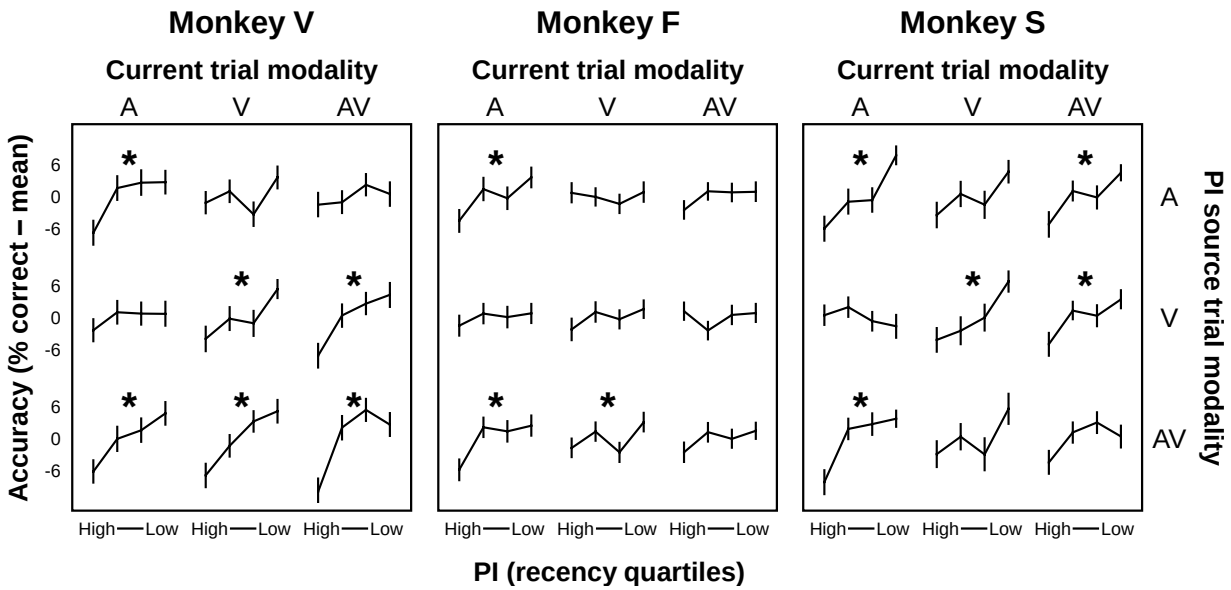

Figure 25. Mean (±SEM) nonmatch accuracy as a function of intertrial proactive interference (PI) arranged by stimulus modality of the current and PI source trials. PI was assessed by identifying the most recent previous trial (*k*) on which the nonmatching test stimulus of the current trial (*n*) had occurred, and trial separation values were transformed into recency quartiles reflecting relatively "high" to "low" PI conditions. Stimulus modalities of the current and PI source trials are arranged along the secondary x- and y-axes, respectively, within panels representing individual subjects (e.g., top-right subplots within each panel represent previous presentations of the nonmatching test sound presented on auditory trial *n* in the context of a previous audiovisual trial *k*). Substantial variation was observed across subjects and modality conditions, however, several outcomes were consistently observed. First, in all cases of significant PI, nonmatch accuracy increased as an inverse function of stimulus repetition recency. Second, PI was never observed in cases where neither the auditory nor visual stimulus was shared between trials (auditory-visual and visual-auditory conditions). Third, PI effects were significant for at least one subject if the current and PI source trials shared at least one common stimulus (all besides the auditory-visual and visual-auditory conditions). Finally, all subjects exhibited significant PI on auditory trials where the PI source trial included the auditory component (auditory-auditory and audiovisual-auditory conditions). *$p < .05$, intertrial PI was significant as assessed by ANOVA comparing nonmatch accuracy among recency quartiles.

match and nonmatch trials tended to reflect the mean overall accuracy values reported above. For Monkey V, nonmatch accuracy was highest on visual trials (67.3%), followed by audiovisual trials (67.3%), and was substantially lower on auditory trials (30.9%; $F[2,444] = 535.7$, $p < .001$; $p < .05$ with Bonferroni correction for all pairwise comparisons); match accuracy was highest on audiovisual trials (85.9%), followed by auditory trials (83.0%), and was lowest for visual trials (76.0%; $F[2,444] = 41.7$, $p < .001$; $p < .05$ with Bonferroni correction for all pairwise comparisons). For Monkey F, nonmatch accuracy was highest on audiovisual trials (75.3%), followed by visual trials (71.1%), and was lowest on auditory trials (67.3%; $F[2,444] = 23.6$, $p < .001$; $p < .05$ with Bonferroni correction for all pairwise comparisons); match accuracy was also highest on audiovisual trials (84.7%), followed by auditory trials (79.8%), and was lowest for visual trials (59.3%; $F[2,444] = 306.3$, $p < .001$; $p < .05$ with Bonferroni correction for all pairwise comparisons). For Monkey S, nonmatch accuracy was highest on audiovisual trials (76.5%), but did not differ significantly between auditory and visual trials (65.6% and 63.8%, respectively; $F[2,216] = 53.8$, $p < .001$; $p < .05$ with Bonferroni correction for post hoc tests comparing audiovisual versus auditory and visual trials); match accuracy was also highest on audiovisual trials (87.3%), followed by auditory trials (84.1%), and was lowest for visual trials (75.9%; $F[2,444] = 50.9$, $p < .001$; $p < .05$ with Bonferroni correction for all pairwise comparisons).

The results of the PI analyses, summarized in Figure 25, revealed considerable variability among subjects and intertrial conditions. Monkey V was most susceptible to PI effects, with significant PI resulting from most, but not all intertrial conditions (auditory-auditory, audiovisual-auditory, visual-visual, audiovisual-visual, visual-audiovisual, and audiovisual-audiovisual). Monkey F was least susceptible to PI, exhibiting significant PI in only three

intertrial conditions (auditory-auditory, audiovisual-auditory, and audiovisual-visual). Monkey S was most similar to Monkey V, with significant PI effects observed in five intertrial conditions (auditory-auditory, audiovisual-auditory, visual-visual, auditory-audiovisual, and visual-audiovisual). Several conclusions are warranted from these outcomes. First, consistent with previous studies, in all instances where significant PI was observed, nonmatch accuracy increased as an inverse function of stimulus repetition recency, i.e., the number of trials separating the nonmatch test of the current trial from the previous trial on which it most recently occurred. Second, at least one common physical stimulus between the current and PI source trials was necessary for an effect: none of the subjects' nonmatch accuracy was significantly affected by recent presentations of the opposite-modality stimulus component on either unimodal trial type (left-middle and top-center subplots), which suggests that our subjects did not form associations among auditory-visual stimulus pairs of significant strength to elicit intertrial PI effects. Third, at least one subject exhibited significant PI effects in all other conditions, i.e., when at least one physical stimulus was shared between the current and PI source trial types (all besides the left-middle and top-center subplots). Finally, the only conditions that consistently resulted in significant PI effects for all subjects were those in which the current trial was auditory (auditory-auditory and audiovisual-auditory), which raises the possibility that unimodal STM in monkeys may be more susceptible to PI in the auditory modality. This was true in our study even though two of three subjects exhibited greater mean overall accuracy for unimodal auditory trials, suggesting that PI-related errors on auditory nonmatch trials were offset by other auditory mnemonic processes such as the higher likelihood of accepting matching auditory sample and test stimuli.

4.4 Discussion

Audiovisual integration underlies a wide range of behaviors in humans and nonhuman primates, including language and communication, recognizing individual conspecifics, and social decision making (Calvert et al., 2004; Kulahci & Ghazanfar, 2013; Stein & Stanford, 2008). Though once believed to be a uniquely human capacity, audiovisual and other forms of multisensory integration have been observed in many other species, and are now considered central to adaptive behavior throughout the animal kingdom (Kulahci & Ghazanfar, 2013; Stein & Stanford, 2008). In recent years, neurobiological studies have revealed similarities in physiological processes and neuroanatomical regions underlying audiovisual integration in humans and nonhuman primates (Romanski, 2012; Stein & Stanford, 2008). Yet, in contrast to the many known behavioral advantages afforded by audiovisual integration in humans (Delogu et al., 2009; Mastroberardino et al., 2008; Shams & Seitz, 2008), relatively little is known about the behavioral consequences of audiovisual integration in other primates (Chandrasekaran et al., 2011; Passingham, 2009). Herein, we build significantly upon the sparse nonhuman primate behavior literature by describing for the first time audiovisual facilitation of memory in a nonhuman primate species (Figure 24). These observations also add to previous reports of qualitative similarities between humans and monkeys in memory processing and other cognitive functions (Matsuzawa, 2001; Zentall & Wasserman, 2012). Taken together, our findings further substantiate monkeys as model species for understanding the biological bases of audiovisual integration and memory in humans (Bigelow & Poremba, 2014; Kulahci & Ghazanfar, 2013; Romanski, 2012; Stein & Stanford, 2008), including the many pathological conditions in which these processes are compromised.

Neuropsychological studies in nonhuman primates and neuroimaging studies in humans have revealed differences in the neural circuitry underlying visual and auditory memory. For instance, visual (but not auditory) memory appears to depend heavily upon the perirhinal and entorhinal cortices, whereas parahippocampal cortex may be especially important for auditory memory (Bigelow & Poremba, 2014; Fritz et al., 2005; Munoz-Lopez et al., 2010). Considering these differences, it is possible that audiovisual facilitation effects reported in the current and previous studies may reflect stronger mnemonic representations resulting from simultaneous activation of both memory pathways. Similarly, cell populations with multisensory integrative properties, such as those in superior temporal sulcus and prefrontal cortex (Brown & Aggleton, 2001; Poremba et al., 2003), may be uniquely recruited during audiovisual trials and thus enhance sensory and mnemonic representations of the memoranda.

An unexpected outcome in the current study was that two subjects (Monkeys F and S) exhibited higher overall mean accuracy on unimodal trials with auditory memoranda, which contrasts sharply with previous studies of auditory and visual memory in nonhuman primates (Fritz et al., 2005; Scott et al., 2012; Wright et al., 1990). This outcome may have resulted from a combination of factors, including the subjects' age (Herndon et al., 1997; Moore et al., 2003) and extensive prior training history with auditory DMS. Nevertheless, one subject (Monkey V) exhibited substantially higher unimodal performance for visual trials, despite virtually identical age and auditory DMS training history. Bearing in mind these outcomes in average performance across the study, it is also worth noting that maximum unimodal performance varied across sessions for all subjects (Figure 24 B and C). Considered together, these outcomes call for a partially revised account of differences in STM performance among sensory modalities in monkeys to accommodate variability among training sessions and account for individual

differences among subjects as well as prior experience and training history.

Significant, negative consequences of PI were observed on trials of all modalities for Monkeys V and S, and for both unimodal trial types for Monkey F (Figure 25). Considering the results of all subjects, however, overall performance outcomes were not consistently tied to differential PI effects among modalities. For instance, nonmatch accuracy on auditory trials was most consistently affected by PI across subjects, even though unimodal accuracy was highest on auditory trials for two of three subjects. Further, audiovisual trials were not consistently less affected by PI than visual trials, even though overall accuracy was highest on audiovisual trials for all subjects. These observations suggest that PI resolution plays a significant role in task performance, but is nevertheless just one of the processes underlying STM.

Chapter 5: The role of the lateral prefrontal cortex in auditory, visual, and audiovisual short-term memory

5.1 Introduction

Important sensory information is often available in the environment for only a brief amount of time, and is typically encountered in a stream of irrelevant information. Moreover, sensory events which hold behavioral significance at one time may quickly become irrelevant in a dynamic contextual environment. Adaptive behavior therefore depends on STM, or the ability to retain internal representations of stimuli that have passed from the sensory environment, while at the same time filtering irrelevant events and continuously monitoring stimulus contingencies. Understanding STM and its underlying neural circuitry has thus been a major focus of research in psychology and neuroscience throughout the past century.

Studies of STM in humans and animals have relied primarily upon the DMS task or one of its variants (D'Amato, 1973; Medin et al., 1976; van Hest & Steckler, 1996). In the prototypical DMS task, a brief sample stimulus is followed by a retention interval, after which the subject must identify the sample from among two test stimuli. In other versions of the task, subjects must report whether or not a single test stimulus was identical to the sample (*same*/*different*). In a simpler version of the task used to study spatial memory, subjects maintain visual fixation at a central point while one of several possible spatial locations are cued. After a retention interval, the subject is rewarded for making a visual saccade or button press that corresponds to the cued location (delayed response).

Studies combining the DMS task with lesions and neurophysiological recordings have

uniformly underscored a critical role for the PFC in integrating and retaining and sensory information in the service of goal directed behavior (Fuster, 2008d; Goldman-Rakic, 1995; Miller & Cohen, 2001). Jacobsen (1935) first reported severe visual STM impairments in monkeys with frontal lobe ablations, and it was subsequently shown that lesions of the lateral division of the PFC were important for the impairment (Meyer et al., 1951; Pribram et al., 1952). Fuster and Alexander (1971) later reported that many cells in the lateral PFC exhibit an elevated firing rate following the sample stimulus which is sustained throughout the retention interval until a behavioral choice is made. Other neurons exhibit a sustained decrease in firing rate during the retention interval, and still others show an intermediate pattern of elevated followed by suppressed firing (Shafi et al., 2007). Many subsequent studies have replicated these findings, further noting that STM performance at the behavioral level is correlated with the level of activity in cells that exhibit delay-related changes in firing rate (Batuev et al., 1979; Fuster, 1973; Watanabe, 1986). Moreover, these delay-related changes in firing rate are not observed in untrained animals (Fuster, 1973). On the basis of these observations, Fuster and others have interpreted the sustained, delay-related changes in firing rate as a neural correlate of stimulus retention (Fuster & Alexander, 1971; Shafi et al., 2007). Cells that exhibit such activity have now been reported in all areas of the lateral PFC, but appear to be most concentrated in and around the principal sulcus (Brodmann's area 46).

Neurons outside of the lateral PFC have also been shown to exhibit delay-related increases in firing rate during visual STM. For example, Miller et al. (1996) compared single-cell activity in the lateral PFC and IT during a visual STM task which included a sample stimulus, followed by a varying number of nonmatching distracters, followed by a matching test stimulus. They reported that, whereas delay-related firing changes in PFC neurons were sustained throughout

each delay period in spite of the distracter stimuli, such changes in IT neurons returned to baseline upon presentation of the first distracter. An additional observation was that the firing rates of neurons in both the lateral PFC and IT were often elevated (but occasionally suppressed) in response to matching versus nonmatching stimuli (Cromer et al., 2011; Miller & Desimone, 1994; Miller et al., 1996; Miller et al., 1991, 1993; Rainer et al., 1999). Because "match enhancement" is observed in IT neurons despite interruption of delay-related activity by distracter stimuli, it has been suggested that these responses may be influenced by the lateral PFC (Miller et al., 1996). These results are consistent with a distributed-network model of STM wherein the lateral PFC takes a central role in biasing representations of behaviorally relevant stimuli in sensory cortical areas (Miller & Cohen, 2001).

Although the vast majority of studies investigating the role of the lateral PFC in STM have focused on the visual modality, several studies have shown that it is also important for auditory STM. Due to the significant challenges associated with training monkeys to perform purely auditory STM tasks (Cohen et al., 2005; Fritz et al., 2005; Scott et al., 2012), some labs have instead trained subjects to associate an auditory sample with a visual test stimulus. Lesions and cooling inactivations of the lateral PFC have been shown to disrupt performance in these tasks (Blum, 1952; Sierra-Paredes & Fuster, 2002), and elevated firing rates have been observed during the retention interval as in purely visual STM tasks (Bodner et al., 1996; Fuster et al., 2000; Joseph & Barone, 1987). It should be noted however, that the delay-related activity observed in these studies may have been related to the visual test stimuli. For example, it has been shown that delay-related activity in PFC neurons can reflect the properties of the anticipated test stimulus as well as the sample stimulus (Rainer et al., 1999). Moreover, Gibson and Maunsell (1997) found that, in IT, an auditory sample stimulus could evoke delay-period

responses when the animals expected to respond to a visual test stimulus. Thus, the auditory-to-visual DMS studies carry the caveat that memory-related neurophysiological activity may reflect activation of cortical areas involved in visual STM.

Several more recent studies have investigated neurophysiological activity in the lateral PFC during STM tasks using purely auditory stimuli for each trial. In some cases, auditory and visual trials have been interleaved throughout the session in order to compare memory-related activity between modalities. Thus, Kikuchi-Yorioka and Sawaguchi (2000) recorded lateral PFC neurons in monkeys performing an oculomotor delayed-response task that presented either a light or a tone in one of four locations. Consistent with previous visuospatial studies, many cells exhibited changes firing rate that were sustained throughout the 3-s delay period. Similar results were obtained in a study by Artchakov et al. (2007) in which subjects identified whether a test tone or light was presented on the same side (left or right) as a sample of the same modality presented several seconds earlier. Of particular note, subpopulations of cells in both of these studies exhibited delay-related changes in activity for only auditory or only visual trials or both, suggesting partial overlap in the circuits underlying auditory and visual STM. Interestingly, however, the proportion of neurons that were active during auditory trials was smaller than the proportion active during visual trials, a finding that is likely related to the inferior performance reported for auditory trials in both studies.

Several additional studies have observed neurophysiological correlates of STM in the lateral PFC for nonspatial auditory stimuli (Lee et al., 2009; Plakke et al., 2013; Russ et al., 2008). Plakke et al. (2013) trained monkeys to perform a Go/No-go version of the *same*/*different* DMS task in which memoranda were selected from a variety of sounds ranging from pure tones to complex vocalizations. Each trial began with a sample sound, which was followed by a 5-s

retention interval, after which a single test sound was presented. A brief wait period (0.5–1 s) followed the test stimulus to aid in distinguishing activity related the sounds from activity related to behavioral responses. The monkeys were trained to press a centrally located button if the sounds were identical (match trials) and otherwise to withhold button presses (nonmatch trials). Comparable to findings in visual DMS tasks, population analyses revealed that matching stimuli elicited enhanced responses compared to nonmatching stimuli. Additional analyses revealed that the enhanced responses were related to the monkeys' perceptual choices, such that erroneous "match" responses on nonmatch trials were similarly associated with an elevated firing rate. During the retention interval, a subpopulation of cells exhibited changes in firing rate, though the proportion was smaller than has been typically reported in visual STM tasks (Shafi et al., 2007), as well as in the auditory-to-visual DMS studies (Bodner et al., 1996; Fuster et al., 2000) and audiospatial studies reviewed above (Artchakov et al., 2007; Kikuchi-Yorioka & Sawaguchi, 2000).

The finding that a larger proportion of delay responsive cells were observed in the auditory-to-visual DMS studies by Bodner et al. (1996) and Fuster et al. (2000) than in the auditory Go/No-go DMS task by Plakke et al. (2013) may have to do with the fact that many of the neurons in the auditory-to-visual DMS studies exhibited acquired auditory-visual associations. Thus, the delay activity observed following the auditory sample stimulus may have reflected anticipation of the visual test stimulus. A second possibility relates to the fact that, because of the Go/No-go contingency used by Plakke et al., the animals may not have anticipated a behavioral action during the retention interval. This is because a "Go" response (button press) was signaled only by a matching test stimulus, which occurred after the retention interval. This view is consistent with a large body of literature describing a role for the PFC in

the anticipation, selection, and execution of motor behavior (reviewed by Fuster, 2008d; Passingham, 1993). The idea that delay-related changes in activity may be partially influenced by anticipation of action is also consistent with the observation that a larger proportion of neurons in the study by Plakke et al. exhibited changes in firing rate during the pre-response wait period on match trials, after which button presses were made. It would also account for the larger proportion of delay responsive neurons reported in the audiospatial studies by Artchakov et al. (2007) and Kikuchi-Yorioka and Sawaguchi (2000), in which a motor response was executed on each trial. Although this speculative interpretation remains to be confirmed by further neurophysiological studies, it receives partial support from human neuroimaging studies reporting activation of the lateral PFC during a memory interval in which an action was selected, but not during a memory interval that did not require action selection (Curtis & D'Esposito, 2003).

On the basis of the data reviewed above, the lateral PFC appears to have a qualitatively similar role in auditory and visual STM. Indeed, neurophysiological recordings coupled with concurrent auditory and visual STM tasks have provided evidence that the networks subserving both forms of STM partially overlap (Artchakov et al., 2007; Kikuchi-Yorioka & Sawaguchi, 2000). However, since the trials in each of these studies were unimodal (auditory or visual), they leave open questions related to STM processing for bimodal, audiovisual events. For instance, a passive exposure study by Sugihara et al. (2006) revealed that many individual neurons in the lateral PFC exhibit audiovisual integrative responses, i.e., enhanced or suppressed responses to audiovisual stimuli relative to their unimodal visual or auditory components. It is not known, however, how the lateral PFC might be involved in retaining representations of audiovisual stimuli in order to guide future behavioral choices. Inasmuch as audiovisual integration is

thought to be important for communication in primates (Romanski & Averbeck, 2009), and has been shown to confer a STM performance advantage in humans (e.g., Mastroberardino et al., 2008) and monkeys (Chapter 4), neurophysiological studies of STM for auditory, visual, and audiovisual stimuli could be of relevance in a wide range of fields.

The current experiment was designed to address several outstanding questions raised in the foregoing review. Individual cells and local cell populations were recorded in the lateral PFC of two nonhuman primate subjects performing a Go/No-go version of the nonspatial DMS task wherein button presses were made following matching but not nonmatching test stimuli. The task included an extended pre-response wait period separating the test stimulus from the response window, similar to the "postponed decision report" paradigm used in previous neurophysiological studies of STM in nonhuman primates (Hernández et al., 2010; Lemus et al., 2007, 2009; Martínez-García et al., 2011). The Go/No-go contingency combined with the pre-response wait period allowed comparison of neurophysiological activity during three types of delay periods: (1) the delay separating sample and test stimuli (sample delay), during which time the behaviorally relevant test stimulus was anticipated but a motor response could not be predicted, (2) the delay separating matching test stimuli and the response window (match delay), during which time a motor response was anticipated, (3) the delay separating nonmatching test stimuli and the response window (nonmatch delay), during which time neither a behaviorally relevant stimulus nor a motor response was anticipated. Differences among all delay types were expected, and in particular, it was hypothesized that a greater proportion of neurons would exhibit elevated changes in firing rate during match delays, which linked a behaviorally relevant sensory event to a prospective action. The memoranda for each trial were presented in either the unimodal auditory or visual or combined audiovisual formats, which allowed us to address two

additional hypotheses. First, it was hypothesized that mnemonic-related neurophysiological activity (delay-related changes in firing rate, and differences between match- nonmatch-evoked responses) would reflect differences in behavioral performance among modalities, e.g., such forms of activity were expected to be most common during audiovisual trials. Second, evidence of physiological audiovisual integration was expected to be observed during the STM task, not only during basic stimulus driven responses, but also during mnemonic-related responses such as delay activity and differences in match/nonmatch firing rates.

5.2 Methods

Subjects and surgery

Two male rhesus macaques (*Macaca mulatta*), each 18 years old at the beginning of the experiment, served as subjects (Monkeys F and S). Prior to recording, both subjects were trained to perform a concurrent audiovisual DMS task (detailed in Chapter 4), and had approximately 10 years of prior experience with auditory DMS (e.g., Ng et al., 2009). Subjects were individually housed with *ad libitum* access to water and controlled feeding schedules, under a 12:12 light:dark cycle. Animals were fed after daily training (Harlan monkey diet plus fruit, vegetables, and treats) and maintained above 85% free feeding weight.

Surgical procedures were performed while the animals were under general anesthesia (isoflurane 1–2%) after being sedated with ketamine (10 mg/kg), with monitoring by veterinary staff for the duration of the surgery. MRI compatible titanium recording chambers (Crist Instruments, Hagerstown, MD) were implanted on the skull using bone screws and dental acrylic

above the principal sulcus of both hemispheres for Monkey F and above the left hemisphere for Monkey S. A titanium head post was affixed centrally over the parietal bones of each animal with bone screws and dental acrylic to enable head restraint during neurophysiological recording. Coordinates for the placement of the chambers and head posts were determined by an atlas of the macaque brain (Paxinos et al., 1999) and a stereotaxic apparatus (David Kopf Instruments, Tujunga, CA). Antibiotics and analgesics were administered under the direction of veterinary staff following surgery. Following surgeries, the animals were scanned with magnetic resonance imaging (MRI: Trio 3T scanner, Siemens Medical Systems, South Iselin, NJ) under ketamine sedation (10 mg/kg) to allow anatomical estimation of the recording locations. The recording chambers were flushed with antiseptics using sterile instruments before and after each experimental session to inhibit infection. All surgeries and procedures conformed to standards provided by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee at the University of Iowa.

Apparatus

Experiments were conducted inside a sound attenuation chamber (Industrial Acoustics Company, Bronx, NY), where subjects sat in a custom primate chair allowing free arm movements. The subjects' head position was fixed during the session by a holder attached to the chair. Images were presented centrally on a monitor (MicroTouch C1500SS, 3M Touch Systems, Methuen, MA) located 32 ± 5 cm from subjects' eye position, and sounds were delivered through a central speaker directly above the monitor. Eye position was continuously monitored at 100 Hz by an infrared primate eye tracking system (ISCAN, Burlington, MA). An acrylic response button (8 × 8 cm) was positioned centrally below the monitor, which was equipped with LED backlights to signal the response window. Rewards were delivered by a pellet

dispenser (Med Associates, Georgia, VT) into a dish immediately below the response button. An overhead LED provided low-level illumination throughout the session, and a second overhead LED served as a trial segregation cue during the intertrial interval (ITI). Custom LabVIEW software (National Instruments, Dallas, TX) controlled all task events.

Short-term memory task

All sessions used the *same*/*different* version of the DMS task (schematic diagram available in Chapter 4, Figure 22). Following a 5-s ITI, a small white fixation point appeared centrally on the screen at eye level, where visual memoranda were subsequently presented (see below). Subjects were required to fixate for 1 s within ~10° of the fixation point to initiate the trial, and maintain fixation within the approximate dimensions of the visual stimuli throughout all stimulus and delay periods until the button light signaled the response window (the same fixation requirements applied to auditory trials even though no images were presented). Trials in which subjects lost fixation were aborted and replaced with a pseudorandomly selected trial of the same trial type. Each trial presented 0.5-s sample and test stimuli separated by a 1.5-s retention interval. Following a 1-s delay after the test stimulus, the response button was illuminated, signaling a 1.5-s response window. The delay separating the test stimulus and response window was included to separate stimulus-evoked and motor-related activity in the neurophysiological recordings. Previous studies suggest that a brief pre-response delay (1–2 s) is not detrimental to performance, but may provide modest facilitation (Lemus et al., 2007). Subjects were trained to press the button following identical test stimuli (match), but to withhold button presses following nonidentical test stimuli (nonmatch). Responses were subject to an asymmetric reinforcement contingency wherein correct match responses were rewarded with a small food pellet and incorrect button presses on nonmatch trials were punished by a 2-s dark

timeout. Responses outside the response window aborted the trial. Memoranda comprised sounds for auditory trials (Figure 22A), images for visual trials (Figure 22B), and sounds and images presented simultaneously for audiovisual trials (Figure 22C). For audiovisual nonmatch trials, both the sound and image of the test stimulus differed from the sample. Within sessions, each trial type (match, nonmatch; auditory, visual, audiovisual) occurred equally often in random order. For occasional sessions in which subjects ceased responding 20 or more trials before the end of the programmed session, the last response was retroactively assigned as the end of the session. Training sessions comprised 120–300 trials, and were retained for analysis only if mean overall accuracy exceeded chance (χ^2 test, $p < 0.05$).

Stimuli

Stimuli serving as memoranda varied from day to day and comprised 12–20 exemplars drawn from a collection of conspecific monkey faces and vocalizations, human faces and vocalizations, heterospecific animal faces and vocalizations, and abstract images and sounds. Each exemplar included image and sound components, presented separately on unimodal trials and together on audiovisual trials as described above. Exemplars were presented equally often as the sample and test (match and nonmatch). Conspecific monkey faces and vocalization stimuli were created from video recordings collected in the primate colony at the University of Iowa (acquired at 30 frames/s; frame size: 1920 × 1080 pixels), and from other labs. Vocalization sounds were extracted from the auditory track of the video recordings. Images that corresponded to the vocalizations were created by selecting one of the frames from the midpoint of the vocalization and cropping the frame around the face. Human and heterospecific vocalization stimuli were created in a similar manner from video recordings obtained in the lab and from open source videos available online. Unlike the face and vocalization pairings, the auditory and visual

components of the synthetic/abstract stimuli were obtained from independent sources. The auditory components comprised complex, artificial sounds generated by electronic synthesizers or downloaded from abstract sound categories (e.g., "science fiction") of sound effects collections available online. The visual components comprised complex, abstract images (e.g., fractals) obtained from open source image collections available online. All image-sound pairings for audiovisual trials were kept constant within and across sessions. Images were normalized for mean luminance (Photoshop, Adobe Systems, San Jose, CA), and presented centrally at eye level in full color at 20 ± 5° viewing angle. Sounds were normalized for root-mean-square (RMS) amplitude (Audition, Adobe Systems, San Jose, CA), and presented centrally at 75 ± 5 dB measured from subjects' ear level.

Neurophysiological recording

For each session, a multielectrode system was used to lower 1–4 insulated tungsten microelectrodes per hemisphere (1–3 MΩ impedance; FHC Inc., Bowdoin, ME) into PFC. Each electrode was held inside a 23-gauge sterile guide cannula, which was positioned with an X-Y grid attached to a micromanipulator, and was lowered into PFC with a computer controlled electrode drive system (NAN Instruments, Nazareth, Israel). Spiking activity was extracted by applying a band-pass filter (0.5–10 kHz) and site-specific amplitude threshold to the raw extracellular signal. The resulting spike waveforms were amplified, digitized, and displayed in real time (OmniPlex, Plexon Inc., Dallas, TX), with spike times saved to hard disk at 40 kHz. Task events such as stimulus presentations and behavioral responses were recorded concurrently with the neurophysiological data. For many recording sites, it was possible to isolate single-unit activity (SUA) from the filtered extracellular signal using a combination of conventional online and offline spike sorting techniques (e.g., principal components analysis, template matching;

OmniPlex and Offline Sorter; Plexon Inc., Dallas, TX). Filtered spiking activity that could not be sorted into SUA was combined into a multiunit activity (MUA) signal on a per electrode basis (Kayser et al., 2008; Perrodin et al., 2014). Data recording was initiated after one or more single units had been isolated prior to beginning the DMS task. Recording sites were not guided by search stimuli or screened for task related activation, resulting in a unit sample that was unbiased with respect to stimulus preference and event related responses. A total of 638 units (294 SUA, 344 MUA) were recorded across 87 sessions (Monkey S, 59 sessions, left hemisphere: 159 SUA, 199 MUA; Monkey F, 28 sessions, left hemisphere: 54 SUA, 51 MUA; right hemisphere: 81 SUA, 94 MUA). Anatomical locations of unit recordings were estimated from the animals' MRIs and stereotaxic surgical coordinates, and are shown superimposed on a generic atlas of the monkey brain in Figure 26. In general, the effects reported below (e.g., sensory evoked responses, delay activity, match nonmatch discrimination effects) were observed across unit types and subjects, and so are evaluated together (Bigelow et al., 2014; Kayser et al., 2008, 2009). The SUA and MUA subsets generally exhibited similar trends in terms of percentages of units responding for one or more modalities and/or task conditions. However, in many (but not all) cases, significant effects were more frequently observed within the MUA subset. For this reason, tables summarizing significant effects by unit type are presented in tandem with graphical summaries of the entire unit population.

Data Analysis

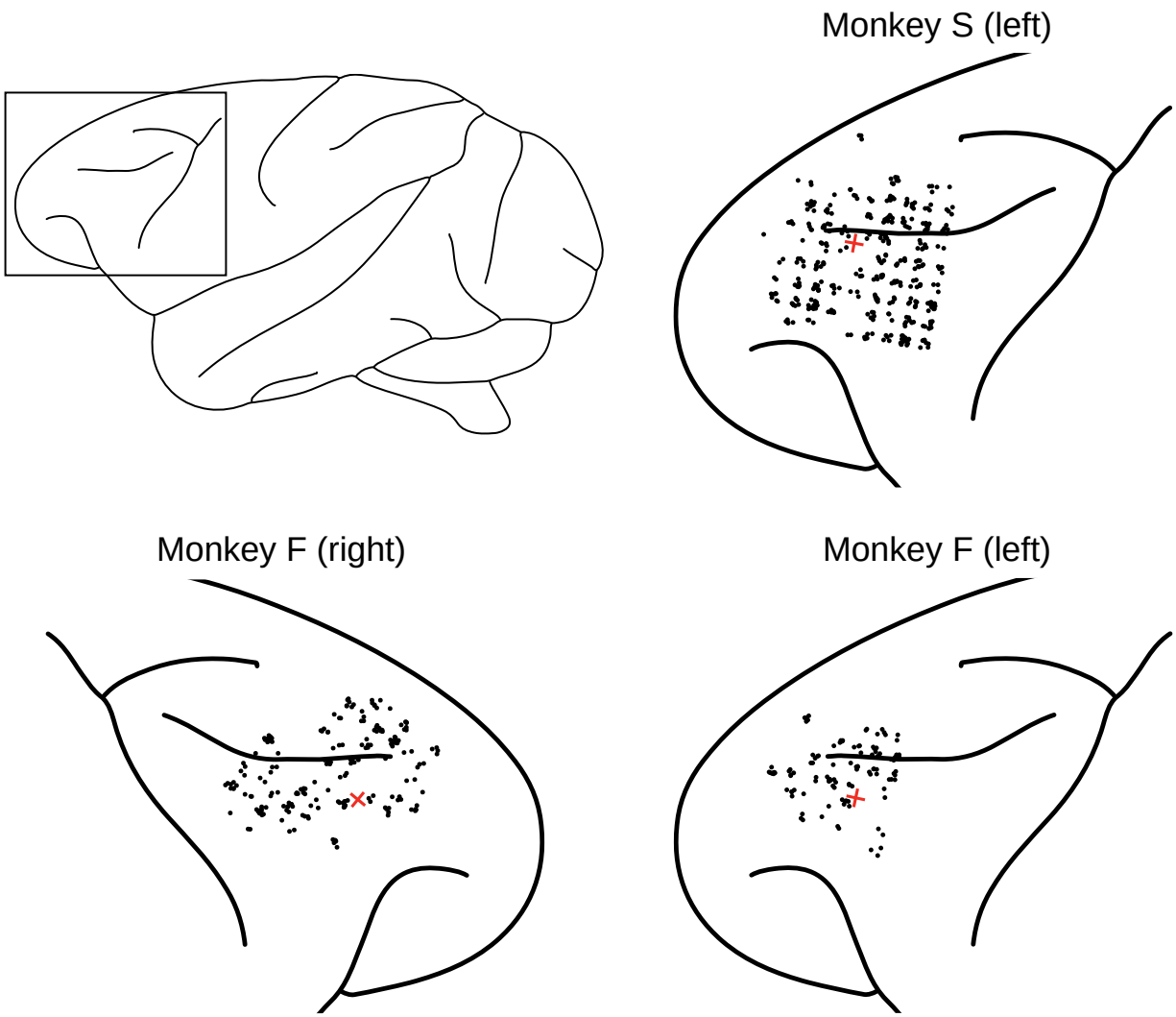

Figure 26. Estimated anatomical locations of all unit recordings for each animal (hemisphere). The recording positions were estimated from the animals' MRIs and stereotaxic surgical coordinates, and are shown superimposed on a generic atlas of the monkey brain (scaled to account for slight variation in subjects' anterior-posterior cerebral dimensions).

The SUA and MUA data were exported to a neurophysiological data analysis program (NeuroExplorer, Nex Technologies, Littleton, MA), wherein spiking activity related to task events such as sample and test stimuli was evaluated using peristimulus time histograms. Unless otherwise noted, raw data were extracted as the mean firing rate for correct trials sampled in 20 ms bins, with data points comprising single trial means for individual unit analyses and session means (collapsed across trials) for population analyses. To compensate for fluctuations in

spontaneous firing rate across units and task conditions, the raw data values were transformed into firing rates minus baseline, defined as the mean firing rate during the 500 ms period before stimulus onset (Perrodin et al., 2011, 2014). In the analyses below, differences in firing rate among conditions were assessed with analysis of variance (ANOVA), except where otherwise specified. For comparisons among relatively broad (e.g., 500 ms), non-overlapping trial segments, differences were tested against a significance threshold of $p < .05$ (two tailed tests, Bonferroni adjustment within modality for multiple comparisons). For other analyses comparing firing rates in brief successive steps (e.g., 20 ms), differences were required to exceed a significance threshold of $p < .01$ (two tailed tests, adjusted for false discovery rate [FDR] against the probability of a streak of $n–k$ consecutive successes in k Bernoulli trials with a success probability of .01, where n corresponds to the minimum number of consecutive significant results required by the analysis and k corresponds to the number of comparisons made for each modality). Cochran's Q tests plus pairwise comparisons were used to evaluate differences among the proportions of units exhibiting significant effects across conditions (two tailed tests evaluating the dichotomous "significant" versus "non-significant" outcomes, $p < .05$, adjusted for multiple comparisons).

Units were classified as exhibiting significant a sensory-evoked response (per modality; SER) if significant tonic (sustained) or phasic (transient) changes in firing rate were detected during one or more of the cue periods (sample, match, or nonmatch; cf. Perrodin et al., 2011, 2014; Scott et al., 2014). All trials (regardless of accuracy) were included in SER analyses, and the sample SER was collapsed across match and nonmatch trials (with Bonferroni adjustment for twofold trial numbers). Tonic responses were defined by significant differences in mean firing rate during the entire 500-ms cue period compared to the 500-ms pre-stimulus baseline period (*p*

< .05, adjusted for multiple comparisons), and phasic responses were defined by significant deviations from baseline for ≥2 consecutive 20-ms bins at any time during the cue period (p < .01, adjusted for FDR as described above). A secondary analysis was performed for only those units exhibiting a significant SER for at least one modality, wherein spiking activity was compared between audiovisual (AV) trials and the unimodal trial type exhibiting the greatest absolute change from baseline during the stimulus period (U_{max}). The analysis was further restricted to only those cues (tonic responses) or portions of the cues (phasic responses) wherein significant SERs were obtained as described above for either modality (AV, U_{max}). Units with significant *trial period* (baseline, stimulus period) × *modality* (AV, U_{max}) interactions were classified as exhibiting significant AV integrative SERs, where interaction effects resulting from either tonic or phasic response analyses were accepted. Following conservative thresholding of phasic SERs (p < .01), interactions were accepted if significant effects were obtained for two or more consecutive 20-ms bins at the p < .05 level (using the same FDR adjustment described above).

Delay-related changes in activity were assessed by ANOVA plus post hoc tests comparing firing rates during the 500-ms pre-sample baseline period to the three successive 500-ms segments of the retention interval separating the sample and test stimuli (p < .05, adjusted for multiple comparisons; Bigelow et al., 2014; Ng et al., 2014; Plakke et al., 2013). The delays separating match/nonmatch test stimuli and the response window were compared with a similar analysis, except that there were only two successive 500-ms delay segments, which were compared to the 500-ms pre-test baseline period. Units exhibiting significant delay activity according to these criteria were further tested for significant AV integrative responses, and as above, only segments of the delays with significant effects for the U_{max} or AV conditions were

considered. Units with significant *trial period* (baseline, delay segment) × *modality* (AV, U_{max}) interactions were thus classified as exhibiting significant delay-related AV integration.

Differences in firing rates evoked by matching and nonmatching test stimuli were compared with ANOVA in a 100-ms sliding window, advancing in 20-ms steps (Apicella et al., 1997; Bigelow et al., 2014; Chandrasekaran & Ghazanfar, 2009; Darbaky et al., 2005; Scott et al., 2014). The sliding window analysis was inclusive of the entire test stimulus period, pre-response delay, and 500-ms pre-test baseline period. Effects were only considered significant in cases where the match-nonmatch (M-NM) difference exceeded a $p < .01$ threshold for two or more consecutive analysis steps (adjusted for FDR as described above). Tests for AV integration effects were conducted, as above, for the subset of units exhibiting significant M-NM differences within the analysis steps where significant differences were obtained for AV and/or U_{max} trials. Following conservative thresholding of the M-NM difference ($p < .01$), interactions were accepted if significant effects were obtained for two or more consecutive analysis steps at the $p < .05$ level (using the same FDR adjustment described above). Under these criteria, significant *trial type* (match, nonmatch) × *modality* (AV, U_{max}) interactions were used to classify units exhibiting M-NM effects that were subject to AV integration.

For display purposes, the averaged firing rates resulting from the sliding window analysis described above are shown for individual unit examples (i.e., boxcar filtered firing rates sampled in 20-ms bins with a 5-bin filter width). A more temporally refined depiction of the population averaged firing rates was obtained by smoothing the spiking data sampled in 1-ms bins with a Gaussian function ($\sigma = 20$ ms).

5.3 Results

Behavior

As reported in Chapter 4, average performance was highest on audiovisual trials, and of the unimodal trial types, average performance was higher for auditory trials, perhaps in part due to carryover from their extensive prior experience with auditory DMS tasks. For the sample of testing sessions paired with the neurophysiological recordings described above, mean accuracy was 80.4% for auditory trials, 73.2% for visual trials, and 85.4% for audiovisual trials, and mean response latencies (correct match trials) were 568 ms for auditory trials, 598 ms for visual trials, and 534 ms for audiovisual trials. Significant effects of modality were confirmed by ANOVA for accuracy ($F[2,172] = 101.3$, $p < .001$) and response latency ($F[2,172] = 51.8$, $p < .001$), with significant differences observed among all pairwise comparisons ($p < 0.05$, adjusted for multiple pairwise comparisons).

Sensory-evoked responses

Example units depicting significant SERs are shown in Figure 27, and a summary of SERs for each modality and cue type (including intersections of SERs among conditions) is depicted graphically by area-proportional Euler diagrams in Figure 28 (Micallef & Rodgers, 2014), along with overall percentages of units with significant responses per condition presented below each diagram. Table 1 presents a similar summary of SERs broken down by unit type (SUA, MUA, Total), which includes the same response intersections among modalities and cue types, as well as overall percentages of each unit type with significant responses per condition presented in the rightmost column. An additional summary of SERs collapsed across cue type is presented

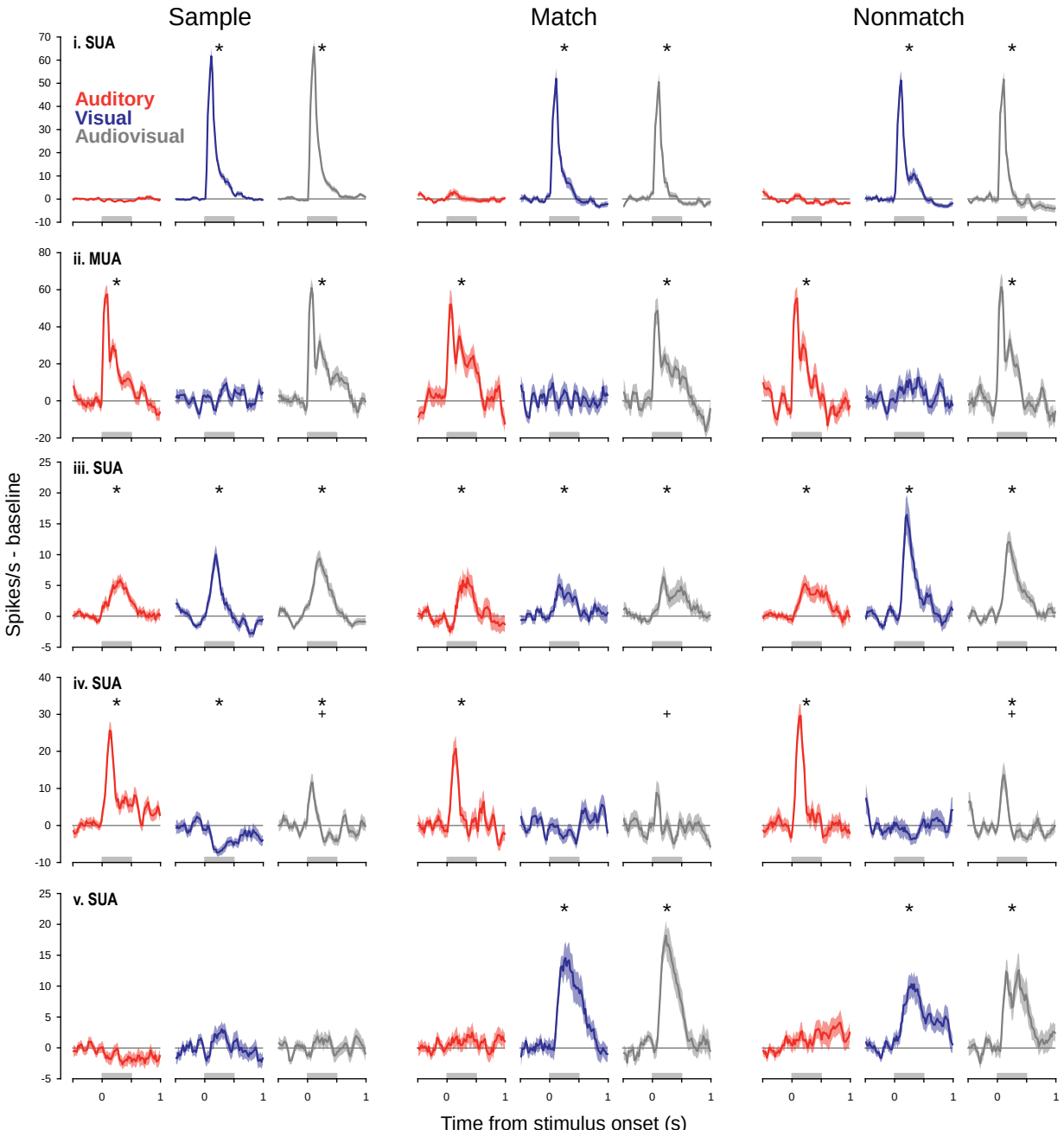

Figure 27. Example units depicting significant sensory-evoked responses (SERs) by cue type and sensory modality. For some units, significant SERs were observed for any cue containing a visual (unit i) or auditory component (unit ii). Other units responded regardless of stimulus modality or cue type (e.g., unit iii). In some cases (e.g., unit iv), a significant interaction was obtained between *trial period* (baseline, cue period) and *modality* (AV, U_{max}). Other units exhibited SERs only for specific combinations of cue type and modality (e.g., unit v responded to match or nonmatching cues with a visual component). Mean (±SEM) firing rates are depicted by dark central lines (plus lighter shaded bands). Stimulus periods are represented by gray bars abutting the abscissae. *p < .05, baseline versus cue period; +p < .05, *trial period* × *modality* interaction.

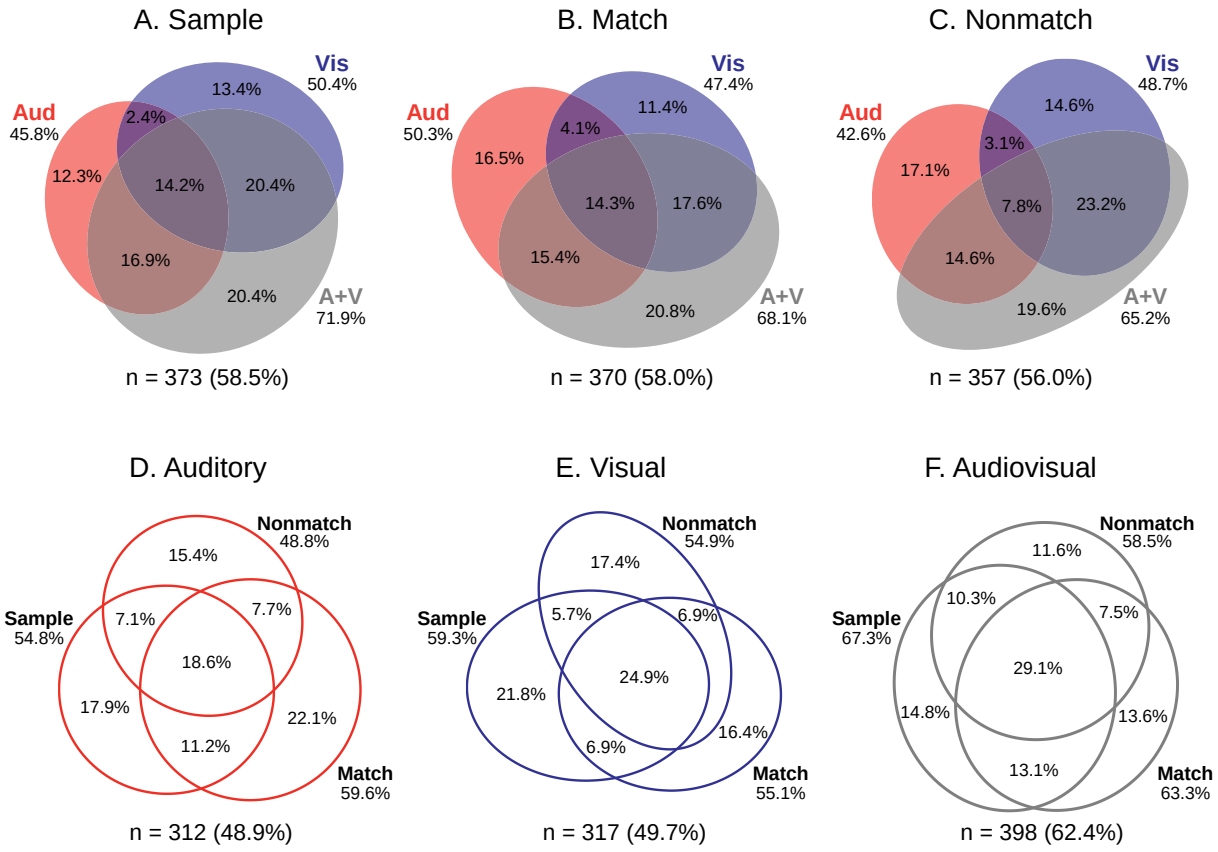

Figure 28. Euler diagram summaries of significant sensory-evoked responses for each modality and cue type represented by ellipses (and overlap) that are area-proportional to the number of units with significant effects for that condition (and overlap with other conditions). (A–C) SERs and overlap among modalities for each cue type (D–F) SERs and overlap among cue types for each modality. Percentages of the subset of units with significant effects for each condition (given below each diagram) are displayed within each fraction of the diagrams. In addition, percentages of responses summed per modality or cue (regardless of overlap with other modalities or cues) are displayed near the outside their respective ellipses. For instance, (A) depicts that, of all units in the sample, 58.5% exhibited significant a sensory-evoked response during the sample for one or more modalities, and of these, 71.9% responded to audiovisual stimuli. Dividing the units with audiovisual responses by modality overlap reveals that 14.2% also responded to both auditory and visual cues, 20.4% responded to visual but not auditory cues, 16.9% responded to auditory but not visual cues, and 20.4% responded exclusively to audiovisual cues. Note that the Euler diagrams are area proportional within but not among conditions (e.g., the total area of A is not exactly proportional to the total area of D). Substantial portions of units responded to more than one stimulus modality (A–C) and cue type (D–F). Approximately equal proportions of units exhibited significant SERs for at least one modality during sample, match, and nonmatch cues (A–C). A larger portion of units exhibited SERs for audiovisual compared to unimodal auditory or visual stimuli (D–F).

Table 1. Sensory-evoked responses by unit type, modality[a], cue type[b], and intersection[c]

	Modality intersections by cue type							Units w/ sig. effect[d]
	Sample							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	13.1%	17.6%	17.6%	3.9%	15.0%	22.2%	10.5%	52.0%
MUA	11.8%	10.5%	22.3%	1.4%	18.2%	19.1%	16.8%	64.0%
Total	12.3%	13.4%	20.4%	2.4%	16.9%	20.4%	14.2%	58.5%
	Match							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	18.5%	16.4%	18.5%	2.7%	15.8%	16.4%	11.6%	49.7%
MUA	15.2%	8.0%	22.3%	4.9%	15.2%	18.3%	16.1%	65.1%
Total	16.5%	11.4%	20.8%	4.1%	15.4%	17.6%	14.3%	58.0%
	Nonmatch							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	16.0%	18.1%	22.2%	2.1%	12.5%	23.6%	5.6%	49.0%
MUA	17.8%	12.2%	17.8%	3.8%	16.0%	23.0%	9.4%	61.9%
Total	17.1%	14.6%	19.6%	3.1%	14.6%	23.2%	7.8%	56.0%

	Cue type intersections by modality							Units w/ sig. effect[d]
	Auditory							
Cue	S	M	N	S ∩ M	S ∩ N	M ∩ N	S ∩ M ∩ N	
SUA	16.5%	26.4%	15.7%	14.0%	9.1%	4.1%	14.0%	41.2%
MUA	18.8%	19.4%	15.2%	9.4%	5.8%	9.9%	21.5%	55.5%
Total	17.9%	22.1%	15.4%	11.2%	7.1%	7.7%	18.6%	48.9%
	Visual							
Cue	S	M	N	S ∩ M	S ∩ N	M ∩ N	S ∩ M ∩ N	
SUA	29.3%	12.8%	14.3%	4.5%	4.5%	10.5%	24.1%	45.2%
MUA	16.3%	19.0%	19.6%	8.7%	6.5%	4.3%	25.5%	53.5%
Total	21.8%	16.4%	17.4%	6.9%	5.7%	6.9%	24.9%	49.7%
	Audiovisual							
Cue	S	M	N	S ∩ M	S ∩ N	M ∩ N	S ∩ M ∩ N	
SUA	15.8%	15.8%	17.6%	12.7%	11.5%	6.1%	20.6%	56.1%
MUA	14.2%	12.0%	7.3%	13.3%	9.4%	8.6%	35.2%	67.7%
Total	14.8%	13.6%	11.6%	13.1%	10.3%	7.5%	29.1%	62.4%

[a]*A = Auditory; V = Visual; AV = Audiovisual*
[b]*S = Sample; M = Match; N = Nonmatch*
[c]*Intersection percentages based on subsets of units with significant effects reported in right column*
[d]*Percentages of units with significant effects based on SUA = 294, MUA = 344, and Total = 638 units*

in Figure 34A (separated by unit type in Table 3), juxtaposed among similar summaries of significant delay-related activity (Figure 34B) and M-NM effects (Figure 34C) which are described in detail below. Collapsing across cue types revealed that the majority of units in our sample exhibited a significant SER during one or more cues for at least one modality (82.6%, Figure 34A). Separating these responses by cue type revealed that, at the population level, approximately equal proportions of units exhibited significant SERs for at least one modality in the sample (58.5%), match (58.0%), and nonmatch (56.0%) cue positions (Figure 28 A–C). Nevertheless, as can be seen in individual unit examples (Figure 27), and summaries of overlapping SERs for each cue type (Figure 28 D–F), most units did not respond indiscriminately across cue types, but instead responded to just one or two of the cue types.

Significant AV integration effects within SERs were evident in our unit population from several results. First, as seen in Figure 34A, a relatively large proportion of units (25.6%) with SERs responded to auditory, visual, and audiovisual stimuli; indeed, this unit subset was larger than any other subset responding exclusively to any one modality or combination of two modalities. In general, these three-way overlapping responses were observed for all cue types (Figure 28 A–C), but were somewhat less common for nonmatch cues. Second, an additional 5.1% of units responded to both auditory and visual cues, and 11.4% responded exclusively to audiovisual cues (Figure 34A). Third, as seen in Figure 28 D–F, the portion of all units responding to audiovisual cues (62.4%, regardless of overlapping responses to other modalities or cues) was significantly larger than the portion of units responding to unimodal auditory (48.9%) and visual cues (49.7%; $p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). Finally, significant AV interaction effects were detected in 11.4% of units with SERs, defined by an interaction of *trial period* (baseline, cue) and *modality* (AV, U_{max}).

Delay activity

Examples of individual units with changes in firing rate during the classically defined retention interval (i.e., the sample delay) are provided in Figure 29. Euler diagrams summarizing percentages of units with significant effects by modality and delay type are available in Figure 30 (separated by unit type in Table 2), and an additional summary collapsed across delay types is provided in Figure 34B (separated by unit type in Table 3). A more detailed population summary of delay activity is depicted in Figure 31, which includes population mean firing rates (Figure 31A) and significant increases and decreases in activity during individual 500-ms segments of each delay type (Figure 31B). In total, 228 units (35.7% of the unit population) exhibited significant changes in firing rate for one or more modality at one or more segments of the sample delay. Sustained changes in firing rate spanning the entire sample delay for one or more modalities were observed in a relatively small portion of these units (11.4%), with the majority responding during one (67.5%) or two (21.1%) segments of the delay. In general, excitatory effects were more frequent during the early delay period and inhibitory effects were more common near the latter portion of the delay nearest to test stimulus onset (Figure 31B). Indeed, suppression effects during the last segment of the sample delay overcame excitatory effects to the extent that population mean firing rates fell significantly below baseline for all modalities (Figure 31A, $p < .05$, ANOVA, adjusted for multiple pairwise comparisons). Suppressed firing rates during the last 500-ms segment of the sample delay was most pronounced for auditory trials, confirmed in both population mean firing rates (Figure 31A; firing rates were significantly lower on auditory trials compared to visual or audiovisual trials, $p < .05$, ANOVA, adjusted for multiple pairwise comparisons) as well as the percentage of units with significant decreases in firing rate (Figure 31B; $p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons).

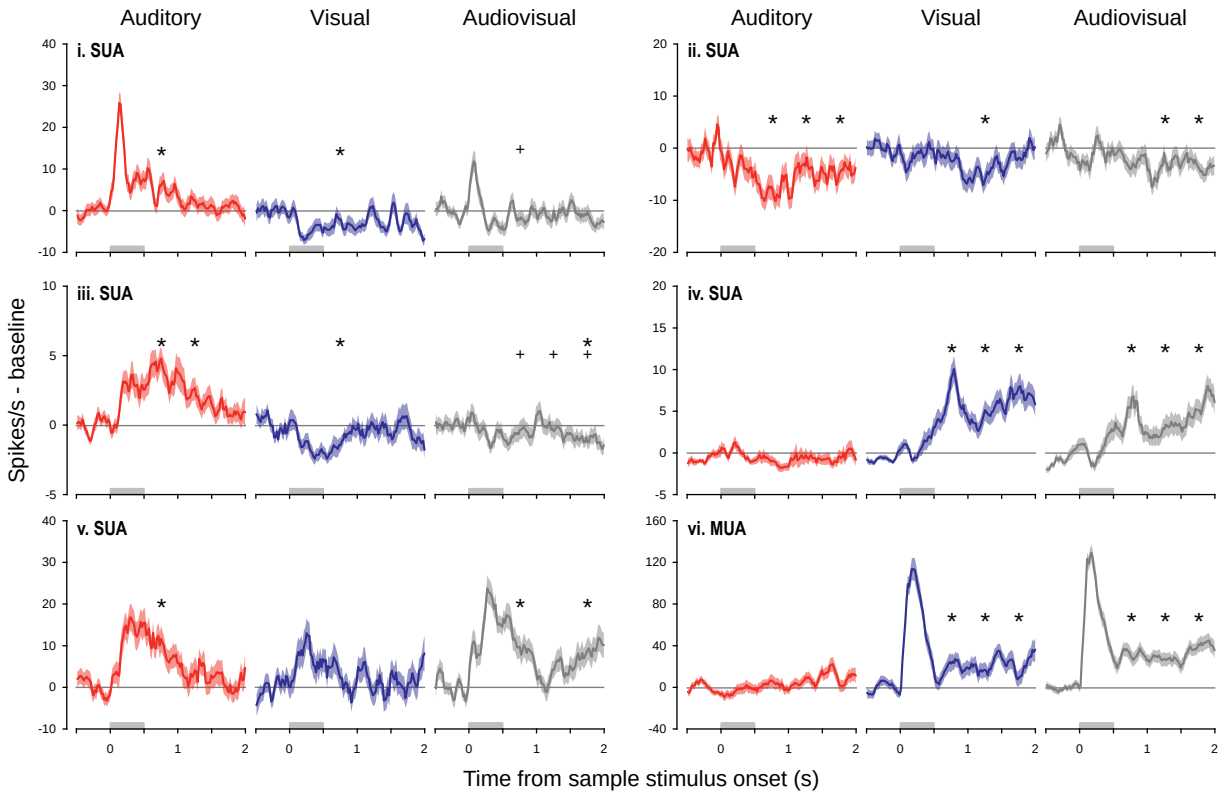

Figure 29. Example units exhibiting significant changes in firing rate for one or more segments of sample delay. Delay activity exceeded baseline for some units (i, iii, iv, v, vi), fell below baseline for others (i, ii, iii), and exhibited combinations of increases and decreases in firing rate for others (i, iii). In some cases (ii, iv, vi), delay-related changes in activity were sustained for the duration of the retention interval, but for the majority, such changes were transient (i, ii, iii, v). In most cases, delay effects were modality dependent (all units shown), and in a subset of these units (i, iii), significant interactions were obtained between *trial period* (baseline, delay segment) and *modality* (AV, U_{max}). Mean (±SEM) firing rates are depicted by dark central lines (plus lighter shaded bands). Stimulus periods are represented by gray bars abutting the abscissae. *$p < .05$, baseline versus cue period; +$p < .05$, *trial period* × *modality* interaction.

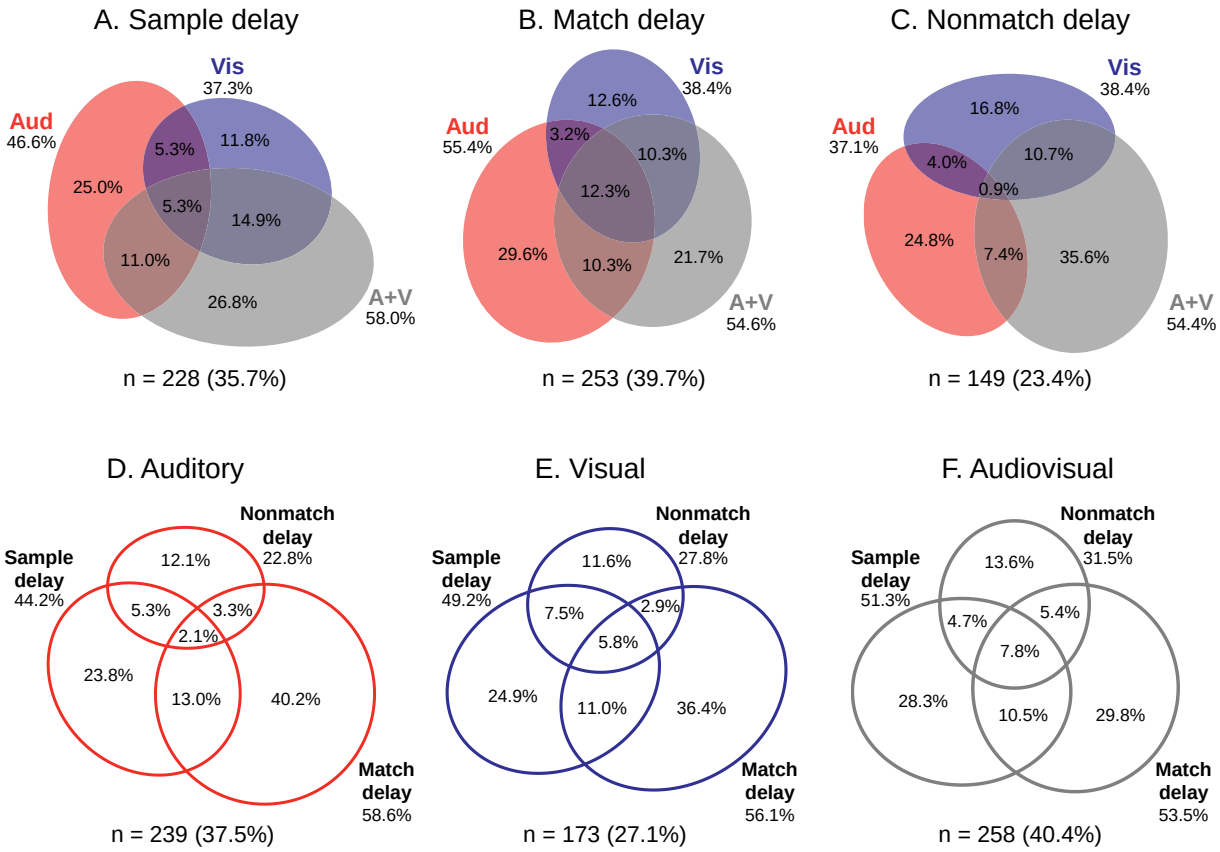

Figure 30. Euler diagram summaries of significant delay-related changes in firing rate for each modality and delay type, using the same conventions as Figure 27. Fewer units exhibited significant responses during the nonmatch delay (C) compared to either the match (B) or sample delays (A). The proportion of units with match delay activity (B) was slightly higher than that with sample delay activity (A), but this difference did not reach significance (see text for details). Significant responses were most common for audiovisual trials during sample and nonmatch delays, whereas during match delays, responses were more likely on auditory and audiovisual trials compared to visual trials. Many units exhibited delay activity for more than one stimulus modality (A–C), and during more than one delay period (D–F), though both forms of overlap among conditions were observed less frequently than in similar SER analyses (see Figure 27).

Table 2. Delay activity by unit type, modality[a], delay type[b], and intersection[c]

	Modality intersections by delay type							Units w/ sig. effect[d]
	Sample delay							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	18.6%	15.7%	23.5%	6.9%	9.8%	19.6%	5.9%	34.7%
MUA	30.2%	8.7%	29.4%	4.0%	11.9%	11.1%	4.8%	36.6%
Total	25.0%	11.8%	26.8%	5.3%	11.0%	14.9%	5.3%	35.7%
	Match delay							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	31.6%	13.7%	16.8%	1.1%	9.5%	14.7%	12.6%	32.3%
MUA	28.8%	12.0%	24.7%	4.4%	10.8%	7.6%	12.0%	45.9%
Total	29.6%	12.6%	21.7%	3.2%	10.3%	10.3%	12.3%	39.7%
	Nonmatch delay							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	14.3%	21.4%	41.4%	4.3%	7.1%	11.4%	0.0%	23.8%
MUA	34.2%	12.7%	30.4%	3.8%	7.6%	10.1%	1.3%	23.0%
Total	24.8%	16.8%	35.6%	4.0%	7.4%	10.7%	0.9%	23.4%

	Delay type intersections by modality							Units w/ sig. effect[d]
	Auditory							
Delay	S	M	N	S ∩ M	S ∩ N	M ∩ N	S ∩ M ∩ N	
SUA	30.5%	41.1%	11.6%	9.5%	3.2%	3.2%	1.1%	32.3%
MUA	19.4%	39.6%	12.5%	15.3%	6.9%	3.5%	2.8%	41.9%
Total	23.8%	40.2%	12.1%	13.0%	5.3%	3.3%	2.1%	37.5%
	Visual							
Delay	S	M	N	S ∩ M	S ∩ N	M ∩ N	S ∩ M ∩ N	
SUA	32.1%	23.8%	13.1%	13.1%	7.1%	4.8%	6.0%	28.6%
MUA	18.0%	48.3%	10.1%	9.0%	7.9%	1.1%	5.6%	25.9%
Total	24.9%	36.4%	11.6%	11.0%	7.5%	2.9%	5.8%	27.1%
	Audiovisual							
Delay	S	M	N	S ∩ M	S ∩ N	M ∩ N	S ∩ M ∩ N	
SUA	29.1%	21.8%	17.3%	10.9%	7.3%	6.4%	7.3%	37.4%
MUA	27.7%	35.8%	10.8%	10.1%	2.7%	4.7%	8.1%	43.0%
Total	28.3%	29.8%	13.6%	10.5%	4.7%	5.4%	7.8%	40.4%

[a] *A = Auditory; V = Visual; AV = Audiovisual*
[b] *S = Sample delay; M = Match delay; N = Nonmatch delay*
[c] *Intersection percentages based on subsets of units with significant effects reported in right column*
[d] *Percentages of units with significant effects based on SUA = 294, MUA = 344, and Total = 638 units*

Similar to outcomes in the SER analyses, AV integration during the sample delay was evident from units exhibiting significant effects for all three modalities (5.3% of units with sample delay effects), or for auditory and visual trials (5.3%), or exclusively audiovisual trials (26.8%). As with SERs, a significantly larger proportion of units exhibited sample delay effects on audiovisual trials (58.0%, regardless of overlapping responses with other modalities or delays; Figure 30B) compared to visual trials (37.3%, $p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). A similar trend was observed between sample delay effects on audiovisual and auditory trials (58.0% vs. 46.6%), but the difference was non-significant after adjusting for multiple comparisons ($p = .09$). In addition to these observations, 8.3% of units with significant sample delay effects also exhibited a significant interaction between *trial period* (baseline, delay segment) and *modality* (AV, U_{max}). Taken together, these findings substantially extend the evidence for audiovisual integrative functions in PFC by demonstrating such processes well after the end of the stimulus (0.5–1.5 s post-stimulus offset).

Extending the definition of "delay activity" to encompass not only the classically defined retention interval separating the sample and test cues, but also the "test delay" separating the test stimulus and response window, revealed that 253 units (39.7% of the total) exhibited significant delay responses for one or more modalities during at least one delay segment following matching test stimuli, and 149 units (23.4% of the total) exhibited significant delay activity following matching test stimuli. Significantly fewer units exhibited significant nonmatch delay activity compared to sample or match delay activity ($p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons), whereas the difference between the number of units with significant sample (35.7%) and match (39.7%) delay responses did not reach significance ($p = .09$).

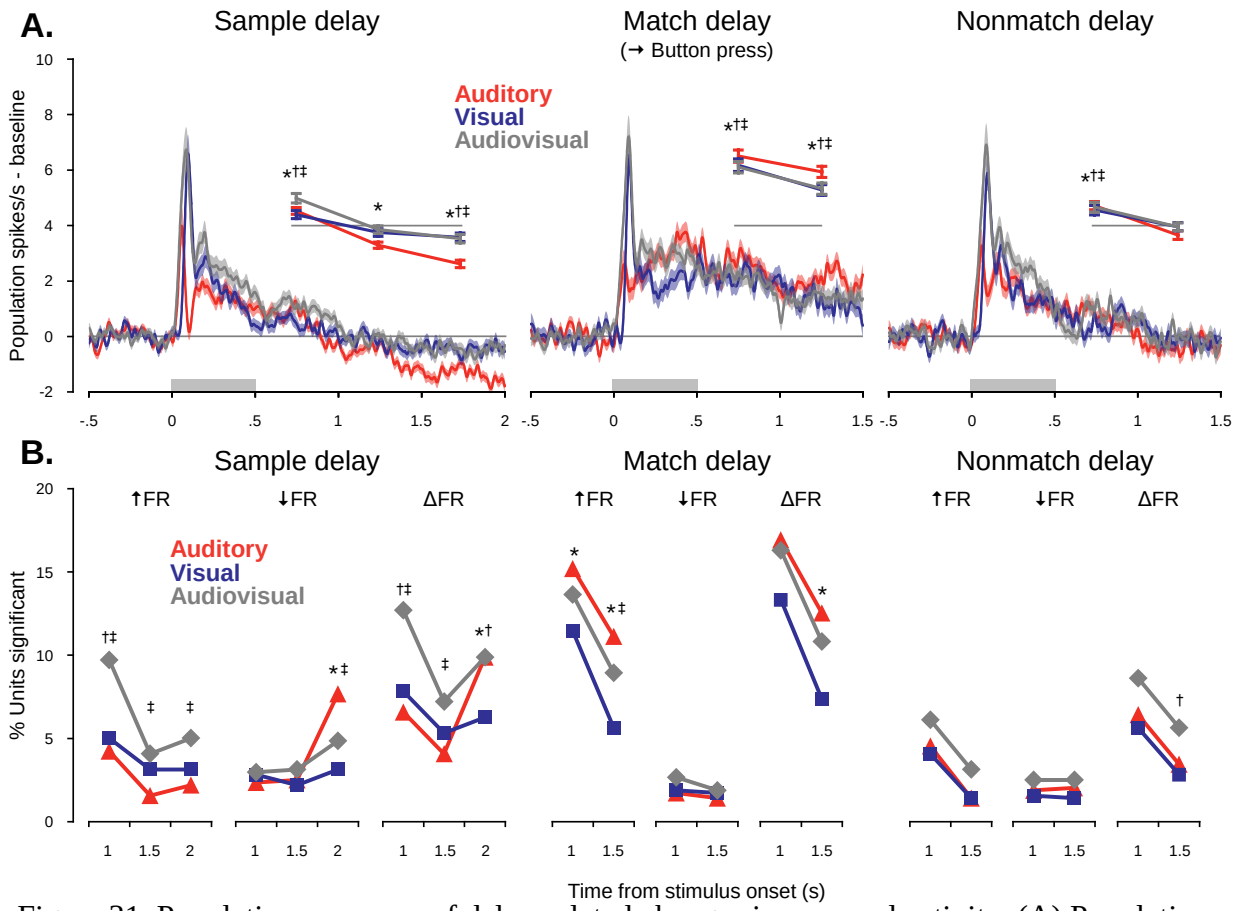

Figure 31. Population summary of delay-related changes in neuronal activity. (A) Population mean firing rates (spikes/s minus baseline) for delay periods separating the sample and test (left panel), the test and response window for match trials (middle panel), and the test and response window for nonmatch trials (right panel). The insets in each panel depict mean (±SEM) firing rates sampled within successive, non-overlapping 500-ms delay periods following stimulus offset (the X and Y scales are the same as those used for the main panels, and the dotted line represents mean pre-stimulus baseline activity). At the population level, sample delays for all modalities were associated with increased firing rates following stimulus offset which then diminished for subsequent delay segments, ultimately falling below baseline before test stimulus onset (firing rates were significantly lower on auditory trials compared to visual or audiovisual trials during the last two delay segments). By contrast, match delays were associated with a sustained increase in firing rate (firing rates were significantly higher on auditory trials compared to visual or audiovisual trials during the last delay segment). For nonmatch delays, firing rates were initially elevated, but returned to baseline values prior to the response window. Mean (±SEM) firing rates are depicted by dark central lines (plus lighter shaded bands). Stimulus periods are represented by gray bars abutting the abscissae. Post hoc comparisons for insets ($p < .05$): *Auditory ≠ baseline, †Visual ≠ baseline, ‡Audiovisual ≠ baseline. (B) In general, the numbers of units exhibiting significant increases and decreases in firing were reflected in the population averages (see text for details). Excitatory and inhibitory effects are designated by ↑FR and ↓FR, respectively, and the sums of these effects are indicated by ΔFR. Pairwise comparisons for insets ($p < .05$): *Auditory ≠ Visual, †Visual ≠ Audiovisual, ‡Audiovisual ≠ Auditory.

Accounting for overlap in units with changes in firing during more than one delay period, inclusion of all three delay types increased the total subset of units with significant "delay activity", as broadly defined, to 61.3% of the population (Figure 34B). As seen in Figure 31B, significant differences were observed in the percentages of units exhibiting changes in firing rate during each delay segment, and moreover, the proportions of excitatory and inhibitory effects were asymmetric among delays. As noted above, population mean firing rates (Figure 31A) revealed diminishing firing rates during the sample delay. By contrast, the delay separating matching test stimuli and the response window was associated with sustained increases in firing rate: all modalities exceeded pre-test baseline period for both match delay segments ($p < .05$, ANOVA, adjusted for multiple pairwise comparisons). On the other hand, mean firing rates for all modalities showed mild but significant elevation following nonmatch test offset ($p < .05$), but then returned to baseline during the final delay segment prior to the response window ($p > .05$, ANOVA, adjusted for multiple pairwise comparisons). In general, the proportions of units exhibiting significant delay activity during each segment of the delay periods (Figure 31B) corresponded with the trends observed in the population mean firing rates (Figure 31A). For match delay periods, the elevated population firing rates reflected the fact that the overwhelming majority of units exhibited excitatory effects (83.0% of units with match delay effects). Here, it is also worth noting that elevated firing was observed more frequently for auditory trials (significantly different from visual trials during the first delay segment, and from visual and audiovisual trials in the second segment; $p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons).

As was the case for the sample delay, units with match and nonmatch delay-related changes in firing rate exhibited audiovisual integrative processes. Of match delay units, 21.7%

responded exclusively during audiovisual trials, 3.2% responded during both auditory and visual trials, and an additional 12.3% responded regardless of modality. In contrast to the sample delay, match delay effects were equally likely for auditory and audiovisual trials, and significantly less likely on visual trials ($p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). Like sample delay units, significant interactions between *trial period* (baseline, delay segment) and *modality* (AV, U_{max}) were observed for 8.3% of units exhibiting match delay effects. Overlapping modality responses and integration effects were also observed during nonmatch delays, though a smaller percentage of units exhibited such effects: 35.6% responded during audiovisual trials only, 4.0% responded for both auditory and visual trials, and just one multiunit (0.7%) responded for all three modalities. Similar to the sample delay, nonmatch delay responses were more likely on audiovisual trials compared to either unimodal trial type ($p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). *Trial period* (baseline, delay segment) × *modality* (AV, U_{max}) interactions were significant for 8.1% nonmatch delay units. For economy of presentation, examples of individual units exhibiting match and nonmatch delay activity are not presented separately, but can be seen in the unit examples depicting M-NM discrimination effects in Figure 32 (all but unit vi exhibited significant match delay effects for one or more modality, and unit viii exhibited a significant interaction; all but units iv, vi, and viii exhibited significant nonmatch delay effects for one or more modality).

Match versus nonmatch discrimination

A sizable subset of our unit population (56.1%) exhibited significant differences in firing rate evoked by matching versus nonmatching test stimuli for at least one modality, either during the stimulus period itself or in the test delay period preceding the response window. Percentages of units with significant M-NM differences are summarized per modality with intersections in

Figure 34C (separated by unit type in Table 3), with example units depicted in Figure 32 and a detailed population summary in Figure 33, which includes population mean firing rates (Figure 33A) and percentages of units with significant M-NM effects plotted over time (Figure 33B). Most units exhibited higher firing rates on match trials ("match enhancement", e.g., Figure 32, units i–vii), but in some cases, nonmatch firing rates were greater ("match suppression", e.g., Figure 32, unit viii). The ratio of units exhibiting match enhancement and suppression effects was found to be significant ($p < .05$, Cochran's Q test), with the largest percentage of enhanced responses observed on auditory trials (70.8%), followed by visual trials (66.0%), and audiovisual trials (59.9%). This outcome is reflected in the population mean firing rates (Figure 33A), wherein a greater difference in raw match and nonmatch firing rates was generally observed on auditory trials (M-NM differences were significantly greater on auditory trials compared to visual or audiovisual trials during the test stimulus period and the second 500-ms test delay period prior to the response window, $p < .05$, ANOVA, adjusted for multiple pairwise comparisons). Indeed, considering the population mean firing rates in isolation might suggest that M-NM discrimination effects were most common for auditory trials, even though overall behavioral performance was highest on audiovisual trials. A closer inspection of both match enhancement and suppression effects reveals that this was not the case. As seen in the percentages of units exhibiting significant M-NM effects during successive analysis steps within the trial period (Figure 33B), enhancement or suppression effects were generally most common for audiovisual trials (confirmed statistically by ANOVA, adjusted for multiple pairwise comparisons within each 500-ms trial segment shown in the insets of Figure 33B), with the exception of more frequent enhancement effects during the test stimulus period on auditory trials. Considering all M-NM discrimination effects for all trial segments together (enhancement

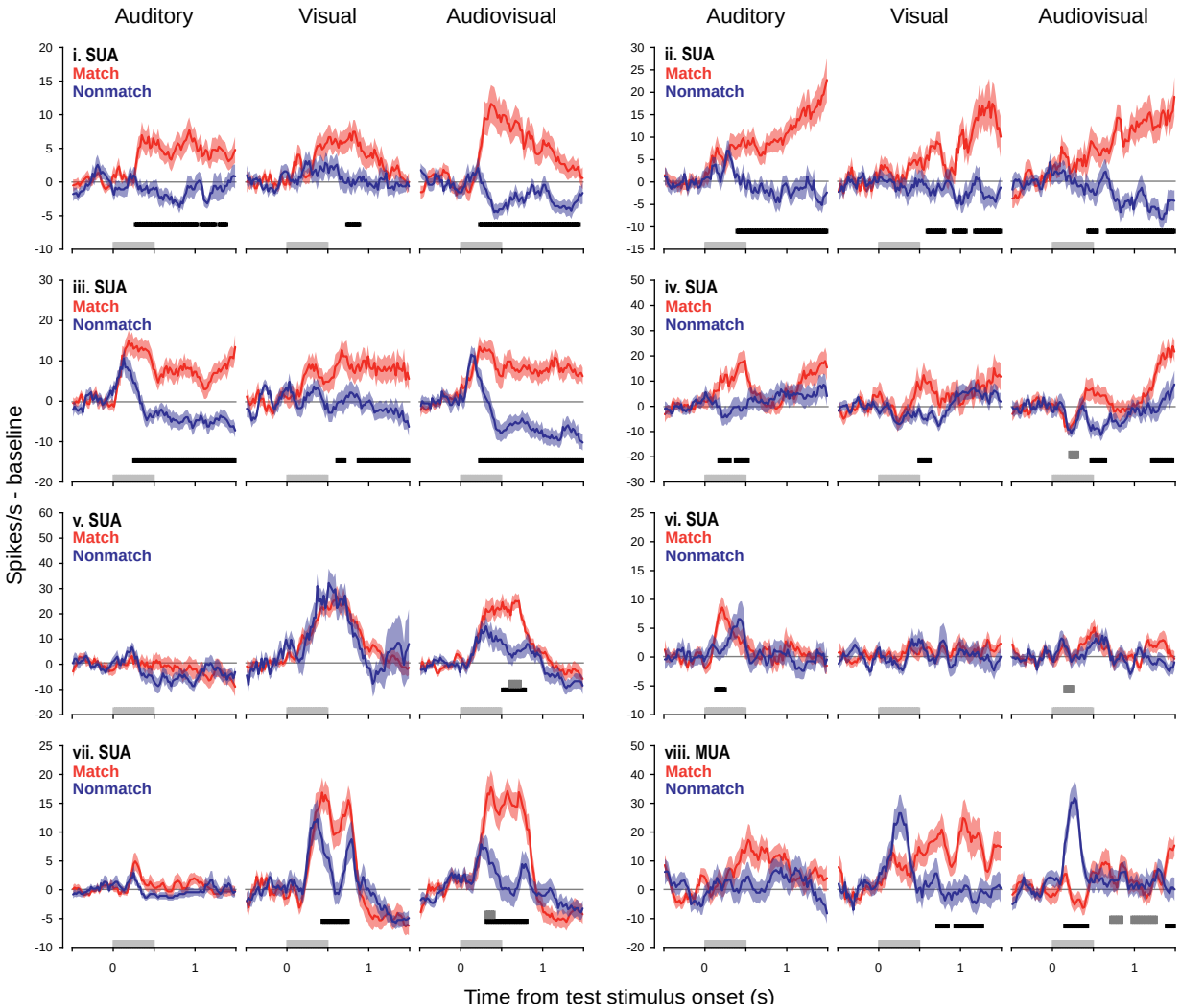

Figure 32. Example units exhibiting significant differences in firing rate elicited by matching versus nonmatching test stimuli (M-NM). Significant M-NM differences were detected for all modalities in some units (i–iv), but just one or two modalities in others (units v–vii). Most units exhibited higher firing rates on match trials ("match enhancement", units i–vii), but in some cases, nonmatch firing rates were greater ("match suppression", unit viii). In some cases (units iv–viii), significant interactions were obtained between *trial type* (match, nonmatch) and *modality* (AV, U_{max}), suggesting audiovisual integrative properties of the M-NM discrimination. Stimulus periods are represented by gray bars abutting the abscissae. The narrow black bands below the firing histograms indicate periods during the trial where significant M-NM differences were obtained in a 100-ms sliding window analysis, advancing in 20-ms steps ($p < .01$, ≥2 consecutive analysis steps) and the thicker gray bands denote periods of significant *trial type* × *modality* interactions.

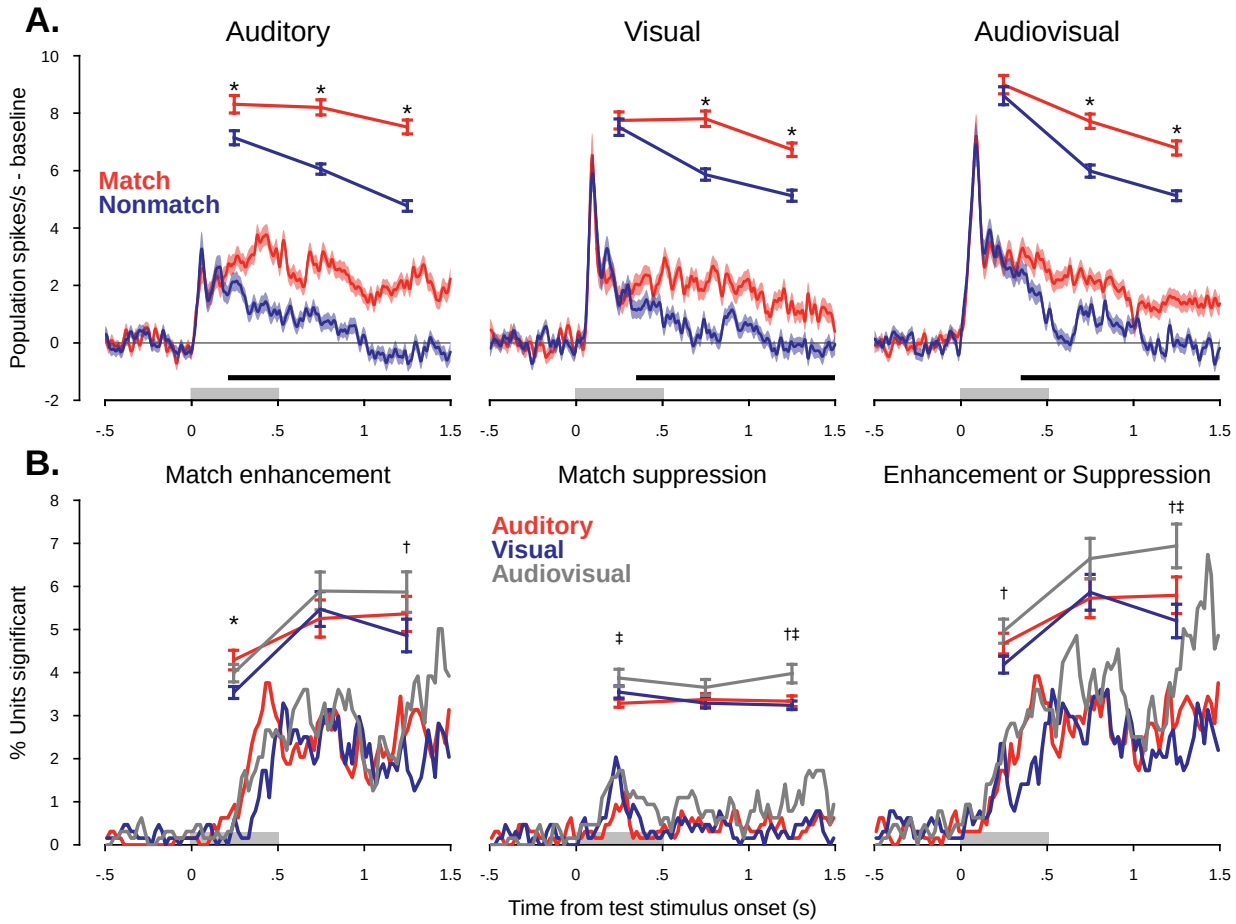

Figure 33. Population summary of differences in neuronal activity evoked by matching and nonmatching test stimuli. (A) Population mean firing rates (spikes/s minus baseline) for auditory (left panel), visual (middle panel), and audiovisual trials (right panel). The insets in each panel depict mean (±SEM) firing rates sampled within successive, non-overlapping 500-ms periods spanning the test stimulus period and the ensuing pre-response delays (the X and Y scales are the same as those used for the main panels). At the population level, differences in firing rates between match and nonmatch trials were greater for auditory than visual or audiovisual trials. Mean (±SEM) firing rates are depicted by dark central lines (plus lighter shaded bands). Stimulus periods are represented by gray bars abutting the abscissae. The narrow black bands below the firing histograms indicate periods during the trial where significant differences were obtained in a 100-ms sliding window analysis, advancing in 20-ms steps ($p < .01$, ≥2 consecutive analysis steps). Post hoc comparisons for insets ($p < .05$): *Match ≠ Nonmatch (B) Percentages of units exhibiting significant match enhancement and suppression effects per analysis step (20 ms) were generally highest for audiovisual trials (with the exception of greater auditory enhancement effects during the test stimulus period). The insets in each panel depict the mean (±SEM) percentages of units with significant effects sampled within successive, non-overlapping 500-ms periods spanning the test stimulus period and the ensuing pre-response delays (the X and Y scales are the same as those used for the main panels). Pairwise comparisons for insets ($p < .05$): *Auditory ≠ Visual, †Visual ≠ Audiovisual, ‡Audiovisual ≠ Auditory.

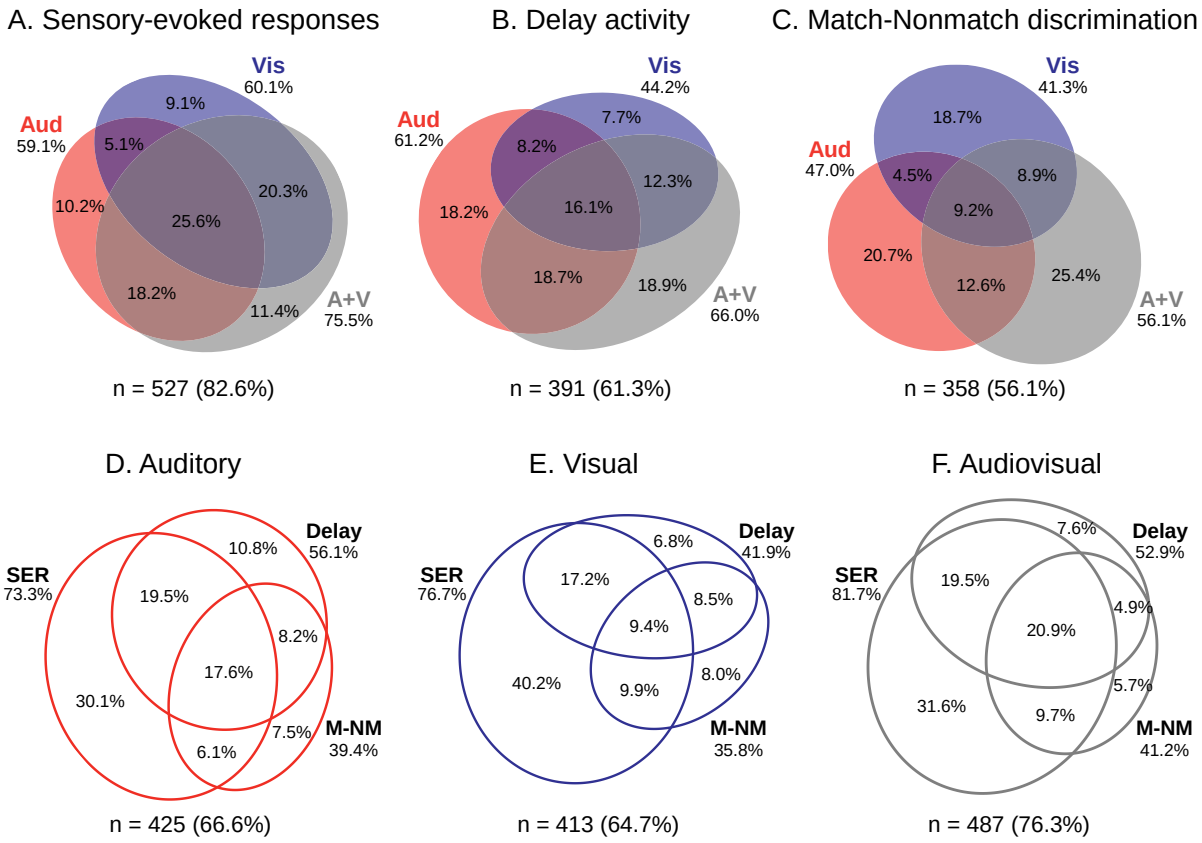

Figure 34. Euler diagram summaries of significant sensory-evoked responses (SERs), delay activity, and match-nonmatch discrimination (M-NM) for each modality and effect type, using the same conventions as Figures 27 and 29. For diagrams A–B and D–F, SERs and delay activity are collapsed across cue type (sample, match, nonmatch; see Figures 27 and 29 for breakdowns by cue type). (A–C) Substantial overlap among modalities was observed for each response type, although in general, overlap was most common for SERs, followed by delay activity, and M-NM effects. (D–F) There was also substantial overlap among response types, i.e., units exhibiting combinations of significant SERs, delay activity, and/or M-NM effects. In general, "response overlap" was most common in units with significant audiovisual effects (F), and the portion of units with at least one significant effect was higher for audiovisual than auditory or visual trials.

Table 3. Significant responses by unit type, modality[a], response type[b], and intersection[c]

	Modality intersections by response type							Units w/ sig. effect[d]
	Sensory-evoked response							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	11.2%	12.9%	14.2%	4.7%	17.2%	20.7%	19.0%	78.9%
MUA	9.5%	6.1%	9.2%	5.4%	19.0%	20.0%	30.5%	85.8%
Total	10.2%	9.1%	11.4%	5.1%	18.2%	20.3%	25.6%	82.6%
	Delay activity							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	17.7%	7.3%	15.2%	7.9%	15.9%	19.5%	16.5%	55.8%
MUA	18.5%	7.9%	21.6%	8.4%	20.7%	7.0%	15.9%	66.0%
Total	18.2%	7.7%	18.9%	8.2%	18.7%	12.3%	16.1%	61.3%
	Match-nonmatch discrimination							
Modality	A	V	AV	A ∩ V	A ∩ AV	V ∩ AV	A ∩ V ∩ AV	
SUA	23.8%	19.0%	28.0%	3.6%	7.7%	8.3%	9.5%	57.1%
MUA	17.9%	18.4%	23.2%	5.3%	16.8%	9.5%	8.9%	55.2%
Total	20.7%	18.7%	25.4%	4.5%	12.6%	8.9%	9.2%	56.1%

	Response type intersections by modality							Units w/ sig. effect[d]
	Auditory							
Response	S	D	M	S ∩ D	S ∩ M	D ∩ M	S ∩ D ∩ M	
SUA	30.3%	12.0%	8.0%	14.9%	7.4%	10.9%	16.6%	59.5%
MUA	30.0%	10.0%	7.2%	22.8%	5.2%	6.4%	18.4%	72.7%
Total	30.1%	10.8%	7.5%	19.5%	6.1%	8.2%	17.6%	66.6%
	Visual							
Response	S	D	M	S ∩ D	S ∩ M	D ∩ M	S ∩ D ∩ M	
SUA	35.7%	7.1%	10.4%	19.8%	7.7%	9.3%	9.9%	61.9%
MUA	43.7%	6.5%	6.1%	15.2%	11.7%	7.8%	9.1%	67.2%
Total	40.2%	6.8%	8.0%	17.2%	9.9%	8.5%	9.4%	64.7%
	Audiovisual							
Response	S	D	M	S ∩ D	S ∩ M	D ∩ M	S ∩ D ∩ M	
SUA	30.5%	9.9%	6.6%	17.4%	11.3%	6.1%	18.3%	72.4%
MUA	32.5%	5.8%	5.1%	21.2%	8.4%	4.0%	23.0%	79.7%
Total	31.6%	7.6%	5.7%	19.5%	9.7%	4.9%	20.9%	76.3%

[a] *A = Auditory; V = Visual; AV = Audiovisual*
[b] *S = Sensory-evoked response; D = Delay activity; M = Match-nonmatch discrimination*
[c] *Intersection percentages based on subsets of units with significant effects reported in right column*
[d] *Percentages of units with significant effects based on SUA = 294, MUA = 344, and Total = 638 units*

or suppression, Figure 34C), the portion of units with significant responses was significantly larger for audiovisual trials (56.1%) than visual trials (41.3%; $p < .05$), and a trend to the same effect was observed for audiovisual versus auditory trials (47.0%; $p = .08$, Cochran's Q test, adjusted for multiple pairwise comparisons).

As in the SER and delay analyses, a significant influence of audiovisual integration was detected for M-NM discrimination effects. Of units with significant effects, 9.2% responded for all three modalities, an additional 4.5% responded for auditory and visual trials, and 25.4% more responded on audiovisual trials alone. Further, for 46.9% units with M-NM effects, a significant *trial type* (match, nonmatch) × *modality* (AV, U_{max}) interaction indicated that the match-nonmatch difference itself was significantly modulated by co-occurrence of sounds and images. For some units, the interactions reflected the fact that match enhancement effects were greater for the audiovisual trials (9.5%, "augmented enhancement"), but in other cases audiovisual match enhancement was either lower than unimodal trials or nonexistent (21.5%, "diminished enhancement"). Parallel effects were observed for inhibitory effects, i.e., audiovisual match suppression was greater in some units (8.4%, "augmented suppression"), but weaker or nonexistent in others (11.5%, "diminished suppression").

As described in detail above, substantial overlap among modalities was observed for SERs, delay-related changes in firing rate, and M-NM discrimination effects (Figures 28 A–C, 30 A–C, and 34 A–C). A direct comparison of the extent of modality overlap for each type of response suggests that integration across modalities in our unit population was most likely for SERs, followed by delay activity, and was least common for M-NM discrimination effects. For instance, 25.6% of units with significant SERs responded to all three modalities, whereas 16.1% of units exhibited significant delay activity for all modalities, and 9.2% exhibited M-NM

discrimination effects for all modalities ($p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). As can be seen in Figure 34 D–F, there were also substantial portions of units that exhibited overlap among response types, i.e., combinations of significant SERs, delay activity, and/or M-NM effects. In general, such "functional overlap" was most common in units with significant audiovisual effects. Thus, 20.9% of units with significant audiovisual effects exhibited significant SERs, delay responses, and M-NM effects, whereas only 17.6% of units with auditory effects and 9.4% of units with visual effects exhibited responses for all three conditions ($p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). Consistent with the results reported above, the proportion of units with any significant effect (SERs, delay responses, or M-NM effects**)** was higher for audiovisual trials (76.3% of all units) than auditory (66.6%) or visual trials (64.7%; $p < .05$, Cochran's Q test, adjusted for multiple pairwise comparisons). Taken together, unit subpopulations with significant responses for multiple conditions highlight widespread integration across sensory modalities and response types at the level of individual units and location cell populations within PFC. Other unit subpopulations responding exclusively within one sensory modality or task condition reveal that units with more specialized functional roles are found within PFC among the units with more general multifaceted response properties. The observation of both selective and integrative response characteristics within SUA and MUA subpopulations (Tables 1–3) adds to growing awareness of both specialized and multifaceted functional roles at the level of single cells and local cell populations (Chandrasekaran & Ghazanfar, 2009; Ghazanfar et al., 2005; Kayser et al., 2008, 2009; Kreiman et al., 2006; Liu & Schreiner, 2007).

5.4 Discussion

Herein, we describe the first unit recordings during a STM task presenting unimodal auditory and visual memoranda on separate trials and compound audiovisual memoranda on other trials. This design allowed assessment of audiovisual integration within both sensory and mnemonic responses using analyses commonly employed in the sensory physiology literature to reveal audiovisual integration of stimulus driven neuronal activity. We contribute to this literature by reporting, for the first time, audiovisual integrative responses during the SERs for units within the dorsal division of the lateral PFC, which were qualitatively comparable to results obtained from other cortical regions such as vlPFC (Sugihara et al., 2006) and STS (Barraclough et al., 2005). However, we extend these findings substantially further by providing evidence for audiovisual integrative influences within mnemonic-related functions ("audiovisual mnemonic integration"), which were observed during the delay periods and in the different firing rates elicited by matching and nonmatching test stimuli. These processes imply that multisensory integration in PFC occurs not only during sensory events themselves, but also for mnemonic representations of those events in the absence of direct sensory stimulation. A distinct but complimentary form of audiovisual mnemonic integration was recently reported by Hwang and Romanski (2015). In their study, all memoranda for each trial had both auditory and visual components, and subjects were trained to detect nonmatching test stimuli which differed from the sample by either the auditory component or visual component or both. For approximately one fifth of units with significant nonmatch detection responses, an interaction was observed for trials in which both stimulus components changed, such that responses elicited by audiovisual changes could not be predicted by responses elicited by either type of unimodal change.

By extending the definition of "delay activity" to include delays separating the test stimulus and response window, we identified a dissociation of delay-related changes in firing, wherein excitatory effects were more often observed during the match delay and inhibitory effects were more common during the sample delay. Across the unit population, the balance of delay excitation and suppression favored elevated, sustained firing rates for all modalities only during the match delay, in which case a button press was anticipated. This outcome is most consistent with the "memory cell" phenomenon first reported by Fuster and Alexander (1971), in which elevated firing rates were observed between a sensory cue and motor response in the delayed response paradigm. Indeed, the absence of elevated population firing rates for other delay types raises the possibility that anticipation of a motor response may be the best predictor of this form of sustained, elevated delay activity.

For the sample delay, population firing rates were initially elevated after sample stimulus offset, but diminished as the delay progressed to the extent that firing rates fell below baseline for all modalities before test stimulus onset. This outcome replicates an earlier study of primary auditory cortical activity during a similar Go/No-go auditory DMS paradigm, wherein firing rates declined during the retention interval (sample delay), eventually falling below baseline prior to test stimulus onset (Bigelow et al., 2014). One speculative interpretation offered in that study was that diminishing firing rates may reflect a mechanism for increasing the signal-to-noise ratio between spontaneous and stimulus-evoked firing rates, thereby increasing the likelihood of detecting and accurately processing an anticipated behaviorally relevant stimulus. The current results imply that such processing may be qualitatively similar across modalities, but may be especially pronounced for audition.

In general, fewer delay responses were observed during the nonmatch delay compared to

sample or match delays, with the population firing returning to baseline for all modalities prior to the response window. Here, it is worth noting that withholding a button press following a nonmatching test stimulus does not imply the absence of motor activity. To the contrary, inhibition of a motor response requires coordinated interaction among primary and sensory motor areas, both cortical and subcortical. Taken together, the diverse physiological responses among the different delay types in our study suggest that "retention" is not a unitary phenomenon, but instead requires different forms of neural activation reflecting cross-temporal connections among behaviorally relevant sensory events and/or motor responses.

The distinctive forms of delay activity observed among modality formats and delay types at the unit and population levels are relevant to a long hypothesized distinction between prospective and retrospective forms of memory (Wasserman, 1985). As first described by Konorski (1967), prospection entails anticipated future events, whereas retrospection requires reflection upon past events. For instance, following presentation of the sensory cue, the delayed response task could theoretically be solved on the basis of prospective memory alone, i.e., subjects could anticipate the correct behavioral response without maintaining any representation of the sensory cue. By contrast, the correct response in the DMS task cannot be anticipated following the initial sensory cue (the sample), and therefore must be selected following retrospective comparison with the test stimulus. Although this parsimonious model fits several behavioral observations (e.g., orienting responses toward the correct response location following the sensory cue in the delayed response task), it does not consistently align with the diverse delay responses of individual units and local populations within PFC. For instance, a prospective account of memory bridging the test stimulus with the appropriate behavioral response might predict uniform activity among modality formats during test delays. This is because the

anticipated behavioral response must be retained, which is uniform among modality formats, but mnemonic representation of the sensory cue itself (the test stimulus) is no longer necessary. In contradiction of this hypothesis, many units exhibited test delay activity that depended not only upon the anticipated behavioral response, but also upon the modality format of the test stimulus (e.g., units that exhibited differential delay responses among modality formats). Moreover, although a strong dissociation was observed among delay types at the population level, exceptions to the observed trends were observed in individual unit responses. Thus, while a prospective memory account might have predicted the population difference between match and nonmatch delay activity (reflecting the different anticipated response types), it would not easily account for individual units with similar responses for both delay types (e.g., excitatory activity for both match and nonmatch delays). Thus, although the proposed distinction between prospection and retrospection has served as an important step toward developing theories of memory that account for different forms of retention, the current data suggest that it may be too general and strict to fully account for the diverse physiological phenomena thought to underlie memory.

Several neurophysiological outcomes were associated with the behavioral "bimodal advantage", i.e., the superior average performance observed on audiovisual trials. First, delay-related changes in firing rate were generally observed most frequently on audiovisual trials, with the exception that match delay effects were equally likely on audiovisual and auditory trials, and less likely on visual trials (Figures 30 A–C, 34B). Second, M-NM discrimination effects were generally most likely on audiovisual trials. Because match enhancement in PFC and other areas has been shown to correlate with behavioral accuracy in previous studies of unimodal STM (Bigelow & Poremba, 2014; Lee et al., 2009; Plakke et al., 2013; Russ et al., 2008), our *a priori*

expectation was that the prevalence of match enhancement among modalities would reflect differences in behavioral accuracy in the current study. This hypothesis turned out to be partially correct, but with an important caveat regarding intermodal differences in match enhancement and suppression ratios. Specifically, suppression effects were less common for auditory trials, such that at the population level, mean firing rate differences between match and nonmatch trials were higher on auditory trials than visual or audiovisual trials (Figure 33A). However, when enhancement and suppression were considered together, more units exhibited M-NM discrimination effects for audiovisual trials (Figure 33B, 34C). Considered together, the association between the physiological and behavioral outcomes among sensory modalities lends support to traditional interpretations of delay activity and M-NM effects as neural substrates of STM.

In summary, we report unit subpopulations in dlPFC with responses reflecting multisensory integrative responses and mnemonic-related activation, as well as the convergence of these two general functions. Comparing overlapping effects among sensory modalities and response types revealed that audiovisual integration occurs for all response types, with the most extensive overlap occurring during direct stimulation (SERs), followed by delay periods, and finally M-NM discrimination responses (Figure 34 A–C). Further, units responding to each sensory modality significant exhibited overlapping effects among response types, with the greatest degree of "functional overlap" observed during audiovisual trials (Figure 34 D–F). The latter observation raises the possibility that units that integrate information across multiple sensory modalities may be more likely to integrate activity underlying multiple functions. Such cross-modal, cross-temporal, and cross-functional processes are highly consistent with theories of PFC function that emphasize integration of sensory and motor functions across time (Fuster,

2008d; Miller & Cohen, 2001).

Chapter 6: General summary

STM is vital to adaptive behavior in humans, nonhuman primates, and other animals. Major effort has been devoted over the last century to understanding the behavioral aspects of STM as well as its underlying neural circuitry. Because of their many homologies with humans, nonhuman primates have served as a primary animal model for understanding STM, particularly its neural substrates. Herein, this animal model has been further substantiated by behavioral evidence from two experiments. First, like nonhuman primates, humans exhibited relatively limited short-term and recognition memory capabilities for auditory information (Chapter 2). Second, like humans, monkeys exhibited superior memory performance for audiovisual memoranda compared to unimodal auditory or visual memoranda (Chapter 4). Unexpectedly, two subjects exhibited superior average unimodal performance on auditory trials, whereas a third subject with similar prior training history exhibited superior visual performance. These outcomes call for increased attention to individual differences and prior experience in future studies of the emergence of differences in memory among modalities.

Most studies of STM have employed the DMS task or one of its derivatives, wherein sample and test stimuli (and/or behavioral responses) are separated by a retention interval, and must therefore be linked by STM processes. Although many studies have emphasized the retention interval as a primary index for STM capabilities, the ability to cope with PI from prior trials has also been shown to be a major determinant of task performance. Herein, a more detailed characterization of PI in the auditory modality has been provided for nonhuman primates in Chapter 3. Several analyses in this chapter suggest that monkeys may rely on a criterion level of familiarity to guide "same" versus "different" decisions required by the task,

similar to findings from earlier studies in human subjects. Considering overall performance generally, the effects of PI in auditory STM were shown to be qualitatively similar to previous studies of visual STM, inasmuch as accuracy increased as a function of stimulus set size and, to a lesser extent, with longer intertrial intervals. Further, as reported in Chapter 4, graded effects of intertrial PI were observed for at least one subject in both the auditory and visual modalities, as well as on audiovisual trials. Intertrial PI effects were most consistently observed across subjects on auditory trials, suggesting auditory STM in nonhuman primates may be somewhat more susceptible to PI. Nevertheless, considering the effects of PI in the context of overall performance outcomes among sensory modalities, PI should be viewed as an important influence, but not the sole determinant of STM performance.

In addition to behavioral studies such as those presented in Chapters 2–4, understanding of STM has been greatly expanded by investigations of its underlying biological mechanisms. Although STM should not be viewed as a unitary function enabled by a single brain region, a large body of evidence has revealed a principal role for primate lateral PFC in STM, particularly for integrating and retaining visual and auditory information for the guidance of future actions. This evidence has primarily emerged from several disciplines of neuroscience: (I) Anatomical connections have been demonstrated between the lateral PFC and visual, auditory, and motor cortices. (II) Sensory physiology has characterized the visual and auditory responsiveness of the lateral PFC and has begun to address audiovisual integrative responses. (III) Neuropsychological investigations in humans and animals have demonstrated deficits in auditory and visual STM following lesions of the lateral PFC. (IV) Neurophysiological recordings in nonhuman primates during visual and auditory STM tasks have consistently reported retention- and recognition-related activity in the lateral PFC. (V) Neuroimaging studies have revealed that the lateral PFC in

humans is activated during visual and auditory STM.

In spite of impressive progress, much of the neural circuitry underlying STM remains to be explored. The current work focused on several outstanding questions regarding the role of the lateral PFC in retaining audiovisual information. First, although it is well known at a basic level that the lateral PFC is centrally involved in retaining representations of auditory and visual events, cellular correlates of integration across these modalities during STM have not yet been identified. Second, although delay-related changes in firing rate have long been considered a direct correlate of retention during STM, it is not clear how the lateral PFC may be differentially engaged during retention when action may or may not be anticipated. Finally, although behavioral studies have identified a performance advantage for audiovisual memoranda during STM, little is understood about how this advantage might be reflected in brain activity. To address these questions, experiments were conducted that coupled neurophysiological recordings in primate lateral PFC during a STM task presenting auditory, visual, and audiovisual memoranda (Chapter 5). This design allowed assessment of audiovisual integrative processing during STM using analyses and techniques commonly employed in the sensory physiology literature to reveal audiovisual integration of stimulus-driven neuronal activity. In addition, the STM task incorporated a Go/No-go behavioral response contingency, which required a button press following identical sample and test stimuli, but no action following nonidentical memoranda. A pre-response wait period separated the test stimulus from the response window, allowing comparison of delay-related changes in activity when motor responses were anticipated (following matching test stimuli) or not (following sample and nonmatching test stimuli).

In general, the outcomes of this experiment, reported in Chapter 5, revealed widespread audiovisual integrative processes within PFC. Thus, many single cells and local cell populations

exhibited significant sensory-evoked responses for both auditory and visual cues, and others responded exclusively to audiovisual cues. Other subpopulations responded to all three modality formats, and still others exhibited significant interactions among responses evoked by audiovisual stimuli and the unimodal stimulus with the maximum response. These outcomes contribute to literature describing audiovisual integration by reporting such processes for the first time within the dorsal division of the lateral PFC. An additional substantial contribution to this literature comes from evidence of audiovisual integrative influences within physiological processes thought to reflect mnemonic representations of stimuli that are no longer present in the environment. Thus, changes in firing rate during the delay periods often occurred for multiple modalities or during audiovisual trials alone, and as with sensory-evoked responses, significant interactions were observed between delay responses on audiovisual trials and the unimodal trial with the maximum response. Similar multisensory responses and integration effects were observed in the differential firing rates evoked by matching and nonmatching test stimuli. Finally, both of these mnemonic-related responses were observed more frequently during audiovisual trials compared to either unimodal trial type, a finding that corresponds to the superior average behavioral performance on audiovisual trials. These outcomes imply that audiovisual integration not only occurs during direct sensory stimulation, but also during mnemonic processing related to events that have passed from the environment.

Evaluating delay-related changes in firing rate separately for different delay types reveled substantial differences in activity during delays in which it was possible to anticipate a behaviorally relevant sensory event (sample delay), a motor response (match delay), or neither (nonmatch delay). Units with robust, elevated firing rates (classically designated as “memory cells”) were most commonly observed during match delays, in which case a button press

subsequently occurred. By contrast, suppressed firing rates were most frequently observed during the sample delay, and the fewest responses of any kind (excitatory or inhibitory) were observed during nonmatch delays. These outcomes partially echo earlier reports of elevated firing rates during STM tasks such as the delayed response paradigm, in which a sensory cue is used to determine a subsequent motor response. However, they call for a dissociation among delay types, wherein (a) sustained, elevated firing is most likely when a motor response is predicted, (b) spontaneous activity is most likely to be inhibited prior to a predictable, behaviorally relevant sensory event, and (c) changes in firing rate, in either direction, are least likely if neither event is expected.

Together, the current results contribute several significant steps toward a full account of STM. In addition to resolving several ambiguities surrounding STM for auditory, visual, and audiovisual information both at the behavioral and neurological levels, the findings presented herein may be of potential benefit for understanding various abnormalities and pathologies in humans. For instance, STM and audiovisual integration at the behavioral and neurophysiological levels are disrupted in neurological disorders such as multiple sclerosis (Litvan et al., 1988) and schizophrenia (Park & Holzman, 1992), as well as in neurodegenerative diseases including Alzheimer's (Becker, 1988) and Parkinson's disease (Lewis et al., 2003). Deficiencies in STM processing can also partially account for deficits in attention, executive function, general intelligence, reading ability, and language comprehension (Baddeley, 2003). Progress in understanding normal STM processing and audiovisual integration, including substrates of these processes within lateral PFC, will contribute to a more complete foundation upon which abnormal functioning can be understood.

References

Artchakov D, Tikhonravov D, Ma Y, Neuvonen T, Linnankoski I, Carlson S (2009) Distracters impair and create working memory-related neuronal activity in the prefrontal cortex. Cereb Cortex 19(11):2680–2689.

Artchakov D, Tikhonravov D, Vuontela V, Linnankoski I, Korvenoja A, Carlson S (2007) Processing of auditory and visual location information in the monkey prefrontal cortex. Exp Brain Res 180(3):469–479.

Baddeley AD (1990) When memory fails. In: Human memory: theory and practice (Baddeley AD, ed), pp 169–89. Needham Heights: Allyn & Bacon.

Baddeley A (2003) Working memory: looking back and looking forward. Nat Rev Neurosci 4(10):829–839.

Baddeley A (1992) Working memory. Science 255(5044):556–559.

Badre D, Wagner AD (2005) Frontal lobe mechanisms that resolve proactive interference. Cereb Cortex 15(12), 2003–2012.

Barbas H, Blatt GJ (1995) Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. Hippocampus 5(6):511–533.

Barbas H, De Olmos J (1990) Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. J Comp Neurol 300(4):549–571.

Barbas H, Ghashghaei HT, Rempel-Clower NL, Xiao D (2002) Anatomic basis of functional specialization in prefrontal cortices in primates. In: Handbook of neuropsychology (Grafman J, ed), pp 1–27. Amsterdam: Elsevier.

Barbas H, Henion TH, Dermon CR (1991) Diverse thalamic projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol 313(1):65–94.

Barbas H, Pandya DN (1989) Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. J Comp Neurol 286(3):353–375.

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17:377–391.

Batuev AS, Pirogov AA, Orlov AA (1979) Unit activity of the prefrontal cortex during delayed alternation performance in monkey. Acta Physiol Acad Sci Hung 53(3):345–353.

Bauer RH, Fuster JM (1976) Delayed-matching and delayed-response deficit from cooling dorsolateral prefrontal cortex in monkeys. J Comp Physiol Psychol 90(3):293–302.

Bauer RH, Steele TL (1985) Short-term memory for haptic cues in monkeys (Macaca mulatta). Anim Learn Behav 13(3):291–302.

Bechara A, Damasio H, Tranel D, Anderson SW (1998) Dissociation of working memory from decision making within the human prefrontal cortex. J Neurosci 18(1):428–437.

Becker JT (1988) Working memory and secondary memory deficits in Alzheimer's disease. J Clin Exp Neuropsychol 10(6):739–753.

Bigelow J, Poremba A (2013a) Auditory memory in monkeys: costs and benefits of proactive interference. Am J Primatol 75:425–434.

Bigelow J, Poremba A (2013b) Auditory proactive interference in monkeys: the roles of stimulus set size and intertrial interval. Learn Behav 41(3):319–332.

Blum RA (1952) Effects of subtotal lesions of frontal granular cortex on delayed reaction in monkeys. Arch Neurol Psychiatry 67(3):375–386.

Bodner M, Kroger J, Fuster JM (1996) Auditory memory cells in dorsolateral prefrontal cortex. Neuroreport 7(12):1905–1908.

Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC (1997) A parametric study of prefrontal cortex involvement in human working memory. Neuroimage 5(1):49–62.

Brown MW, Aggleton JP (2001) Recognition memory: what are the roles of the perirhinal cortex and hippocampus? Nat Rev Neurosci 2:51–61.

Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM (1999) Dissociation between the effects of damage to perirhinal cortex and area TE. Learn Mem 6(6):572–599.

Burwitz L (1974) Proactive interference and directed forgetting in short-term motor memory. J Exp Psychol 102(5):799–805.

Butters N, Pandya D (1969) Retention of delayed-alternation: effect of selective lesions of sulcus principalis. Science 165(899):1271–1273.

Campbell RJ, Harlow HF (1945) Problem solution by monkeys following bilateral removal of the prefrontal areas. V. Spatial delayed reactions. J Exp Psychol 35(2):110–126.

Carmichael ST, Price JL (1995a) Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J Comp Neurol 363(4):615–641.

Cavada C, Compañy T, Tejedor J, Cruz-Rizzolo RJ, Reinoso-Suárez F (2000) The anatomical connections of the macaque monkey orbitofrontal cortex. A review. Cereb cortex. 10(3):220–242.

Cermak LS (1970) Decay of interference as a function of the intertrial interval in short-term memory. J Exp Psychol 84(3):499–501.

Cohen JS, Reid S, Chew K (1994) Effects of varying trial distribution, intra- and extramaze cues, and amount of reward on proactive interference in the radial maze. Anim Learn Behav 22(2):134–142.

Cohen MA, Evans KK, Horowitz TS, Wolfe JM (2011) Auditory and visual memory in musicians and nonmusicians. Psychon Bull Rev 18(3):586–591.

Cohen MA, Horowitz TS, Wolfe JM (2009) Auditory recognition memory is inferior to visual recognition memory. Proc Natl Acad Sci USA 106(14):6008–6010.

Cohen YE, Russ BE, Gifford GW III (2005) Auditory processing in the posterior parietal cortex. Behav Cogn Neurosci Rev 4(3):218–231.

Colombo M, D'Amato MR (1986) A comparison of visual and auditory short-term memory in monkeys (Cebus apella). QJ Exp Psychol B 38(4):425–448.

Constantinidis C, Procyk E (2004) The primate working memory networks. Cogn Affect Behav Neurosci 4(4):444–465.

Corballis MC (1966) Rehearsal and decay in immediate recall of visually and aurally presented items. Can J Psychol 20:43–51.

Cothros N, Köhler S, Dickie EW, Mirsattari SM, Gribble PL (2006) Proactive interference as a result of persisting neural representations of previously learned motor skills in primary motor cortex. J Cogn Neurosci 18(12):2167–2176.

Craig KS, Berman MG, Jonides J, Lustig C (2013) Escaping the recent past: Which stimulus dimensions influence proactive interference? Mem Cognit 41(5):650–670.

Cromer JA, Roy JE, Buschman TJ, Miller EK (2011) Comparison of primate prefrontal and premotor cortex neuronal activity during visual categorization. J Cogn Neurosci 23(11):3355–3365.

Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7(9):415–423.

D'Amato MR (1973) Delayed matching and short-term memory in monkeys. In: The psychology

of learning and motivation: advances in research and theory (Bower GH, ed), pp 227–269. New York: Academic Press.

D'Amato MR, Colombo M (1985) Auditory matching-to-sample in monkeys (Cebus apella). Anim Learn Behav 13:375–382.

D'Amato MR, Worsham RW (1974) Retrieval cues and short-term memory in capuchin monkeys. J Comp Physiol Psychol 86(2):274–282.

Delogu F, Raffone A, Belardinelli MO (2009) Semantic encoding in working memory: Is there a (multi)modality effect? Memory 17:655–663.

D'Esposito M, Postle BR, Jonides J, Smith EE (1999) The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. Proc Natl Acad USA 96(13):7514–7519.

De Rosa E, Hasselmo ME (2000) Muscarinic cholinergic neuromodulation reduces proactive interference between stored odor memories during associative learning in rats. Behav Neurosci 114(1):32–41.

Dum RP, Strick PL (2003) An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. J Neurophysiol 89(1):634–639.

Dunnett SB, Martel FL (1990) Proactive interference effects on short-term memory in rats: I. Basic parameters and drug effects. Behav Neurosci 104(5):655–665.

Edhouse WV, White KG (1988) Sources of proactive interference in animal memory. J Expl Psychol: Anim Behav Process 14(1):56–70.

Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. Annu Rev Neurosci 30:123–152.

Feredoes E, Tononi G, Postle BR (2006) Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. Proc Natl Acad Sci USA 103(51):19530–19534.

Franz SI (1907) On the function of the cerebrum: the frontal lobes. Arch Psychol 2:1–64.

Fritz J, Mishkin M, Saunders RC (2005) In search of an auditory engram. Proc Natl Acad Sci USA 102(26):9359–9364.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1993a) Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic "scotomas". J Neurosci 13(4):1479–1497.

Funahashi S, Chafee MV, Goldman-Rakic PS (1993b) Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. Nature 365(6448):753–756.

Fuster JM (2008a) Anatomy of the prefrontal cortex. In: The prefrontal cortex (Fuster JM, ed), pp 7–58. Amsterdam; Boston: Academic Press/Elsevier.

Fuster JM (2008b) Human neuropsychology. In: The prefrontal cortex (Fuster JM, ed), pp 171–219. Amsterdam; Boston: Academic Press/Elsevier.

Fuster JM (2008c) Neuroimaging. In: The prefrontal cortex (Fuster JM, ed), pp 285–331. Amsterdam; Boston: Academic Press/Elsevier.

Fuster JM (2008d) Overview of prefrontal functions: the temporal organization of behavior. In: The prefrontal cortex (Fuster JM, ed), pp 333–385. Amsterdam; Boston: Academic Press/Elsevier.

Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. Science 173(997):652–654.

Fuster JM, Alexander GE (1973) Firing changes in cells of the nucleus medialis dorsalis associated with delayed response behavior. Brain Res 61:79–91.

Fuster JM, Bauer RH (1974) Visual short-term memory deficit from hypothermia of frontal cortex. Brain Res 81(3):393–400.

Fuster JM, Bauer RH, Jervey JP (1985) Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. Brain Res 330(2):299–307.

Fuster JM, Bodner M, Kroger JK (2000) Cross-modal and cross-temporal association in neurons of frontal cortex. Nature 405(6784):347–351.

Fuster JM, Jervey JP (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. Science 212(4497):952–955.

Fuster JM, Jervey JP (1982) Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. J Neurosci 2(3):361–375.

Gemba H, Sasaki K (1988) Changes in cortical field potentials associated with learning processes of audio-initiated hand movements in monkeys. Exp Brain Res 70(1):43–49.

Ghazanfar AA, Hauser MD (2001) The auditory behaviour of primates: a neuroethological perspective. Curr Opin Neurobiol 11(6):712–720.

Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. (2005) Multisensory integration of

dynamic faces and voices in rhesus monkey auditory cortex. J Neurosci 25(20):5004–5012.

Gibson JR, Maunsell JH (1997) Sensory modality speci fi city of neural activity related to memory in visual cortex. J Neurophysiol 78(3):1263–1275.

Gleitman H, Jung L (1963) Retention in rats: The effect of proactive interference. Science 142(3600):1683–1684.

Goldman PS, Rosvold HE (1970) Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. Exp Neurol 27(2):291–304.

Goldman PS, Rosvold HE, Vest B, Galkin TW (1971) Analysis of the delayed-alternation deficit produced by dorsolateral prefrontal lesions in the rhesus monkey. J Comp Physiol Psychol 77(2):212–220.

Goldman-Rakic PS (1987) Motor control function of the prefrontal cortex. Ciba Found Symp 132:187–200.

Goldman-Rakic PS (1995) Cellular basis of working memory. Neuron 14(3):477–485.

Grant DS (1975) Proactive interference in pigeon short-term memory. J Exp Psychol: Anim Behav Process 1(3):207–220.

Grant DS (1976) Effect of sample presentation time on long-delay matching in the pigeon. Learn Motiv 7(4):580–590.

Grant DS (1981) Intertrial interference in rat short-term memory. J Exp Psychol: Anim Behav Process 7(3):217–227.

Greene RL (1992) Human Memory. Earlbaum, Hillsdale, NJ.

Hackett TA, Stepniewska I, Kaas JH (1999) Prefrontal connections of the parabelt auditory cortex in macaque monkeys. Brain Res 817(1–2):45–58.

Hampton RR, Shettleworth SJ, Westwood RP (1998) Proactive interference, recency, and associative strength: Comparisons of black-capped chickadees and dark-eyed juncos. Anim Learn Behav 26(4):475–485.

Harlow HF, Bromer JA (1938) A test apparatus for monkeys. Psychol Rec 2:434–436.

Harlow HF, Dagnon J (1943) Problem solution by monkeys following bilateral removal of the prefrontal areas. I. The discrimination and discrimination-reversal problems. J Exp Psychol 32(4):351–356.

Hartshorne JK (2008) Visual working memory capacity and proactive interference. PLoS One 3(7):e2716.

Hashiya K, Kojima S (2001) Acquisition of auditory–visual intermodal matching-to-sample by a chimpanzee (*Pan troglodytes*): comparison with visual–visual intramodal matching. Anim Cogn 4:231–39.

Hayes KJ, Thompson R (1953) Nonspatial delayed response to trial-unique stimuli in sophisticated chimpanzees. J Comp Physiol Psychol 46(6):499–500.

Hendrikx AJ (1986) Short-term proactive interference revisited. Bull Psychon Soc 24(5):358–60.

Herman LM (1975) Interference and auditory short-term memory in the bottlenosed dolphin. Anim Learn Behav 3(1):43–48.

Herman LM, Bailey DR (1970) Comparative effects of retroactive and proactive interference in motor short-term memory. J Exp Psychol 86(3):407–415.

Hernández A, Nácher V, Luna R, Zainos A, Lemus L, Alvarez M, Vázquez Y, Camarillo L, Romo R (2010) Decoding a perceptual decision process across cortex. Neuron 66(2):300–314.

Hogan DE, Edwards CA, Zentall TR (1981) Delayed matching in the pigeon: Interference produced by the prior delayed matching trial. Anim Learn Behav 9(3):395–400.

Hwang J, Romanski LM (2015) Prefrontal neuronal responses during audiovisual mnemonic processing.J Neurosci 35(3):960–971.

Isseroff A, Rosvold HE, Galkin TW, Goldman-Rakic PS (1982) Spatial memory impairments following damage to the mediodorsal nucleus of the thalamus in rhesus monkeys. Brain Res 232(1):97–113.

Ito SI (1982) Prefrontal unit activity of macaque monkeys during auditory and visual reaction time tasks. Brain Res 247(1):39–47.

Jacobsen CF (1935) Functions of the frontal association area in primates. Arch Neurol Psychiatry 33:558–69.

Jarrard LE, Moise SL (1971) Short-term memory in the monkey. In Cognitive processes of nonhuman primates (Jarrard LE, ed), pp 1–24. New York: Academic Press.

Jensen AR (1971) Individual differences in visual and auditory memory. J Educ Psychol 62:123–131.

Jitsumori M, Wright AA, Shyan MR (1989) Buildup and release from proactive interference in a rhesus monkey. J Exp Psychol: Anim Behav Process 15(4):329–337.

Jones EG, Powell TP (1970) An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain 93(4):793–820.

Jonides J, Nee DE (2006) Brain mechanisms of proactive interference in working memory. Neuroscience 139(1):181–193

Joseph JP, Barone P (1987) Prefrontal unit activity during a delayed oculomotor task in the monkey. Exp Brain Res 67(3):460–468.

Kane MJ, Engle RW (2000) Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. J Exp Psychol Learn Mem Cogn 26(2):336–358.

Kelly RM, Strick PL (2003) Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J Neurosci 23(23):8432–8444.

Keppel G, Underwood BJ (1962) Proactive inhibition in short-term retention of single items. J Verb LearnVerb Be 1(3):153–161.

Kikuchi-Yorioka Y, Sawaguchi T (2000) Parallel visuospatial and audiospatial working memory processes in the monkey dorsolateral prefrontal cortex. Nat Neurosci 3(11):1075–1076.

Kirkpatrick EA (1894) An experimental study of memory. Psychol Rev 1:602–609.

Kojima S (1985) Auditory short-term memory in the Japanese monkey. Int J Neurosci 25(3–4):255–262.

Konorski J (1967) Integrative activity of the brain. Chicago: University of Chicago Press.

Kubota K, Iwamoto T, Suzuki H (1974) Visuokinetic activities of primate prefrontal neurons during delayed-response performance. J Neurophysiol 37(6):1197–1212.

Kubota K, Niki H (1971) Prefrontal cortical unit activity and delayed alternation performance in monkeys. J Neurophysiol 34(3):337–347.

Larsson M, Bäckman L (1998) Modality memory across the adult life span: evidence for selective age-related olfactory deficits. Exp Aging Res 24(1):63–82.

Lashley KS (1950) In search of the engram. Symp Soc Exp Biol 4:454–82.

Lee JH, Russ BE, Orr LE, Cohen YE (2009) Prefrontal activity predicts monkeys' decisions

during an auditory category task. Front Integr Neurosci 3:16.

Lemus L, Hernández A, Luna R, Zainos A, Nácher V, Romo R (2007) Neural correlates of a postponed decision report. Proc Natl Acad Sci USA 104(43):17174–17179.

Lemus L, Hernández A, Romo R (2009) Neural encoding of auditory discrimination in ventral premotor cortex. Proc Natl Acad Sci USA 106(34):14640–14645.

Lewicki MS (1998) A review of methods for spike sorting: the detection and classification of neural action potentials. Network 9(4):R53–78.

Lewis SJ, Cools R, Robbins TW, Dove A, Barker RA, Owen AM (2003) Using executive heterogeneity to explore the nature of working memory deficits in Parkinson's disease. Neuropsychologia 41(6):645–654.

Litvan I, Grafman J, Vendrell P, Martinez JM, Junqué C, Vendrell JM, Barraquer-Bordas JL (1988) Multiple memory deficits in patients with multiple sclerosis. Exploring the working memory system. Arch Neurol 45(6):607–610.

Loess H, Waugh NC (1967) Short-term memory and intertrial interval. J Verb Learn Verb Be 6(4), 455–460.

Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 23(10):3963–3971.

Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. Nature 390(6657):279–281.

Lu MT, Preston JB, Strick PL (1994) Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. J Comp Neurol 341(3):375–392.

Maki WS, Moe JC, Bierley CM (1977) Short-term memory for stimuli, responses, and reinforcers. J Exp Psychol: Anim Behav Process 3(2):156–177.

Makovski T, Jiang YV (2008) Proactive interference from items previously stored in visual working memory. Mem Cognit 36(1):43–52.

Martin-Elkins CL, Horel JA (1992) Cortical afferents to behaviorally defined regions of the inferior temporal and parahippocampal gyri as demonstrated by WGA-HRP. J Comp Neurol 321(2):177–192.

Martínez-García M, Rolls ET, Deco G, Romo R (2011) Neural and computational mechanisms of postponed decisions. Proc Natl Acad Sci USA 108(28):11626–11631.

Mason M, Wilson M (1974) Temporal differentiation and recognition memory for visual stimuli in rhesus monkeys. J Exp Psychol 103(3):383–390.

Mastroberardino S, Santangelo V, Botta F, Marucci FS, Olivetti Belardinelli M (2008) How the bimodal format of presentation affects working memory: an overview. Cogn Process 9(1):69–76.

Matsuzawa T, ed. (2001) Primate origins of human cognition and behavior. New York: Springer.

May CP, Hasher L, Kane MJ (1999) The role of interference in memory span. Mem Cogn 27:759–767.

McGuire PK, Bates JF, Goldman-Rakic PS (1991b) Interhemispheric integration: II. Symmetry and convergence of the corticostriatal projections of the left and the right principal sulcus (PS) and the left and the right supplementary motor area (SMA) of the rhesus monkey. Cereb Cortex 1(5):408–417.

Mecklinger A, Weber K, Gunter TC, Engle RW (2003) Dissociable brain mechanisms for inhibitory control: Effects of interference content and working memory capacity. Cogn Brain Res 18(1):26–38.

Medin DL (1980) Proactive interference in monkeys: Delay and intersample interval effects are noncomparable. Anim Learn Behav 8(4)L553–560.

Medin DL, Roberts WA, Davis RT, eds. (1976) Processes of animal memory. New Jersey: Erlbaum.

Meredith MA, Stein BE (1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophysiol 56(3):640–662.

Meudell PR (1977) Effects of length of retention interval on proactive interference in short-term visual memory. J Exp Psychol: Hum Learn Mem 3(3):264–269.

Meunier M, Bachevalier J, Mishkin M, Murray EA (1993) Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. J Neurosci 13(12):5418–5432.

Meyer DR, Harlow HF, Settlage PH (1951) A survey of delayed response performance by normal and brain damaged monkeys. J Comp Physiol Psychol 44(1):17–25.

Micallef L, Rodgers P (2014) eulerAPE: Drawing Area-Proportional 3-Venn Diagrams Using Ellipses. PLoS ONE 9(7): e101717.

Miller BL, Cummings JL, eds (1999) The human frontal lobes: functions and disorders. New

York: Guilford Press.

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202.

Miller EK, Desimone R (1994) Parallel neuronal mechanisms for short-term memory. Science 263(5146):520–522.

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16(16):5154–5167.

Miller EK, Li L, Desimone R (1991) A neural mechanism for working and recognition memory in inferior temporal cortex. Science 254(5036):1377–1379.

Miller EK, Li L, Desimone R (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. J Neurosci 13(4):1460–1478.

Milner B (1972) Disorders of learning and memory after temporal lobe lesions in man. Clin Neurosurg 19:421–446.

Mishkin M (1957) Effects of small frontal lesions on delayed alternation in monkeys. J Neurophysiol 20(6):615–622.

Mishkin M, Delacour J (1975) An analysis of short-term visual memory in the monkey. J Exp Psychol: Anim Behav Process 1(4):326–334.

Mohedano-Moriano A, Pro-Sistiaga P, Arroyo-Jimenez MM, Artacho-Pérula E, Insausti AM, Marcos P, Cebada-Sánchez S, Martínez-Ruiz J, Muñoz M, Blaizot X, Martinez-Marcos A, Amaral DG, Insausti R (2007) Topographical and laminar distribution of cortical input to the monkey entorhinal cortex. J Anat 211(2):250–260.

Moise SL (1976) Proactive effects of stimuli, delays, and response position during delayed matching from sample. Anim Learn Behav 4(1-A):37–40.

Munoz-Lopez MM, Mohedano-Moriano A, Insausti R (2010) Anatomical pathways for auditory memory in primates. Front Neuroanat 4:129.

Monsell S (1978) Recency, immediate recognition memory, and reaction time. Cogn Psychol 10(4):465–501.

Münsterberg H (1894) Studies from the Harvard Psychological Laboratory. Psychol Rev 1:34–60.

Murray EA, Mishkin M (1998) Object recognition and location memory in monkeys with

excitotoxic lesions of the amygdala and hippocampus. J Neurosci 18(16):6568–6582.

Nauta WJH (1972) Neural associations of the frontal cortex. Acta Neurobiol Exp 32(2):125–140.

Nęcka E (1992) Cognitive analysis of intelligence: the significance of working memory processes. Pers Individ Dif 13(9):1031–1046.

Nelson CN, Bignall KE (1973) Interactions of sensory and nonspecific thalamic inputs to cortical polysensory units in the squirrel monkey. Exp Neurol 40(1):189–206.

Ng CW, Plakke B, Poremba A (2013) Neural correlates of auditory recognition memory in the primate dorsal temporal pole. J Neurophysiol 111(3):455–469.

Niki H (1974b) Prefrontal unit activity during delayed alternation in the monkey. I. Relation to direction of response. Brain Res 68(2):185–196.

Niki H (1974c) Prefrontal unit activity during delayed alternation in the monkey. II. Relation to absolute versus relative direction of response. Brain Res 68(2):197–204.

Nissen HW, Riesen AH, Nowlis V (1938) Delayed response and discrimination learning by chimpanzees. J Comp Psychol 26:361–386.

Ongür D, An X, Price JL (1998) Prefrontal cortical projections to the hypothalamus in macaque monkeys. J Comp Neurol 401(4):480–505.

Overman WH Jr, Doty RW (1980) Prolonged visual memory in macaques and man. Neurosci 5(11):1825–1831.

Park S, Holzman PS (1992) Schizophrenics show spatial working memory deficits. Arch Gen Psychiatry 49(12):975–982.

Passingham R (1993) The Frontal Lobes and Voluntary Action. New York: Oxford University Press.

Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. Nat Rev Neurosci 6(2):97–107.

Paxinos G, Huang X-F, Toga AW (2000) The Rhesus Monkey Brain in Stereotaxic Coordinates.

San Diego: Academic.

Penney CG (1975) Modality effects in short-term verbal memory. Psychol Bull 82:68–84.

Penney CG (1989) Modality effects and the structure of short-term verbal memory. Mem Cogn 17(4):398–422.

Perrodin C, Kayser C, Logothetis NK, Petkov CI (2011) Voice cells in the primate temporal lobe. Curr Biol 21(16):1408–1415.

Perrodin C, Kayser C, Logothetis NK, Petkov CI (2014) Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. J Neurosci 34(7):2524–2537.

Peters J, Suchan B, Köster O, Daum I (2007) Domain-specific retrieval of source information in the medial temporal lobe. Eur J Neurosci 26(5):1333–1343.

Petrides M (2005) Lateral prefrontal cortex: architectonic and functional organization. Philos Trans R Soc Lond B Biol Sci 360(1456):781–795.

Petrides M, Pandya DN (1994) Comparative architectonic analysis of the human and the macaque frontal cortex. In: Handbook of neuropsychology (Boller F, Grafman J, eds), pp 17–58. Amsterdam: Elsevier.

Petrides M, Pandya DN (1999) Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur J Neurosci 11(3):1011–1036.

Petrides M, Pandya DN (2002a) Association pathways of the prefrontal cortex and functional observations. In: Principles of frontal lobe function (Stuss DT, Knight RT, eds), pp 31–84. New York: Oxford University Press.

Plakke B, Ng CW, Poremba A (2013) Neural correlates of auditory recognition memory in primate lateral prefrontal cortex. Neurosci 244:62–76.

Poremba A, Bigelow J (2013) Neurophysiology of attention and memory processing. In: Neural Correlates of Auditory Cognition, Springer Handbook of Auditory Research, Vol 45 (YE Cohen, AN Popper, and RR Fay, eds), pp 215–250. New York: Springer.

Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, Mishkin M (2003) Functional mapping of the primate auditory system. Science 299(5606):568–572.

Porrino LJ, Crane AM, Goldman-Rakic PS (1981) Direct and indirect pathways from the amygdala to the frontal lobe in rhesus monkeys. J Comp Neurol 198(1):121–136.

Postle, BR, Berger JS, Goldstein JH, Curtis CE, D'Esposito M (2001) Behavioral and neurophysiological correlates of episodic coding, proactive interference, and list length effects in a running span verbal working memory task. Cogn Affect Behav Neurosci 1(1), 10–21.

Postle BR, Brush LN, Nick AM (2004) Prefrontal cortex and the mediation of proactive interference in working memory. Cogn Affect Behav Neurosci 4(4):600–608.

Preuss TM (1995) Do rats have prefrontal cortex? The Rose–Woolsey–Akert program reconsidered. J Cogn Neurosci 7(1):1–24.

Preuss TM, Goldman-Rakic PS (1989) Connections of the ventral granular frontal cortex of macaques with perisylvian premotor and somatosensory areas: anatomical evidence for somatic representation in primate frontal association cortex. J Comp Neurol 282(2):293–316.

Rainer G, Rao SC, Miller EK (1999) Prospective coding for objects in primate prefrontal cortex. J Neurosci 19(13):5493–505.

Rempel-Clower NL, Barbas H (1998) Topographic organization of connections between the hypothalamus and prefrontal cortex in the rhesus monkey. J Comp Neurol 398(3):393–419.

Reynolds TJ, Medin DL (1981) Stimulus interaction and between-trials proactive interference in monkeys. J Exp Psychol: Anim Behav Process 7(4):334–347.

Roberts WA (1980) Distribution of trials and intertrial retention in delayed matching to sample with pigeons. J Exp Psychol: Anim Behav Process 6(3):217–237.

Roberts WA, Kraemer PJ (1982) Some observations of the effects of intertrial interval and delay on delayed matching to sample in pigeons. J Exp Psychol: Anim Behav Process 8(4):342–353.

Roitblat HL, Harley HE (1988) Spatial delayed matching-to-sample performance by rats: Learning, memory, and proactive interference. J Exp Psychol: Anim Behav Process 14(1):71–82.

Rolls ET (1989) Information processing in the taste system of primates. J Exp Biol 146:141–164.

Romanski LM, Averbeck BB (2009) The primate cortical auditory system and neural representation of conspecific vocalizations. Annu Rev Neurosci 32:315–346.

Romanski LM, Bates JF, Goldman-Rakic PS (1999a) Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol 403(2):141–157.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999b) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2(12):1131–1136.

Rosene DL, Van Hoesen GW (1977) Hippocampal efferents reach widespread areas of cerebral cortex and amygdala in the rhesus monkey. Science 198(4314):315–317.

Roth PL (1994) Missing data: a conceptual review for applied psychologists. Pers Psychol 47(3):537–560.

Russ BE, Ackelson AL, Baker AE, Cohen YE (2008) Coding of auditory-stimulus identity in the auditory non-spatial processing stream. J Neurophysiol 99(1):87–95.

Ruusuvirta T (2000) Proactive interference of a sequence of tones in a two-tone pitch comparison task. Psychon Bull Rev 7(2):327–331.

Ruusuvirta T, Astikainen P, Wikgren J (2002) Proactive interference of differently ordered tone sequences with the accuracy and speed of two-tone frequency comparisons. Music Percept 19(4):551–563.

Ruusuvirta T, Wikgren J, Astikainen P (2008) Proactive interference in a two-tone pitch-comparison task without additional interfering tones. Psychol Res 72(1):74–78.

Sands SF, Wright AA (1980) Primate memory: Retention of serial list items by a rhesus monkey. Science 209(4459):938–940.

Schechter PB, Murphy EH (1975) Response characteristics of single cells in squirrel monkey frontal cortex. Brain Res 96(1):66–70.

Schulman S (1964) Impaired delayed response from thalamic lesions. Studies in monkeys. Arch Neurol 11:477–499.

Scott BH, Mishkin M, Yin P (2012) Monkeys have a limited form of short-term memory in audition. Proc Natl Acad Sci USA 109(30):12237–12241.

Scott BH, Mishkin M, Yin P (2013) Effect of acoustic similarity on short-term auditory memory in the monkey. Hear Res 298:36–48.

Shafi M, Zhou Y, Quintana J, Chow C, Fuster J, Bodner M (2007) Variability in neuronal activity in primate cortex during working memory tasks. Neuroscience 146(3):1082–1108.

Shams L, Seitz AR (2008) Benefits of multisensory learning. Trends Cogn Sci 12:411–417.

Shimamura AP (2000) The role of the prefrontal cortex in dynamic filtering. Psychobiology 28(2):207–218.

Sierra-Paredes G, Fuster JM (2002) Reversible impairment of an auditory–visual association

task. In: Virtual Lesions (Lomber G, Galuske R, eds), pp 239–245. Oxford: University Press.

Siwek DF, Pandya DN (1991) Prefrontal projections to the mediodorsal nucleus of the thalamus in the rhesus monkey. J Comp Neurol 312(4):509–524.

Spaet T, Harlow HF (1945) Problem solution by monkeys following bilateral removal of the prefrontal areas. II. Delayed reaction problems involving use of the matching-from-sample method. J Exp Psychol 32(5):424–434.

Squire LR, Schmolck H, Stark SM (2001) Impaired auditory recognition memory in amnesic patients with medial temporal lobe lesions. Learn Mem 8(5):252–256.

Squire LR, Zola-Morgan S (1991) The medial temporal lobe memory system. Science 253:1380–1386.

Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 9(4):255–266.

Stepien LC, Cordeau JP (1960) Memory in monkeys for compound stimuli. Am J Psychol 73(3):388–395.

Sternberg S (1966) High-speed scanning in human memory. Science 153(3736):652–654.

Sugihara T, Diltz MD, Averbeck BB, Romanski LM (2006) Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. J Neurosci 26(43):11138–11147.

Suzuki WA, Amaral DG (1994) Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. J Comp Neurol 350(4):497–533.

Suzuki WA, Zola-Morgan S, Squire LR, Amaral DG (1993) Lesions of the perirhinal and parahippocampal cortices in the monkey produce long-lasting memory impairment in the visual and tactual modalities. J Neurosci 13(6):2430–2451.

Thompson RK, Herman LM (1981) Auditory delayed discriminations by the dolphin: Nonequivalence with delayed-matching performance. Anim Learn Behav 9(1):9–15.

Thompson VA, Paivio A (1994) Memory for pictures and sounds: Independence of auditory and visual codes. Can J Exp Psychol 48(3):380–398.

Treichler FR, Hamilton DM, Halay MA (1971) The influence of delay interval on severity of the spatial alternation deficit in frontal monkeys. Cortex 7(2):143–151.

Underwood BJ (1957) Interference and forgetting. Psychol Rev 64(1):49–60.

Uylings HBM, Van Eden CG (1990) Qualitative and quantitative comparison of the prefrontal cortex in rat and in primates, including humans. In: The prefrontal cortex: its structure, function and pathology. (Uylings HBM, Van Eden CG, De Bruin JPC, Corner MA, Feenstra MGP, eds), pp 31–62. Amsterdam: Elsevier.

Vaadia E, Benson DA, Hienz RD, Goldstein MH Jr (1986) Unit study of monkey frontal cortex: active localization of auditory and of visual stimuli. J Neurophysiol 56(4):934–952.

van Hest A, Steckler T (1996) Effects of procedural parameters on response accuracy: Lessons from delayed (non-)matching procedures in animals. Cogn Brain Res 3(3–4):193–203.

Van Hoesen G, Pandya DN, Butters N (1975) Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. II. Frontal lobe afferents. Brain Res 95(1):25–38.

Visscher KM, Kahana MJ, Sekuler R (2009) Trial-to-trial carryover in auditory short-term memory. J Exp Psychol: Learn Mem Cogn 35(1):46–56.

Visscher KM, Kaplan E, Kahana MJ, Sekuler R (2007) Auditory short-term memory behaves like visual short-term memory. PLoS Biology 5, e56.

Ward G, Avons SE, Melling L (2005) Serial position curves in short-term memory: functional equivalence across modalities. Memory 13:308–317.

Warren JM, Cornwell PR, Warren HB (1969) Unilateral frontal lesions and learning by rhesus monkeys. J Comp Physiol Psychol 69(3):498–505.

Warren JM, Nonneman AJ (1976) The search for cerebral dominance in monkeys. Ann N Y Acad Sci 280:732–744.

Wasserman EA (1985) Prospection and retrospection as processes of animal short-term memory. In: Animal Memory (DF Kendrick, M Rilling, MR Denny, eds.), pp 53–75. Hillsdale, NJ: Erlbaum.

Wasserman EA, Zentall TR, eds. (2006) Comparative cognition: experimental explorations of animal intelligence. New York:Oxford University Press.

Watanabe M (1986) Prefrontal unit activity during delayed conditional go/no-go discrimination in the monkey. I. Relation to the stimulus. Brain Res 382(1):1–14.

Wegener JF (1964) Auditory discrimination behavior of normal monkeys. J. Auditory Res. 4:81–106.

Weiskrantz L, Mihailovic L, Gross CG (1960) Stimulation of frontal cortex and delayed alternation performance in the monkey. Science 131:1443–1444.

Whitney P, Arnett PA, Driver A, Budd D (2001) Measuring central executive functioning: what's in a reading span? Brain Cogn 45:1–14.

Wickens DD (1970) Encoding categories of words: An empirical approach to meaning. Psychol Rev 77:1–15.

Wixted JT, Rohrer D (1993) Proactive interference and the dynamics of free recall. J Exp Psychol: Learn Mem Cognit 19(5):1024–1039.

Wollberg Z, Sela J (1980) Frontal cortex of the awake squirrel monkey: responses of single cells to visual and auditory stimuli. Brain Res 198(1):216–220.

Woloszyn L, Sheinberg DL (2009) Neural dynamics in inferior temporal cortex during a visual working memory task. J Neurosci 29(17):5494–5507.

Wood JN, Grafman J (2003) Human prefrontal cortex: processing and representational perspectives. Nat Rev Neurosci 4(2):139–147.

Worsham RW (1975) Temporal discrimination factors in the delayed matching-to-sample task in monkeys. Anim Learn Behav 3(2):93–97.

Wright AA (1998) Auditory and visual serial position functions obey different laws. Psychon Bull Rev 5(4):564–584.

Wright AA (1999) Auditory list memory and interference processes in monkeys. J Exp Psychol: Anim Behav Process 25(3):284–296.

Wright AA (2006) Memory processing. In: Comparative cognition: Experimental explorations of animal intelligence (EA Wasserman, TR Zentall, eds.), pp 164–185. New York: Oxford University Press.

Wright AA (2007) An experimental analysis of memory processing. J Exp Anal Behav 88(3):405–433.

Wright AA, Katz JS, Ma WJ (2012) How to be proactive about interference: lessons from animal memory. Psychol Sci 23(5):453–458.

Wright AA, Santiago HC, Sands SF, Kendrick DF, Cook RG (1985) Memory processing of serial lists by pigeons, monkeys, and people. Science 229(4710):287–289.

Wright AA, Shyan MR, Jitsumori M (1990) Auditory same/different concept learning by

monkeys. Anim Learn Behav 18(3):287–294.

Wright AA, Urcuioli PJ, Sands SF (1986) Proactive interference in animal memory research. In: Theories of animal memory (D. F. Kendrick, M. Rilling, & R. Denny, eds.), pp 101–125 New Jersey: Erlbaum.

van Hest A, Steckler T (1996) Effects of procedural parameters on response accuracy: lessons from delayed (non-)matching procedures in animals. Cogn Brain Res 3:193–203.

Vogt BA, Pandya DN (1987) Cingulate cortex of the rhesus monkey: II. Cortical afferents. J Comp Neurol 262(2):271–289.

Yerkes RM, Yerkes DN (1928) Concerning memory in the chimpanzee. J Comp Psychol 8:237–271.

Yeterian EH, Pandya DN (1991) Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. J Comp Neurol 312(1):43–67.

Yonelinas AP (2002) The nature of recollection and familiarity: A review of 30 years of research. J Mem Lang 46(3):441–517.

Zald DH, Andreotti C (2010) Neuropsychological assessment of the orbital and ventromedial prefrontal cortex. Neuropsychologia 48(12):3377–3391.

Zentall TR, Hogan DE (1974) Memory in the pigeon: Proactive inhibition in a delayed matching task. Bull Psychon Soc 4(2–A):109–112.