4-6-2009

Statistical and Data Mining Methodologies for Behavioral Analysis in Transgenic Mouse Models of Alzheimer's Disease: Parallels with Human AD Evaluation

Ralph E. Leighty
University of South Florida

Statistical and Data Mining Methodologies for Behavioral Analysis in Transgenic Mouse Models of Alzheimer's Disease: Parallels with Human AD Evaluation

by

Ralph E. Leighty

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Cellular Biology, Microbiology, and Molecular Biology
College of Arts and Sciences
University of South Florida

Co-Major Professor: Gary Arendash, Ph.D.
Co-Major Professor: Huntington Potter, Ph.D.
Gordon Fox, Ph.D.
Patrick Bradshaw, Ph.D.
Chuanhai Cao, Ph.D.

Date of Approval:
April 6, 2009

Keywords: neuropathology, neural networks, caffeine, GRK5, GMCSF

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

STATISTICAL AND DATA MINING METHODOLOGIES FOR BEHAVIORAL ANALYSIS IN TRANSGENIC MOUSE MODELS OF ALZHEIMER'S DISEASE: PARALLELS WITH HUMAN AD EVALUATION

Ralph E. Leighty

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of human senile dementia. Alzheimer's represents a significant public health concern, having widespread social and economic implications. Consequently, protocols for early detection and therapeutic intervention (both behavioral and pharmacologic) constitute important targets for medical investigation. Furthermore, contemporary research depends upon comprehensive neurobehavioral assessment and advanced statistical and computational analytic methodologies for characterizing AD-associated sensorimotor and cognitive impairment, as well as evaluating therapeutic efficacy. This dissertation introduces data mining-based techniques (decision trees, neural networks, support vector machines) for behavioral analysis in both nontransgenic and Alzheimer's transgenic mice, to evaluate the cognitive benefits of long-term caffeine treatment. Both treatment and transgenic effects are identified through advanced statistical (discriminant analysis) and data mining approaches. In addition, a novel mouse-based cognitive assessment paradigm, adapted from a human interference learning AD-diagnostic protocol, is implemented to evaluate both genetic (GRK5) and therapeutic (GM-CSF) effects in mice, against an Alzheimer's transgenic background. Data mining techniques are shown to be comparable to con-

ventional statistical analyses, often providing complementary diagnostic information. Indeed, comparisons between data mining-based and multivariate statistical analyses, with respect to groupwise discriminability, support the use of both methodologies in neurobehavioral research. Future work involving both data mining-based and multivariate statistical analyses of cognitive-behavioral data is discussed, emphasizing the need for longitudinal studies, repeated-measure designs, and spatiotemporal modeling for evaluating the time-course of both human AD and AD-like pathology in transgenic mouse models.

CHAPTER 1

BACKGROUND

1.1 History of Alzheimer's Disease

Alzheimer's disease (AD; OMIM 104300) is a progressive, degenerative neurological disorder, characterized by sensorimotor, perceptual, and cognitive impairment and the development of distinctive neuropathologic lesions. The disorder was first described in 1907 by the German psychiatrist Alois Alzheimer (1864-1915), who reported the case of a 51-year-old woman who initially presented with paranoia, confusion, and memory and language impairment (Alzheimer, 1907; Stelzmann et al., 1995; Maurer et al., 1997). Her mental deterioration continued until her death four and one-half years later. Postmortem neurohistological examination revealed proteinaceous extracellular neuritic plaques and intracellular neurofibrillary tangles; Alzheimer was the first to identify the latter finding (Graeber et al., 1998; Graeber, 1999). The presence of both neuropathologic markers represents confirmatory evidence of Alzheimer's disease, although neither is specific for AD (Markesbery, 1997; Selkoe, 2001; Adlard and Cummings, 2004).

The leading cause of senile dementia is Alzheimer's disease, affecting approximately ten percent of people over the age of 65 years and forty percent of those over 80 years of age (Evans et al., 1989). Currently, about five million Americans have Alzheimer's disease, and the number may increase to as many as sixteen million by 2050 (Hebert et al., 2003; Alzheimer's Association, 2008). The increasing prevalence

of Alzheimer's disease represents a significant public health concern (Brookmeyer et al., 1998), with widespread social and economic consequences.

1.2 Behavioral Characterization of Alzheimer's Disease

Reflecting the progressive neuropathology, the time course of Alzheimer's disease corresponds to a graded decline in mental function occurring over a period of between five and twenty years (Locascio et al., 1995), which resembles a retrogression of normal development (Reisberg et al., 1999). Alzheimer's disease has been called "death ... by a thousand subtractions" (Shenk, 2001) in recognition of this insidious and, ultimately, fatal trajectory of attrition. Overt behavioral (including cognitive) manifestations reflect the extent and progress of the underlying neuropathology (Weiner et al., 2005).

Alzheimer's disease, in the early stage, is marked by significant impairment in short-term (working) memory (McKhann et al., 1984; Forstl and Kurtz, 1999). This condition transcends casual forgetfulness or "absentmindedness" in daily routine. Patients begin experiencing difficulty with tasks requiring concentration and sustained attention, planning and organization. Psychiatric symptoms, including depression, delusions, and anxiety, may appear toward the end of this stage (Hart et al., 2003) and persist throughout the patient's life. These deficits impede the performance of daily activities, and contribute to increased accident and hazard risks to patients.

In the moderate stage of Alzheimer's disease, further declines in short-term memory and thinking skills compromise the patient's ability to care for him- or herself. Erosion of the integrity of long-term memory leads to further disorientation and confusion. This confusion, in turn, causes agitation, anxiety, restlessness, and aggression. Delusions and hallucinations of increasing severity, compulsive wandering and hoarding behaviors are observed as well (Devanand et al., 1997).

Nearly complete loss of mental function, in addition to profound motor impairment, characterizes the advanced stage of Alzheimer's disease. These patients are in extremely fragile health, often becoming bedridden and sinking into a minimally-responsive vegetative state before finally succumbing to secondary illnesses (e.g., respiratory and/or cardiovascular pathology, opportunistic infection) (Souren et al., 1995). After clinical diagnosis of AD, the average duration of patient survival is between five and eight years (Bracco et al., 1994).

1.3 Pathological Characterization of Alzheimer's Disease

Before Alzheimer described the association between progressive dementia and cortical lesions, many age-associated cerebrovascular structural anomalies had been reported in patients. However, it was unclear whether the lesions were the cause of the dementia. Elucidation of the underlying mechanisms of Alzheimer's disease developed from several lines of inquiry into the neurophysiological correlates of the progressive behavioral (mainly cognitive) dysfunction observed in AD patients.

Originally described by Alzheimer in 1907, the distinguishing neuropathological markers of AD are intercellular senile plaques and intracellular neurofibrillary tangles (NFTs) (e.g., Mattson, 2004). The plaques are comprised of variable-length (38 to 42 residues) beta amyloid protein (Roher et al., 1986) and the tangles consist of hyperphosphorylated cytoskeletal tau protein (Grundke-Iqbal et al., 1986). These features are found throughout the cerebral cortex, amygdala, and hippocampus. Within these specifically vulnerable brain regions, synaptic disruption and neuronal cell loss occur (e.g., Stern et al., 2004), with concomitant neurochemical changes. The underlying mechanism for the pathogenesis of Alzheimer's disease emphasizes the role of beta amyloid protein (Hardy, 1997; Tanzi and Bertram, 2005; Hardy, 2006), identified as the core peptide in plaques (Glenner and Wong, 1984a, 1984b;

Masters et al., 1985). This protein appears highly conserved across species (Selkoe et al., 1987); antibodies to a partial peptide fragment (28 amino acids) of human amyloid cross-reacted with cerebrovascular deposits and neuritic plaques from aged mammals (dog, monkey, orangutan, and polar bear). It is through the accumulation of abnormal beta amyloid protein that Alzheimers disease is considered a protein misfolding disease (Cohen and Kelly, 2003; Hashimoto et al., 2003), sharing aggregation kinetics properties with prion diseases (e.g., scrapie, Creutzfeldt-Jakob disease, bovine spongiform encephalopathy) (Griffith, 1967; Come et al., 1993; Cohen, 1999; Hayashi et al., 2004).

With progressive loss of cortical and hippocampal neurons, concomitant brain atrophy is observed in AD brains. The cerebral cortical atrophy is hemispherically asymmetric, being more pronounced in the left hemisphere than in the right (Gee et al., 2003; Thompson et al., 2003), particularly in the anterior- and posterolateral-temporal and dorsolateral-prefrontal regions (Gee et al., 2003). The cerebellum, occipital region, and sensorimotor cortex are largely spared (Thompson et al., 2003). Interruption of connectivity between association cortices may arise from the loss of pyramidal neurons of the CA1, subiculum, and entorhinal cortex of the hippocampus (Morrison and Hof, 1997). Within the hippocampus, substantial loss of CA1 neurons is observed in mild to severe AD patients (Price et al., 2001). Loss of hippocampal neurons in the CA1 and subiculum are associated with the formation of neurofibrillary tangles (Rossler et al., 2002).

Beta amyloid is produced through a series of protease-mediated cleavages from amyloid precursor protein (APP; Goldgaber et al., 1987) into several isoforms of between 365 and 750 residues (Vassar et al., 1999). APP is a transmembrane protein (Kang et al., 1987), having a long extracellular amino-terminus and a shorter intracellular carboxy-terminus (Maccioni et al., 2001). APP is encoded by a gene

within the Down's syndrome region of chromosome 21 (Tanzi et al., 1987). Normally, alpha-secretase cleaves APP near the center of the beta amyloid domain, resulting in a non-amyloidogenic polypeptide fragment (Selkoe, 2001). However, two pathogenic forms of beta amyloid (Aβ), beta amyloid[1-40] (Aβ_{40}) and beta amyloid[1-42] (Aβ_{42}), result from cleavage of APP by beta-secretase at Met^{671} followed by gamma-secretase at Val^{711} or Ile^{713}, respectively (Citron et al., 1996; Bossy-Wetzel et al., 2004). Missense mutations encoded in APP near secretase cleavage sites lead to increased beta-amyloid production (Selkoe, 2001). Neither APP nor Aβ is specific for Alzheimers disease; amyloid precursor protein and beta amyloid are found in AD patients as well as non-AD individuals (Selkoe, 2001).

Regiospecific accumulation of Aβ_{42} results in diffuse plaques, which occur mainly within the association and limbic cortices (Roher et al., 2000). Monomeric beta amyloid form alpha-helices which undergo destabilization and conformal transformation into beta-helical dimers (Serpell, 2000), each consisting of a hydrophobic core surrounded by hydrophilic residues. Adjacent dimers bind together into protofilaments which, in turn, form beta-sheets, and subsequently coalesce into fibrils (Roher et al., 2000; Walsh and Selkoe, 2004). Compact (dense) plaques subsequently may form from additional conformal changes in fibril aggregates (e.g., Kayed et al., 2003), and trigger astrocytic and microglial activation (Yamaguchi, et al., 1988), which represent an inflammatory response (Selkoe, 2001). Subsequent production of free radicals by microglial mitochondria leads to neuronal dysfunction and cell death through oxidative damage, such as lipid peroxidation (Selkoe, 2001; Reddy, 2006). Astrocytes produce two inflammatory proteins – α-antichymotrypsin (ACT) and apolipoprotein E (ApoE) – which participate in amyloid plaque formation (Selkoe, 2001; Potter et al., 2001). Both ACT and ApoE have been shown to promote amyloid formation and deposition as "pathologic chaperones" both *in vitro* and *in vivo* (Sanan

et al., 1994; Wisniewski et al., 1994; Bales et al., 1999). Furthermore, it has been demonstrated that ACT and ApoE act either together or independently to promote both diffuse and compact beta amyloid plaques without influencing monomeric beta amyloid levels (Nilsson et al., 2004).

Intracellular NFTs are comprised of aggregates of paired helical filaments (PHF), alpha-helices of hyperphosphorylated microtubular tau protein (Alonso et al., 2001; Canevari et al., 2004; Sobow et al., 2004). Tau, which normally binds tubulin, is involved in microtubule formation and stabilization within neurons (Mudher and Lovestone, 2002). Following phosphorylation by protein kinases cdk5 or GSK3beta, tau dissociates from microtubules and forms PHFs (Maccioni et al., 2001). Finally, accumulating PHFs destabilize intraneuronal microtubules and eventually replace the microtubules with pathogenic tangles (Mudher and Lovestone, 2002).

1.4 Diagnosis of Alzheimer's Disease

Dementia – the progressive deterioration of cognitive function – may result directly from primary diseases of the brain, or as a consequence of other disease states. The most common types of dementia, distinguished by age of onset and specific pattern of structural and functional pathology, include: Alzheimer's disease, frontotemporal dementia, HIV-associated dementia, Lewy body dementia, and vascular dementia. Dementia is often comorbid with Huntington's disease, Parkinson's disease, progressive supranuclear palsy, neurosyphilis, and several prion disorders (e.g., Creutzfeld-Jakob disease, Gerstmann-Straussler-Scheinker syndrome). In addition, certain environmental toxins (e.g., lead, industrial solvents), metabolic disorders (e.g., hypothyroidism, vitamin B12 deficiency) and brain injuries (e.g., subdural hematoma, normal-pressure hydrocephalus) can produce dementia which, in some cases, may be reversible with treatment. Several psychiatric conditions, including

depression, may resemble dementia as well. Differential diagnosis of Alzheimer's disease, therefore, requires a complete personal and family medical history, and consideration of coexisting medical/psychiatric conditions which share symptoms in common with Alzheimer's disease. The complete diagnostic process involves several stages in increasing order of invasiveness.

General cognitive impairment appears gradually over time, and probable Alzheimer's disease is diagnosed through behavioral assessment, typically as clinical observation and interview-based examination, over a period of weeks or months. The overt symptoms of Alzheimer's disease are called the "Four As" – agnosia (the inability to recognize objects or familiar people), amnesia (the inability to remember), aphasia (linguistic incompetence; the inability to understand or communicate using language), and apraxia (the inability to perform activities). Physical and psychological examinations are necessary for differential diagnosis, to rule out alternative disorders with similar presentations, such as clinical depression. Interviews with the patient, family members, and caregivers also provide diagnostic information. In addition, electroencephalographic and brain imaging techniques are used to supplement diagnosis. Currently, the definitive diagnosis of Alzheimer's disease requires microscopic examination for the characteristic lesions at autopsy.

1.4.1 Psychometric Assessment

Evaluation of probable Alzheimer's disease in both clinical and community-living settings requires standardized psychometric testing instruments. Cognitive assessment inventories commonly used in clinical settings include the Mini-Mental State Examination (MMSE; Folstein et al., 1975) and the Blessed Orientation-Memory-Concentration test (BOMC; Katzman et al., 1983). The MMSE consists of a thirty-point questionnaire of items representing five cognitive components: time and place

orientation, registration (ability to identify and remember three common objects), attention and mental calculation (e.g., serial subtraction), short-term memory recall (of the three objects presented earlier), and language ability (e.g., repeat phrase, follow directions given by examiner). Scores are compared against norms established for age and educational level (Crum et al., 1993). The BOMC contains six items, addressing three cognitive domains: time orientation, concentration (analogous to the MMSE attention and mental calculation component), and memory recall. The similar content of these instruments is reflected in the strong intercorrelation between examinee scores (Fillenbaum et al., 1987; Zillmer et al., 1990). High-throughput screening techniques, abbreviated versions of mental test batteries, are under development for efficient, large-scale evaluation of prospective patient groups. Mnemonic measures, for instance, are useful for AD-screening as well as in differential diagnosis. For example, AD patients score lower than MCI patients who, in turn, score lower than age-matched depressed patients on visual association cued-recall tests (Dierckx et al., 2007), using discriminant analysis (95% overall accuracy). Semantic fluency tasks (e.g., free-response naming of exemplars within a target category) are particularly sensitive discriminators between normal aging and Alzheimer-associated deficits (Cerhan et al., 2002; Salmon et al., 2002). For example, in tests of verbal fluency (names of animals, words having same first-letter), normal-aged individuals generate more words and produce larger clusters within categories, relative to mild-AD patients (Gomez and White, 2006). In addition, comprehensive cognitive assessment protocols have been established (e.g., CERAD; Morris et al., 1993) for clinical evaluation, therapeutic monitoring, and epidemiological studies of AD patients. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) has developed a multimetric neuropsychological battery for diagnosing probable Alzheimer's disease, as well as for characterizing the severity of cognitive impairment (Morris

et al., 1993; Strauss and Fritsch, 2004), consisting of the MMSE and six additional performance measures: modified Boston Naming Test, verbal fluency test, constructional praxis (design copying), and three word-list learning measures (immediate recall, delayed recall, and word recognition). Factor analytic studies of the CERAD inventory (Strauss and Fritsch, 2004; Jones and Ayers, 2006) suggest a single factor representing overall cognitive performance in probable-Alzheimer's subjects, however a two-factor solution was obtained when age-matched, non-AD subjects are included in the analysis (Jones and Ayers, 2006). Indeed, researchers have extracted two- (e.g., Jacobs et al., 1994), three- (e.g., Kanne et al., 1998), and even five-factor (e.g., O'Donnell et al., 1988) solutions using multiple psychometric instruments – including the MMSE (Folstein et al., 1975), SPMSQ (Pfeiffer, 1975), and ACAD (Sevush et al., 1991); the apparent multidimensionality of cognition in Alzheimer's disease has been interpreted in terms of either premorbid heterogeneity or AD-related heterogeneity (Fisher et al., 1999; Sevush et al., 2003). Two hundred sixty-one patients with probable-AD completed the MMSE, SPMSQ, and ACAD examinations (Sevush et al., 2003); subsequent factor analysis (principal components, Varimax rotation) identified two general-cognitive factors, related to Alzheimer's progression, and a third factor reflecting premorbid characteristics (correlated with demographic variables). In addition, Alzheimer's disease has been considered a heterogeneous disorder, having diverse overt behavioral and cognitive manifestations (Cummings, 2000). The cognitive impairment associated with Alzheimer's disease tends to increase in both severity and scope, mirroring the insidious progression of the underlying neuropathology (Morris et al., 1989; Cummings, 2000). In contrast, various psychological and behavioral symptoms of Alzheimer's pathology may appear during different stages (Jost and Grossberg, 1996). The multidimensional character of Alzheimer's disease, for instance, was suggested by a factor analysis (principal component) of eighteen

standard behavioral/cognitive measures in an AD outpatient (DSM-IV, NINCDS-ADRDA criteria) sample (N=244; Spalletta et al., 2004), which identified seven distinct factors (and their associated behavioral vs. cognitive dimension): general cognitive (cognitive), hyperactivity (behavioral), psychosis (behavioral), constructive ability (cognitive), anxiety (behavioral), mood-excitement (behavioral), and mood-depression/apathy (behavioral). These results, collectively, suggest that behavioral and cognitive components represent independent dimensions of AD (Spalletta et al., 2004), based on multiple conventional neuropsychological assessment techniques, and underscore the primacy of a single underlying "global" construct of general cognition. Finally, although both psychometric (e.g., MMSE) and behavior-based (e.g., DAFS) evaluation protocols have demonstrated effectiveness for differential diagnosis and therapeutic assessment in Alzheimer's disease, interpreting the complex interplay between cognitive and behavioral processes – as revealed through factor analyses – remains a controversial area of research (e.g., Ownby et al., 2004).

Cooperative efforts to standardize clinical diagnosis of Alzheimer's disease, by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA), led to the development of a set of criteria based on clinical observation, neuropsychological testing, and histopathologic evidence (McKhann et al., 1984; Blacker et al., 1994). The presence of a dementia syndrome, in addition to progressive cognitive impairment, occurring in the absence of other diseases capable of producing dementia, is required for a diagnosis of probable Alzheimer's disease; post-mortem microscopic examination of brain tissue is required to confirm definite Alzheimer's disease. Recent advances in diagnostic imaging techniques, however, have prompted the revision of the NINCDS-ADRDA criteria, and a new proposal is under review (Dubois et al., 2007). Finally, neuropsychological tests are modest predictors of daily functional sta-

tus in Alzheimer's patients. In a study (Farias et al., 2003) involving 42 outpatients diagnosed with possible- or probable-AD, cognitive performance (immediate- and delayed-recall, attention, visuospatial, executive functioning, and praxis measures) was significantly correlated with functional status (DAFS scale; Loewenstein et al., 1989). The DAFS scale (Loewenstein et al., 1989, 1995) is a highly-reliable (both inter-rater and test-retest) behavior-based evaluation of daily living skills in familiar domains, such as: telling time (reading a clock), handling money (writing a check, balancing a checkbook, making change), and household routines (addressing and mailing a letter, using a telephone). The DAFS is a useful screening device for differential diagnosis among normal-aged, depressed, and probable-AD individuals; AD patients score significantly lower than either normal-aged or depressed individuals on all functional measures except telling time (Loewenstein et al., 1989).

1.4.2 Biomarkers

Advances in bioinformatics methodology enable investigators to examine gene-level (genomic) and protein-level (proteomic) features of normal and pathological phenotypes, including neurological syndromes (Mirnics and Pevsner, 2004). Manipulations of genetic expression in transgenic animals, for instance, which alter neuronal structure and/or molecular signalling pathways are manifested through distinct cognitive genotypes (Flint, 1999). Candidate-gene association studies and genome-wide scans suggest that human memory is a polygenic cognitive trait (de Quervain et al., 2003; Egan et al., 2003; Papassotiropoulos et al., 2005; de Quervain and Papassotiropoulos, 2006), with heritability estimated at 50% (McClearn et al., 1997). For example, a genome-wide scan (Papassotiropoulos et al., 2006) involving over 500,000 single-nucleotide polymorphisms in pooled DNA from 341 young adults, stratified by performance in a verbal delayed-recall task, linked better performance (after both

5-min delay and 24-hr delay) with the presence of the KIBRA rs17070145 T allele; an independent sample of 256 adults associated this allele with superior performance on standard tests of episodic memory (Buschke's Selective Reminding Test, Rey Auditory Verbal Learning Test), although no significant allele-dependent differences were observed in the Wisconsin Card Sorting Test or the Paced Auditory Serial Attention Task (Papassotiropoulos et al., 2006). Hence, a single genetic mutation (KIBRA rs17070145) is associated with individual differences in episodic memory, but not attention, executive function, or working memory (Papassotiropoulos et al., 2006). Furthermore, gene-associated differences in brain function (e.g., memory, intelligence) are partly determined by gene-influenced structural differences in the brain (Toga and Thompson, 2005; Zimmer, 2008). Additionally, genetic evidence from animal studies underscores the association between genes and cognitive ability. Genetic differences in spatial learning ability in mice has been examined using multitrait analysis (Steinberger et al., 2003) in DBA/2 and C57BL6/J mouse strains, and reveals two quantitative trait loci, on chromosomes 4 and 12, are associated with Morris water maze retention (time spent and number of annular crossings in former platform-containing quadrant), as well as spatial learning rate (swim path distance, latency to reach platform location), measured in the probe trial (Steinberger et al., 2003).

Techniques based on data mining (e.g., Bayesian statistical models) have been applied to human brain-specific gene network databases compiled from genomic and proteomic studies in order to identify candidate genes for neurological disorders (Liu et al., 2006). Genome screening in known-AD families, for example, has identified mutations in three genes associated with the early-onset familial AD (APP, PSEN1, and PSEN2; e.g., Tanzi and Bertram, 2001) and a late-onset AD-associated polymorphism of ApoE (ApoE-ϵ4; chromosome 19q13). In addition, linkage analysis

suggests as many as twelve genetic loci contribute to AD (1q23, 3p26, 4q32, 5p14, 6p21, 6q27, 9q22, 10q24, 11q25, 14q22, 15q26, and 21q22; Blacker et al., 2003). Microarray analysis of plasma samples from normal-control and AD-individuals identified eighteen proteins which differ significantly in concentration (i.e., ANG-2, CCL5, CCL7, CCL15, CCL18, CXCL8, EGF, G-CSF, GDNF, ICAM-1, IGFBP-6, IL-1α, IL-3, IL-11, M-CSF, PDGF-BB, TNF-α, and TRAIL-R4), and distinguish between the groups with approximately 90% accuracy (Ray et al., 2007). These signaling proteins are associated through two independent regulatory pathways: one set is related by TNF-α (tumor necrosis factor) and M-CSF (monocyte-colony stimulating factor), while the other set centers around EGF (epidermal growth factor) (Ray et al., 2007). The involvement of these proteins in immunoreactivity, hematopoiesis, and apoptosis is consistent with recent findings showing increased hematopoietic cell activity in the Alzheimer brain (both human AD and in Alzheimer's transgenic mice; (Wyss-Coray, 2006; Simard et al., 2006; Britschgi and Wyss-Coray, 2007), as well as Alzheimer-associated abnormalities in apoptotic pathways (LeBlanc, 2005). The diagnostic utility of microarray analysis is underscored by postmortem frontal cortical RNA analyses in normal-aged and AD patients (Walker et al., 2004). A data mining-based classifier (BioMiner^{TM}) confirmed the expression of 17 neuropathology-associated genes (DTNA, B2M, APLP1, C4B, LIMS2, IGHM, GRP58, KRT8, ATF4, RANGAP, GSTM2, TU3A, ADD3, FTL, HBB, CLU, and HBG2) in cortical samples from AD patients (Walker et al., 2004).

Similarly, gene profiling (cDNA microarray) of 18-month-old Alzheimer's transgenic mice (APPsw) and age-matched control animals has identified 52 differentially expressed genes (Jee et al., 2005). Subsequent examination of age-related patterns in genetic up- and down-regulation, comparing aged (18 month-old) and young (1 month-old) transgenic and normal mice, revealed 48 and 40 differentially expressed

genes between aged and young animals of each group, respectively (Jee et al., 2007). Age-associated changes in the genetic profiles are consistent with inflammatory response, increased oxidative stress, and decreased neurotrophic support (Lee et al., 2000). Differential 2D electrophoresis of whole-cortical tissue from 14-month-old Alzheimer's transgenic mice (APP/PS1) identified significant differences in the concentrations of 15 proteins, relative to wild-type animals (Sizova et al., 2007). In addition to intraneuronal and membrane proteins associated with synaptic function and axonal growth, proteins involved in glial response and inflammation, cholesterol metabolism, and oxidative response were included (Sizova et al., 2007).

1.4.3 Neurological Basis of Memory Systems

Anatomical components of the mammalian nervous system form a complex, interconnected network. Understanding neuropathologies, such as Alzheimer's disease, requires thorough study of the relationship between structure and function at different scales. Indeed, memory – the storage and retrieval of learned information – has been examined and modeled in humans (and other animals) on different scales, within both normal and pathological states, and in a variety of contexts. On a cellular scale, for example, simultaneous stimulation of two neurons enhances synaptic transmission between the neurons through a complex cascade of intracellular protein transcriptional and molecular signalling events (Kandel, 1979; Sweatt, 1999; Kandel, 2001), consequently improving long-term interneuronal communication ("long-term potentiation," LTP; Lynch, 2004). Direct-stimulation studies, for example, show that short-term memory formation is dependent upon LTP in the hippocampus (e.g., Bliss and Lomo, 1973), and pharmacologic agents which interfere with hippocampal LTP (e.g., by blocking NMDA receptors) in rats also impair spatial memory performance (Morris water maze) (Morris et al., 1986). In addition, mice receiv-

ing medial frontal cortical injections of anisomycin (a protein synthesis inhibitor) exhibit spatial memory acquisition deficits (radial maze nonmatching-to-place task) (Touzani et al., 2007), suggesting the role of protein synthesis during learning. A macroscale, information processing model of memory uses storage duration as the distinguishing criterion for a tripartite organization of memory systems, consisting of: sensory memory (registry of ambient environmental stimuli, typically available for a very brief duration; Sperling, 1960); short term memory (limited-capacity temporary store, Baddeley and Hitch's "working memory," 1974; Miller, 1956; Brown, 1958; Peterson and Peterson, 1959); and, long term memory (high-capacity, "permanent" reference memory; Landauer, 1986). Long term memory is subsequently divided into declarative (explicit) memory and nondeclarative (implicit) memory components, each having distinct neuroanatomical substrates (e.g., Schacter, 1987; Thompson and Kim, 1996; Squire and Zola, 1997; Miyashita, 2004). Indeed, much of what is known about the structure of human memory systems is inferred from clinical studies of amnesics (e.g., Shallice and Warrington, 1970; Cohen et al., 1985; Graf and Schacter, 1987; Periani et al., 1993), who exhibit dissociation syndromes among component systems. The inability to form new long-term memories through experience (anterograde amnesia), for example, is associated with hippocampal lesions or temporal lobe damage (either deliberately by surgery, or accidentally through injury) (e.g., Kandel and Pittenger, 1999). Declarative memory involves medial temporal lobe, hippocampal, and prefrontal cortical structures (Kandel and Pittenger, 1999), and consists of factual information (semantic memory) and autobiographical, experience-based knowledge (episodic memory); the semantic-episodic distinction was suggested by Tulving (1972, 1985). By contrast, nondeclarative memory involves the cerebellum and striatum, and encompasses procedural, skill-based knowledge, including conditioning; associative priming (Meyer and Schvaneveldt, 1971), artificial gram-

mar learning (correct identification of rule-consistent strings of letters formed from a finite-state grammar; Reber, 1967), category learning, and sequence learning (e.g., serial reaction time task) are examples of implicit memory phenomena.

Within a comprehensive assessment of learning and memory function, domain-specific memory tasks provide evidence for the differential effects of normal and pathological (i.e., AD) aging on human memory systems. For example, comprehensive memory tests administered to early-to-moderate AD patients and age-matched normal individuals (Stopford et al., 2007) identified five distinct memory domains (personal memory, recognition, verbal recall, visual recall, and working memory) in which impairment within one domain may occur independently of the other domains. In a comparison of differential impairment across memory systems, Mitchell (1989) examined the effect of normal aging in two adult cohorts (cross-sectional study; 19-32 y/o "young" and 63-80 y/o "older"; N = 48, each) using multiple performance measures: episodic memory (free recall, recognition, intrusions during recall, and longitudinal change in recall; Underwood et al., 1978); procedural memory (repetition priming latency; Graf and Schacter, 1985); and, semantic memory (vocabulary, picture-naming latency and errors). Although no significant age-related differences were observed for either procedural or semantic memory, younger adults performed better in recall and recognition measures of episodic memory (Mitchell, 1989); additionally, factor analysis of the multimetric assessment confirmed the tripartite memory taxonomy. Episodic memory (free recall, word recognition) deficits have been observed in preclinical AD patients (Backman et al., 2001), up to six years prior to diagnosis, as well as deficits in perceptual speed and executive functioning (Backman et al., 2005). No short-term memory (digit span) impairments were found in preclinical AD cases, relative to age-matched normal individuals, however (Backman et al., 2001).

1.4.4 Electroencephalography

A landmark discovery in neurology, the construction of a minimally-invasive electrical interface between the living nervous system and a recording apparatus, enabled investigators to monitor ensemble neural activity in real-time. Electroencephalography (EEG), first described by Hans Berger in 1924, involves recording electrical potentials from an array of electrodes distributed across the scalp surface following a standardized grid pattern (Rowan and Tolunsky, 2003). The potentials recorded from each electrode reflect the collective electrical activity (i.e., APs, EPSPs and IPSPs) of the underlying local population of neurons, to a depth of several centimeters. The EEG, therefore, represents regional brain electrical potential (e.g., microvolts) as a function of time (e.g., milliseconds). Four characteristic signal patterns have been described in the human EEG, somewhat arbitrarily based on frequency ranges, which correspond to different mental states. Beta activity ("fast-wave"), indicative of focused mental activity and attention, has relatively high frequency (greater than 12 Hz) and low amplitude. Alpha activity, recorded in relaxed, awake subjects with their eyes closed, exhibits a frequency of 8 to 12 Hz with amplitude between 10 and 50 microvolts. The prominence of the alpha rhythm in the occipital region, and its responsiveness to opening and closing of the eyes, suggests an association with visual processing (Adrian, 1935). Theta waves of between 4 and 7 Hz, and delta ("slow-wave") activity, of less than 4 Hz frequency, are typical of drowsiness and sleep. Visual inspection is the conventional method for interpreting EEGs recorded from subjects. The presence (or absence) of specific features or patterns of activity, as depicted on a graphical recording trace, represents neurological diagnostic criteria. Human EEGs recorded during a broad range of behavioral states – both normal and pathological – have been examined in this manner. For example, in-

creased frontal midline theta activity is associated with working-memory intense tasks (Gevins et al., 1997, 1998). Sleep-staging, for instance, traditionally relies on visual recognition (by an expert) of unique waveforms and sustained periods of slow-wave activity in the EEG trace (e.g., Rechtschaffen and Kales, 1968). An alternative, computational approach, nonlinear dynamical analysis (NDA), utilizes mathematical techniques adapted from engineering for studying complex, time-based physical phenomena (Pradhan and Dutt, 1993; Pritchard and Duke, 1995). Using this approach, the dynamics of the EEG signal are characterized mathematically in terms of "dimensional complexity" (represented as the correlation dimension, D2; Grassberger and Procaccia, 1983). Dimensional complexity is a summary statistic, capturing the overall character of cortical activity (incl. nonlinear communication between individual neurons, as well as between neuron populations) in a single measure. It is noteworthy that the presence of dynamically chaotic activity throughout the brain may be essential for – and indicative of – normal functioning, including consciousness (Fell et al., 2003). The spatiotemporal complexity of the EEG mirrors the intricate neurophysiological substrate for behavioral and cognitive processing. Indeed, within the nervous system, synchrony and coherence are often associated with pathology (e.g., epileptiform spiking and seizures) (Skarda and Freeman, 1987; Doyon, 1992).

In addition to providing theoretical insights on spatiotemporal phenomena in the brain, dimensional complexity measures can be used to describe cortical states (for identifying neural correlates of behavior and cognitive activity), to examine therapeutic efficacy and experimental treatment effects, and to assist in the diagnosis of neuropathology. For example, levels of consciousness can be reliably distinguished through NDA techniques, facilitating computer-assisted automated EEG monitoring and recording. In humans, as sleep proceeds from NREM ("non-REM") Stage 1 through NREM Stage 4, a monotonic decrease in measured dimensional

complexity (D2) has been observed (Pradhan and Sadasivan, 1996). In addition, D2 measures during REM are slightly higher than D2 during waking. Automated discrimination among NREM stages using D2 measures from sampled EEG data is comparable or superior to expert-staged hypnograms (Pradhan and Sadasivan, 1996). Lower D2 measures, relative to resting conditions, are observed during meditative states (Aftanas and Golocheikine, 2002) particularly over the midline frontal and central regions. Sleep-deprived subjects also exhibit lower D2, which may represent sub-optimal information processing capacity (Jeong et al., 2001b). In addition, task-dependent changes in D2 may be related to individual intellectual ability (e.g., Jausovec and Jausovec, 2000; Lutzenberger et al., 1992; Molle et al., 1999), consistent with a resource-demand model of functional processing in the brain. NLDA-based studies of neuropathologies, including Alzheimer's disease, suggest that dimensional complexity may be a more sensitive diagnostic marker than traditional expert-based EEG interpretive methods (e.g., Gallez and Babloyantz, 1991; Jeong et al., 1998; Jeong, 2004). Decreased D2 was found in clinical patients with posttraumatic stress disorder (PTSD), suggestive of impaired cortical information processing (Chae et al., 2004). Patients diagnosed with schizophrenia do not exhibit initial transient D2 changes, compared with non-schizophrenic individuals, when performing a cognitively-demanding task (Kirsch et al., 2000), suggesting impaired ability to respond adaptively to changing cognitive demands.

The primary EEG findings in AD patients include: increased diffuse slow activity and slowing of the dominant posterior rhythm (Brenner et al., 1988; Jeong, 2004), decreased alpha and beta activity (Letemendia and Pampiglione, 1958; Jeong, 2004), increased theta and delta activity (Brenner et al., 1986; Giaquinto and Nolfe, 1986), and decreased synchronization (coherence) in the alpha and beta bands (Dunkin et al., 1994; Koenig et al., 2005). These abnormalities correlate with disease severity

(Hughes et al., 1989; Kowalski et al., 2001), and reflect concomitant structural and functional impairment of the cerebral cortex. Underscoring cholinergic involvement in Alzheimers disease is the finding that young, healthy subjects receiving the muscarinic (cholinergic) antagonist scopolamine also exhibit either/both increased slow-wave and decreased fast-wave activity (Ebert and Kirch, 1998; Ebert et al., 2001) and transitory Alzheimer-like memory impairment (Wesnes et al., 1988). In addition, AD patients receiving the cholinergic agonist nicotine show increased fast-wave activity and decreased slow-wave activity (Knott et al., 2000), and mild-AD patients receiving long-term treatment with donepezil show decreased theta activity (Kogan et al., 2001). NLDA-based studies, using single-channel as well as multi-channel EEG, show decreased D2 in AD, suggesting reduced nonlinear cellular communication or coupling between cortical regions (Jelles et al., 1999; Jeong et al., 2001a).

1.4.5 Diagnostic Medical Imaging

Medical imaging techniques provide useful diagnostic information. These methods include: computed tomography (CT), magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and single photon emission computed tomography (SPECT). For example, anatomical (structural) changes in the brain are detected using CT and MRI, while physiological (metabolic) correlates of brain function are studied using PET and SPECT. Regiospecific atrophy of the hippocampus and entorhinal cortex have been identified in longitudinal studies of Alzheimers patients using CT and MRI (Jack et al., 1997; Xu et al., 2000; Barnes et al., 2004; Pennanen et al. 2004) underscoring the utility of these methods for monitoring the progress of AD in patients over time. Localized decreases in glucose metabolism, measured through PET studies, can identify areas compromised by Alzheimers disease (Minoshima, 2003). Modified PET techniques,

using a radiolabeled ligand which binds to beta amyloid, have been used to localize amyloid plaques (Klunk et al., 2003; Klunk et al., 2004). Correlation between *in vivo* and *in vitro* PET imaging studies using the radiotracer Pittsburgh Compound-B (Ikonomovic et al., 2008), for instance, show highly selective binding for insoluble (fibrillar) beta amyloid deposits. A similar approach is being developed for PET imaging of neurofibrillary tangles (Mathis et al., 2004). Additionally, a refinement of SPECT (HMPAO-^{99}Tc SPECT; DeFigueiredo et al., 1995) has demonstrated utility for differential diagnosis of AD from both vascular- and fronto-temporal dementias (Dougall et al., 2004), with 71.3% (75.9%) and 71.5% (78.2%) sensitivities (and specificities), respectively.

Early perceptual and information processing events, and their neural substrates, are examined through diagnostic medical imaging techniques. Moreover, diagnostic imaging during task performance helps reveal the neuroanatomical substrates, connectivity, and mechanisms which underlie complex cognitive behavior. For example, the Stroop color word interference task, in which subjects identify congruence between color-word (i.e., name of a color) and word-color (i.e., color of a printed word) stimuli presented visually (Stroop, 1935), requires coordinated visual and verbal processing; AD individuals exhibit deficits in this task, relative to age-matched normal subjects (Bondi et al., 2002; Amieva et al., 2004). The neural circuitry involved in the Stroop color word interference task has been explored using fMRI (Peterson et al., 1999; Leung et al., 2000). Regiospecific activation patterns within the frontal and anterior cingulate cortex follow a temporal sequence correlated with visuospatial and verbal processing of the target stimuli (Leung et al., 2000). Left prefrontal cortical activation was identified during verbal fluency tasks using fMRI in normal individuals (Schlosser et al., 1998), as well. Additionally, neuroimaging can reveal parallels between nascent behavioral (functional) impairment and neuropathology, and sug-

gest patterns of functional compensation during early AD (e.g., Posner et al., 1997). For example, age-matched normal and early-AD individuals performed an episodic memory task while undergoing functional magnetic resonance (fMRI) imaging (Gron and Riepe, 2004); in a graded progression, increasing functional impairment (poorer cognitive performance) was associated with decreasing posteromedio-temporal activation, across normal and AD subjects. Intact operational components of working memory – executive control (information encoding and retrieval processes) and active maintenance (preserving immediate availability of information) – are necessary for short-term storage and manipulation of information (e.g., Baddeley and Hitch, 1974; Baddeley, 1992). Functional MRI studies in humans performing working memory tasks indicate both frontal and parietal cortical involvement in the active maintenance component (Cohen et al., 1997). Positron emission tomography (PET) studies of human subjects during episodic memory tasks (yes-no recognition) reveal increased activity within the right prefrontal cortex and anterior cingulate (Nyberg et al., 2000). In addition, PET during an episodic memory task involving normal-aged and mildly cognitively-impaired (MCI) patients showed different patterns of regional activation between the two groups (Moulin et al., 2007); the task consisted of two successive encoding trials (semantically-related pairs of words) followed by a retrieval trial. Although both groups exhibited left frontal lobe activation during initial encoding, only the normal-aged individuals showed activation during the secondary encoding. In addition, MCI individuals showed activation in the visual cortex during secondary encoding, while normal-aged individuals did not. During retrieval, MCI individuals did not exhibit right frontal or left temporal activation (in contrast to normal-aged individuals), although they showed more extensive left frontal cortical activation, relative to normal-aged individuals (Moulin et al., 2007).

An important goal of medical research in Alzheimer's disease is the development of early diagnostic techniques, to optimize the efficacy of treatment and intervention. Methods which are both sensitive to early behavioral and pathological indications and selective for Alzheimer's disease are highly desirable. Content analysis, for instance, represents a potential early predictor, providing a means for identifying the progressive decline of linguistic abilities associated with Alzheimer's disease (Forbes et al., 2004; Garrard et al., 2005; Venneri et al., 2005). Careful examination of an author's written work (e.g., the final novel of British writer Iris Murdoch) completed prior to the onset of characteristic Alzheimer symptoms can reveal lexical deficits suggestive of incipient cognitive impairment (Garrard et al., 2005). However, whether semantic and syntactic components of verbal ability decline separately or in parallel remains unclear (Garrard et al., 2005). Alzheimer's disease, like many other nucleating pathologies (e.g., prion diseases; Prusiner, 1991; Jarrett and Lansbury, 1993; DeMager et al., 2002), exhibits very subtle manifestations in its nascent form; prodromal detection of AD is complicated because neuropathology precedes behavioral deficits.

1.5 Risk Factors for Alzheimer's Disease

Aging and family history have been identified as major risk factors for Alzheimer's disease. However, Alzheimer's disease is not an inevitable consequence of aging. For example, slight atrophy and sparse accumulations of hyperphosphorylated tau, but no beta-amyloid plaques, were found in the cortex and hippocampus of a 115-year-old woman whose cognitive abilities remained intact (MMSE scores comparable to 60-75 y/o) until her death from cancer (den Dunnen et al., 2008). Although the cause of Alzheimer's disease is unknown, it is likely a multifactorial disorder possessing a combination of dispositional (e.g., genetic) and situational (e.g., environmental)

determinants (e.g., Mattson et al., 2002). Two variants of Alzheimer's disease have been identified: a late-onset (i.e., 65 years of age, or older) "sporadic" form, which encompasses the majority of cases, and an early-onset (between 30 and 65 years of age) "familial" form, which accounts for approximately five percent of cases. These variants share similar behavioral and pathological manifestations, and differ only in their age of onset.

Several inalterable factors contribute to the development of Alzheimer's disease. The likelihood of developing Alzheimer's disease increases with age. Females are at higher risk than are males. The presence of the ApoE4 allele of the ApoE gene increases risk in a dose-dependent manner, and lowers the age of onset of AD (Veurink et al., 2003; Yao et al., 2004). In addition, the early-onset variant of Alzheimer's disease is associated with mutations in several autosomal-dominant genes (Selkoe and Podlisny, 2002; Pardo and van Duijn, 2005; Chai, 2007).

Life experiences and lifestyles can influence both the likelihood of developing Alzheimer's disease as well as the progress of the disorder. A history of traumatic head injury, particularly during early adulthood, is associated with increased risk (Fleminger et al., 2003). Poor cardiovascular health (high blood pressure, high cholesterol) is also associated with increased risk. A longitudinal study (Buchman et al., 2006) examined the relationship between body mass index and several age-associated dementing neuropathologies (AD, cerebral infarction, and Lewy body disease), and showed a significant association with Alzheimer's pathology even after correcting for dementia, chronic diseases, and physical activity. The relative risk of incident dementia was estimated (Cox proportional hazard model) in a cohort study (Stern et al., 1994; N=593 non-demented individuals, aged 60+ y/o), in which increased risk of dementia was associated with low education (RR 2.2), low lifetime occupational attainment (RR 2.25), or both (RR 2.87); the reduced risk of inci-

dent AD associated with increased educational and/or occupational attainment may reflect either the limitations of diagnostic testing (education influences test performance) or an acquired cognitive reserve which delays (or masks) the onset of clinical dementia (Stern et al., 1994). For reasons which remain unclear, higher educational attainment is also associated with accelerated cognitive decline with advancing age in AD (Wilson et al., 2004), particularly in the executive speed and memory cognitive domains (Scarmeas et al., 2006a). Environmental toxins may also play a role, particularly during early development. Relatively high concentrations of copper, for example, have been found associated with human amyloid plaques and neuropil (Kowalik-Jankowska et al., 2002). Additionally, aged (23 y/o) monkeys exposed to lead as infants exhibit greater expression of APP and BACE1, as well as beta amyloid deposits in the frontal association cortex (Wu et al., 2008b); concomitant decrease in DNA methyltransferase activity is suggestive of an epigenetic imprinting mechanism.

1.6 Risk Reduction Strategies

Beneficial health-promoting lifestyles, including increased levels of physical activity (Larson et al., 2006), mental stimulation (e.g., reading, puzzle-solving; Verghese et al., 2003), and social interaction are associated with decreased risk for developing Alzheimer's disease, and may have long-term protective benefits (Fratiglioni et al., 2004; Kramer and Erickson, 2007). Elderly (aged 65+ years) individuals who engage in regular (i.e., at least three times per week) physical exercise (e.g., walking, bicycling, swimming) are significantly less likely to develop dementia, compared to individuals who exercise less frequently (Larson et al., 2006). Participants (N=176; 70+ year-olds) in a year-long group investigative program who engaged in moderate daily exercise reported higher subjective measures of quality-of-life, physical well-being, and physical self-perception, relative to more-sedentary participants (Fox et

al., 2007). Interestingly, although voluntary participation in leisure physical activities is associated with reduced AD risk, neither occupational (at-work) nor commuting (to/from workplace) physical activity is associated with reduced risk for either dementia or AD (Rovio et al., 2007). A five-year study (Wilson et al., 2007) involving 700 aged individuals examined the relationship between daily cognitive activity level and AD risk, and reported that more frequent engagement in cognitive activity was associated with reduced AD incidence, and that a cognitively-inactive person was 2.6 times more likely to develop AD.

Neuroprotective and/or therapeutic effects of dietary modification have been demonstrated in both laboratory (using Alzheimer's transgenic animal models) and clinical (human) studies. A "Mediterranean diet," low in saturated fat and enriched with fruits and vegetables, is associated with reduced risk for Alzheimer's disease (Scarmeas et al., 2006b). In addition, specific dietary micronutrient components can influence AD risk. The increased intake of vitamins C and E is associated with decreased AD risk (Masaki et al., 2000; Morris et al., 1998); the antioxidant properties of these vitamins help protect neurons from membrane lipid peroxidation-induced damage (Berman and Brodaty, 2004). Polyphenol components (e.g., phenolic acid, flavonoids) of phytochemical extracts from certain spices and fruits (incl. blueberries and pomegranates) also contain antioxidants having beneficial effects (Joseph et al., 1998; Aggarwal and Shishodia, 2004; Joseph et al., 2005; Hartman et al., 2006). Diets richer in sources of omega-3 fatty acids (e.g., DHA, docosahexaenoic acid), such as fish, are also associated with decreased AD risk (Bourre, 2004; Morris et al., 2005). Regular intake of alcohol and nicotine are also associated with decreased AD risk, although the potential risks from chronic use or abuse of these agents is under investigation (e.g.,. Letenneur et al., 2004).

Regular consumption of coffee, tea, or other caffeine-containing products may protect against the development of Alzheimer's disease and help maintain cognitive functioning throughout life. Caffeine is a non-selective (A_1, A_{2A}) adenosine receptor antagonist which increases alertness and arousal (e.g., Nehlig et al., 1992; Fredholm et al., 1999). In a Portuguese retrospective study (20-year), daily caffeine intake was significantly associated with reduced risk of AD (Maia and DeMendonca, 2002). A Canadian prospective study in aged adults showed that coffee consumption is associated with reduced AD risk (Lindsay et al., 2002). A cross-sectional study in aged adults (N = 890 female, mean age 72.6 yrs; N = 638 male, mean age 73.3 yrs) examined the relation between mental ability on standardized tests and self-reported coffee consumption, and showed a significant association between lifetime coffee consumption and better cognitive performance in women (Johnson-Kozlow et al., 2002). No significant relation between coffee intake and cognitive performance was found in men, nor between decaffeinated coffee consumption and cognitive performance by either sex (Johnson-Kozlow et al., 2002). By contrast, a study (Van Gelder et al., 2007) involving 676 European men, all born between 1900 and 1920, used MMSE scores in a mixed longitudinal model to explore the association between coffee consumption (in cups/day) and ten-year cognitive decline. A significant difference in MMSE-score decline was observed between coffee-consumers versus non-consumers, with the least cognitive decline occurring in individuals consuming three cups per day (Van Gelder et al., 2007). The optimal orally-administered dose of caffeine in healthy adult humans for enhancing both sensorimotor-cognitive (attentional, performance) and subjective (mood) experience is approximately 250 mg (Kaplan et al., 1997). Studies in Alzheimer's transgenic mice (APPsw) strongly support the putative association between caffeine consumption and both cognitive-protection and AD-amelioration. Arendash et al. (2006) administered caffeine (1.5

mg/day p.o.; equivalent to 500 mg/day in human) to APPsw mice for six months, beginning at four months of age. Caffeine-treated animals performed superior to age-matched transgenic controls, and comparably to non-transgenic animals, across multiple measures of spatial learning, working and reference memory, and recognition/identification (Arendash et al., 2006). Significantly decreased hippocampal Aβ was found in caffeine-treated transgenic animals, as well (Arendash et al., 2006). In addition, although dietary caffeine supplementation did not affect either cortical nor hippocampal adenosine receptor densities, caffeine-treated animals exhibited normal brain adenosine levels, which may explain the observed cognitive-protection effect (Arendash et al., 2006). Additional studies (e.g., Dall'Igna et al., 2007) suggest the neuroprotective mechanism of caffeine involves blockade of A_{2A} receptors in the brain.

Chronic intake of non-steroidal anti-inflammatory drugs (NSAIDs; e.g., ibuprofen, indomethacin, sulindac) is associated with a decreased AD risk (McGeer et al., 1996; Stewart et al., 1997; Zandi et al., 2002). The primary anti-inflammatory mechanism of these agents involves inhibition of the two isoforms of cyclooxygenase (COX1 and COX2) (Marnett and Kalgutkar, 1999), however, NSAID administration reduces Aβ_{42} levels independently of COX inhibition (Weggen et al., 2001). Studies in Alzheimers transgenic mice provide a more complex portrait, however. Although ibuprofen, flurbiprofen, indomethacin, and sulindac reduce brain Aβ_{42} levels in mice, many other NSAIDs (including aspirin, naproxen, and ketoprofen) do not (Eriksen et al., 2003). Nevertheless, modulation of the endogenous inflammatory response (e.g., by suppressing microglial activity) remains an active area of pharmaceutical research (e.g., Kitzawa et al., 2004; McGeer and McGeer, 2007).

1.7 Genetics of Alzheimer's Disease

The early-onset "familial" variant of Alzheimer's disease results from mutations within any of three autosomal-dominant genes (Price and Sisodia, 1998; Rocchi et al., 2003; Brouwers et al., 2006): amyloid precursor protein (APP, chromosome 21; Murrell et al., 1991; Mullan et al., 1992), presenilin-1 (PS1, chromosome 14; Haas et al., 1999), and presenilin-2 (PS2, chromosome 1). As described earlier, the polymorphic apolipoprotein E gene (ApoE, chromosome 19) confers differential vulnerability to the more-common, late-onset, "sporadic" form of AD. Additional genes, identified through microarray analyses of genome-wide screens, are currently under investigation to elucidate their roles in Alzheimer's disease. In addition, the association between environmental factors operating through epigenetic mechanisms and late-onset AD has received recent attention (e.g., Wu et al., 2008a).

The amyloid precursor protein (APP) gene is located on chromosome 21 (locus 21q21.3) and contains 19 exons (Kang et al., 1987; St George-Hyslop et al., 1987; Tanzi et al., 1988). Multiple isoforms of the protein are caused by alternative splicing (Yoshikai, 1990). In addition, several mutations in the APP gene have been identified in families through pedigree analysis. One variant of early-onset AD involves a valine-to-phenylalanine substitution at residue 717 (V717F; Murrell et al., 1991). This mutation, known as the "London" variant, increases cleavage of APP by gamma-secretase, resulting in higher levels of Aβ_{42}. Another variant of familial AD, known as the "Swedish" variant (Mullan et al., 1992), involves the double mutation K670N (i.e., Lys to Asn) and M671L (Met to Leu). This double point-mutation increases cleavage of APP by beta-secretase, resulting in elevated levels of both Aβ_{40} and Aβ_{42}.

The presenilin-1 (PS1) gene was localized to chromosome 14 (locus 14q24.3) through a pedigree analysis of 34 individuals with early-onset AD (Van Broeck-

hoven, 1992; Campion, 1995; Alzheimer's Disease Collaborative Group, 1995), and specifically described by Sherrington et al. (1995). The PS1 gene consists of 14 exons and the coding region is estimated at 60 kb (Rogaev, 1997; Del-Favero, 1999). Presenilin-1 is an integral membrane protein that cleaves Notch1 (Ikeuchi, 2002), and may play a role during embryonic somitogenesis (Koizumi, 2001). Several pathogenic allelic variants have been found, including M146L (Met to Leu) (Sherrington et al., 1995; Morelli et al., 1998), M146V (Met to Val) (ADCG, 1995) and M146I (Met to Ile) (Jorgensen et al., 1996; Gustafson et al., 1998). In addition, the mutation E318G (Glu to Gly) results in a predisposition to FAD (Taddei, 2002). Mutations of PS1 decrease the age of AD onset to the early 40s and 50s by selectively increasing gamma-secretase cleavage of peptides C99 and C83, which subsequently increases brain Aβ_{42} levels (Selkoe, 2001).

The presenilin-2 (PS2) gene is located on chromosome 1 (locus 1q31-q42). The PS2 gene contains 12 exons of which 10 are coding regions. The 448-residue polypeptide encoded by the primary transcript of PS2 shares 67% homology with PS1 (Levy-Lahad, 1996). Both PS1 and PS2 genes share structural and functional similarity, and their associated proteins are expressed in the same regions within mammalian neurons (Kovacs, 1996). Like PS1, several pathogenic allelic variants have been identified, including N131I (Asn to Ile), M239V (Met to Val) (Rogaev, 1995), and D439A (Asp to Ala) (Lleo, 2001). Tomita (1997) has suggested that this mutation results in increased amyloid plaque formation through altered APP metabolism.

The apolipoprotein E (ApoE) gene is located on chromosome 19 (locus 19q13.2). ApoE is polymorphic, and allelic variants of ApoE confer differential predisposition to human aging-associated cognitive disorders, including memory impairment and AD. The presence of the ApoE4 variant is associated with a more rapid decline in episodic memory, relative to ApoE2 and ApoE3 (Wilson et al., 2002). The ApoE4 variant

may be associated with increased risk for both AD (Corder et al., 1993; Gureje et al., 2006) and Creutzfeldt-Jakob disease (Amouyel, 1994), in a gene dose-dependent manner (i.e., more copies of the gene confers greater risk). Interaction effects between ApoE and cellular-level microenvironmental variables have been observed, as well. Studies in Alzheimer's transgenic mice, for example, suggest that concurrent presence of latent herpes simplex virus type-1 (HSV-1) in the brain may exacerbate the risk associated with ApoE4 (Miller and Federoff, 2008), by promoting more-favorable conditions for AD-associated neuronal degeneration. Mild-to-moderate AD patients receiving antidiabetic rosiglitazone treatment who are ApoE4-negative show modest cognitive improvement, while ApoE4-positive patients do not (Risner et al., 2006). Additionally, the deleterious effects of ApoE4 may be modulated by the expression of additional genes. Using a genome-wide single nucleotide polymorphism (SNP) analysis to identify latent genetic interactions, Reiman et al. (2007) found an association between the ApoE4 and GRB-associated binding protein 2 (GAB2) genes which quadruples the risk for late-onset AD in ApoE4-positive individuals. GAB2 helps to regulate tau phosphorylation, through the AKT/GSK3beta kinase cascade.

Although the role of the tau gene in Alzheimer's pathology remains largely unclear, evidence from transgenic mouse studies demonstrates the interaction between tau and APP expression on subsequent pathology and behavior. For instance, although no change in beta-amyloid levels or plaque burden was detected in transgenic mice with both human APP overexpression and reduced tau expression (i.e., either homogeneous or heterogeneous tau-knockout), these animals performed comparably to wild-type animals in the Morris water maze (Roberson et al., 2007).

The neuronal sorting receptor (SOR1) gene is associated with late-onset AD (Rogaeva et al., 2007). SOR1 promotes recycling of APP from the cell surface, and underexpression of SOR1 permits beta amyloid production from APP through

an alternative pathway. In addition, a point mutation in a mitochondrial enzyme-coding sequence has been associated with AD (Lin et al., 1992), and genes from chromosomes 10 (locus 10p3) (Zubenko, 2001) and 20 (Olson et al., 2002) have been implicated.

Epigenetics, the study of heritable regulation of gene expression without alteration of the DNA sequence (Waddington, 1953; Weinhold, 2006; Dolinoy et al., 2007), is a growing area in bioinformatics research supported by genome-wide scans of families (Bock and Lengauer, 2008). Epigenetic inheritance has been proposed as one mechanism underlying non-genotoxic carcinogenesis (e.g., peroxisome proliferator-activated receptor; Gonzalez, 2002), asthma (Vercelli, 2004), autism (Lopez-Rangel and Lewis, 2006), and possibly late-onset AD (Bassett et al., 2002, 2006). Additionally, modifications of gene expression during early mammalian development caused by nutritional, environmental, and even behavioral factors can profoundly alter the resulting phenotype (Dolinoy et al., 2007); epigenetic inheritance explains, in part, how individuals with identical genetic makeup may exhibit divergent phenotypic trajectories over the lifespan (Fraga et al., 2005). Different patterns of DNA methylation by maternal and paternal alleles during early development, for example, can result in gene expression determined by parent-of-origin ("imprinting") (Lewin, 2008, p. 833). Two extremely rare human genetic disorders, Prader-Willi syndrome (PWS; OMIM 176270) and Angelman syndrome (AS; OMIM 105830), illustrate the disparate phenotypic effects of genomic imprinting: Partial or complete deletion of 15q11-q13 (i.e., bands 11 through 13 of the long arm of chromosome 15) from the maternal copy of the chromosome results in AS, characterized by delayed motor, speech and cognitive development, seizures, and eccentric affective behavior (Angelman syndrome was once known, pejoratively, as "happy puppet syndrome") (Magenis et al., 1987); similar corruption of the paternal copy results in PWS, characterized by hyperpha-

gia, hypotonia, and impaired learning ability (Ledbetter et al., 1981). Because DNA methylation events can occur throughout the entire lifespan (Bjornsson et al., 2008), epigenetic modulation of individual phenotype represents an ongoing process.

1.8 Animal Models of Alzheimer's Disease

Animal models are useful in research for identifying characteristic biochemical, histopathological, and behavioral manifestations of neurological disorders (e.g., Gallagher and Rapp, 1997). Additionally, these models can be used to elucidate the functional role of genes or proteins and the impact of genetic manipulation on phenotypic expression, as well as to evaluate therapeutic efficacy of pharmacologic and/or behavioral interventions. The assessment and analytical methodologies for animal models should closely resemble those of their human counterparts, as well (e.g., MRI in living transgenic mice, Jack et al. 2007). An ideal model for Alzheimer's disease, for example, should manifest progressive cognitive impairment and AD-characteristic neuropathology (Janus and Westaway, 2001; Dodart et al., 2002; Bloom et al., 2005; Brasnjevic et al., 2006). In addition, reproducibility of results by different investigators is necessary for establishing a valid animal model.

In some cases, naturally-occurring examples of Alzheimer's-like neuropathology already exist, which can be studied as model systems of disease (Nakayama et al., 2004). For instance, cognitive dysfunction syndrome, a geriatric neurological disorder in domestic cats (Levine et al., 1987) and dogs (Ruehl et al., 1995; Cummings et al., 1996a, 1996b; Borras et al., 1999), is symptomatically reminiscent of human Alzheimer's disease (e.g., confusion, sensory disorientation, reduced social interaction, vocalization, disruption of sleep-wake cycle, and urinary incontinence). Histopathological changes similar to human AD, including extraneuronal deposition of beta amyloid (primarily Aβ_{42}) protein and intraneuronal accumulation of hyper-

phosphorylated tau protein (although no mature neurofibrillary tangles), have been observed in aged (10+ year-old) cats (Head et al., 2005; Cummings et al., 1996c; Gunn-Moore et al., 2006). The highest concentrations of Congophilic staining were found in the deep cortical and anterior cerebral regions, and intense neuronal AT8-immunoreactivity was detected. In aged (8-18 y/o) dogs, scattered diffuse (but not compact) beta amyloid plaques were detected in the cerebral cortex using immunohistochemical staining, in addition to extensive cerebrovascular amyloid deposition (Borras et al., 1999). Both cats and dogs respond positively to antioxidant-enriched diets, with improved cognitive performance (Cotman et al., 2002; Milgram et al., 2005) and increased longevity (Cupp et al., 2006). In addition to domestic canine and feline species, both aged nonhuman primates (chimpanzees and orangutans; Gearing et al., 1994, 1997) and polar bears (Tekirian et al., 1996) exhibit diffuse Aβ_{42}-specific plaques in the cortex and hippocampus; tau-immunoreactive paired helical filaments have also been found in aged chimpanzee cortical tissue (Rosen et al., 2008). Both neurofibrillary tangles and neuritic plaques have been found in the cerebral cortex of aged sheep (Nelson et al., 1994). In addition, argyrophilic neurofibrillary tangles, as well as both neuritic and diffuse plaques, were found in the cortex and hippocampus of an aged wolverine (Roertgen et al., 1996).

Transgenic animals, by contrast, represent artificial (constructed) model systems in which one or more genes are deleted, mutated, and/or overexpressed (Jaenisch, 1988; Lee et al., 1996). The resulting genetically-engineered animals are useful subjects for observing and measuring the pathogenesis and progression of disorders, as well as for examining the effects of preventive or therapeutic interventions. For studying many human neurodegenerative diseases, including AD, transgenic mice are commonly used (e.g., Emilien et al., 2000; Wong et al., 2002). These mice are produced through pronuclear injection of an appropriate cDNA transgene into a

single-cell mouse embryo which is implanted into a pseudopregnant female mouse. Successful incorporation of the transgene into the mouse genome is verified by genotyping the offspring by polymerase chain reaction (PCR) at weaning. Using this protocol, multiple lines of transgenic mouse models have been developed and maintained for studying Alzheimer's disease. The selection of an appropriate background mouse strain is an important consideration, however (Gerlai, 1996; Carlson et al., 1997; Crawley et al., 1997; Picciotto and Wickman, 1998; Nguyen et al., 2000; Pugh et al., 2004). Inbred FVB/N or C57BL/6J lines carrying the APP transgene, for instance, are more prone to premature death preceded by behavioral anomalies (neophobia, decreased Y-maze spontaneous alternation activity), relative to outbred hybrids (Carlson et al., 1997). Inbred CBA/J mice exhibit learning impairment in the Morris water maze, which may be related to aberrant hippocampal electrophysiology, specifically, impaired long-term potentiation in the CA1 region (Nguyen et al., 2000).

1.8.1 PDAPP Transgenic Mouse Model

The PDAPP transgenic mouse model includes a single-mutation human APP gene, with phenylalanine replacing valine at residue 717 (V717F). The hAPP minigene, using a human platelet derived growth factor b (PDGF-b) promotor, is incorporated into mice having a C57B6 x DBA/2 F1 hybrid strain background (Games et al., 1995; Rockenstein et al., 1995).

Pathological Characterization

A significant increase in transgenic APP levels is detected in the cerebral cortex and hippocampus between 4 and 18 months of age (Games et al., 1995), including a 17-fold increase in hippocampal beta amyloid (primarily Abeta[1-42]) between 4 and 8 months of age (Johnson-Wood et al., 1997). At three months of age, significant

hippocampal atrophy is detected (Dodart et al., 2000). At 4-5 months, enhanced paired-pulse facilitation (PPF) and rapidly-decaying long-term potentiation (LTP) are noted (Larson et al., 1999). In addition, decreased proliferation of neurons in the hippocampal subgranular zone (SGZ) occurs during the first year (Donovan et al., 2006), suggesting that altered neurogenesis contributes to decreases in hippocampal function in these animals. Neither neurofibrillary tangles nor paired helical filaments have been detected in the PDAPP transgenic mouse model (Masliah et al., 1996; Masliah et al., 2001).

Although three- to four-month-old homozygous PDAPP mice exhibited some extent of mature beta amyloid deposits, not all heterozygous PDAPP mice manifest mature deposits (Dodart et al., 2000). Additional studies have shown that between four to six months of age, no significant pathology is observed in heterozygotic animals (Games et al., 1995). However, beta amyloid deposits are found exclusively in the hippocampus, corpus callosum, and cerebral cortex of transgenic animals in an age-dependent fashion beginning at six months (Dodart et al., 2000). Eight- to twelve-month-old heterozygous PDAPP mice exhibit structural pathologic features reminiscent of human AD (Masliah et al., 1996), including similarities in amyloid deposition, dystrophic neurites, and glial cell reactivity. Thioflavin S-positive beta amyloid deposits are detected in the hippocampus and entorhinal cortex between 12 and 15 months, with little or no detectable deposits earlier (Reilly et al., 2003). By 16-17 months of age, increased plaque density was observed within the entorhinal cortex, dentate gyrus, and CA1 of the hippocampus (Chen et al., 2000), and diffuse plaques were found within the hippocampus and entorhinal cortex.

An interaction among human APP, ApoE, and ACT expression in the PDAPP mouse model has been shown to influence beta amyloid plaque deposition and character (Nilsson et al., 2004). PDAPP mice lacking both ApoE and ACT have no

compact beta amyloid deposits, and only minimal diffuse deposits. Overexpression of ACT in ApoE-knockout PDAPP mice increases diffuse beta amyloid deposition, but compact deposits remain absent. Expression of ApoE produces higher diffuse beta amyloid deposition, relative to ApoE-knockouts, and is associated with the presence of compact deposits. The presence of both ApoE and ACT produces the highest levels of both diffuse and compact beta amyloid deposition.

Behavioral Characterization

Behavioral manifestations and deficits precede amyloid plaque deposits, with radial arm maze impairment (reference memory errors) detected at three months, compared to wildtype controls (Dodart et al., 1999). Object recognition memory impairment is noted at six months, with significant decline identified at 9-10 months (Dodart et al., 1999; Dodart et al., 2000).

Correlation-, factor-, and discriminant-analyses of behavioral measures obtained from comprehensive task batteries support the existence of uniquely identifiable sensorimotor and cognitive domains, as well as domain-specific pathological correlations in the PDAPP mouse model (Leighty et al., 2004). Signficant correlations are observed both within and between water maze tasks (Morris maze, platform recognition, and radial arm water maze), and between these tasks and Congophilic beta amyloid deposition within the cerebral cortex and hippocampus of 15 month-old PDAPP mice (Leighty et al., 2004). In particular, spatial memory tasks have identified domain-specific deficits in cognitive performance. Progressive deficits in water maze serial spatial memory are seen between 13 and 18 months, and non-progressive deficits in water maze spatial reference memory are detected between 13 and 18 months, and as early as 3-4 months, compared with nontransgenic mice (Chen et al., 2000).

The pattern of cognitive impairment in 15 month-old PDAPP mice is influenced by human ApoE and ACT expression (Nilsson et al., 2004). No impairment in the Morris maze and radial arm water maze tasks are observed in PDAPP mice lacking both ApoE and ACT. Expression of ApoE results in substantial impairment in both the Morris maze and radial arm water maze tasks. Overexpression of ACT in ApoE-knockout mice results in modest impairment in the radial arm water maze task, but normal Morris maze performance.

1.8.2 APPsw Transgenic Mouse Models

A double mutation of the APP gene, occurring at residues 670 (K670N) and 671 (M671L), known as the "Swedish" mutation is the basis for several transgenic mouse models. The models differ in strain backgrounds, which produces marked differences in cognitive and sensorimotor phenotypic characteristics. The FVB/N mouse strain, for example, exhibits susceptibility to hippocampal and cortical apoptosis and astrogliosis (Moechars et al., 1996) as well as retinal degeneration (Vinores et al., 2003), and the APPsw model having this background is particularly vulnerable to hippocampal cell death (Mohajeri et al., 2004). Sturchler-Pierrat et al. (1997) describe the generation of another APPsw model called the APP23 mouse, produced by insertion of the hAPP751 cDNA into the XhoI site of an expression vector containing the murine Thy1.2 glycoprotein gene and promotor. The Tg2576 mouse model (Hsiao et al., 1996; Chapman et al., 1999) introduces mutant human APP through a hamster prion protein gene, PrP, promotor on a C57BL6 x C57BL6/SJL strain background.

Pathological Characterization

Young APP23 mice (6 months) exhibit sparse beta amyloid deposits, however, substantial plaque deposition (including dense deposits) are seen at 24 months in

diverse regions, including the neocortex, hippocampus, and thalamus (Sturchler-Pierrat et al., 1997). Both heterozygous and homozygous APP23 mice exhibit extensive plaque deposits between 14-18 months of age in the neocortex and hippocampus (Calhoun et al., 1998), although no significant neuronal loss was present. Additionally, no neurofibrillary tangles were identified through immunostaining and Bielschowski silver staining (Schwab et al., 2004). Nine month-old animals which received intrahippocampal injections of human beta amyloid extract at five months of age exhibit extensive hippocampal plaque deposition (Meyer-Luehmann et al., 2006). The induction of beta-amyloidogenesis at this earlier timepoint within a susceptible host demonstrates the nucleant ("seeding") aspect of Alzheimer's pathology (Riek, 2006).

Neuropathology in the Tg2576 mouse was first studied by Hsiao et al. (1996). Increased levels of Aβ_{40} were found as early as two months of age, with a five-fold increase by 11 to 13 months of age. Elevated levels of Aβ_{42} were also found, with a 14-fold increase by 11 to 13 months of age. Amyloid deposits within the frontal, temporal, and entorhinal cortices, as well as the hippocampus and cerebellum were found in 11 to 13 month-old mice (Hsiao et al., 1996). Similarly, brain levels of both Aβ_{40} and Aβ_{42} increased from 4-8 months of age through 12 months, with beta amyloid deposits within the hippocampus and cortex by the later time-point (Pratico et al., 2001). By 18 months of age, Tg2576 mice have widespread beta amyloid deposition in the hippocampus and neocortex (Pratico et al., 2001). Benzing et al. (1999) also found extensive brain beta amyloid deposits in 18 month-old Tg2576 mice, predominantly within the temporal and cingulate cortices, accompanied by astrocytes and microglia. Beta amyloid deposits within the cingulate cortex, entorhinal cortex, and hippocampal region CA1 were detected in 16 month-old Tg2576 mice (Irizarry et al., 1997). Positive plaque-associated astrocyte activation was found, consistent with

Benzing et al. (1999). King and Arendash (2002b) noted extensive synaptophysin staining of the hippocampus and neocortex in 19 month-old Tg2576 mice, relative to age-matched nontransgenic animals; decreased staining was observed within the plaque cores, with increased staining around the periphery. Additionally, Westerman et al. (2002) suggested an association between the presence of insoluble beta amyloid and Morris water maze retention performance deficits in Tg2576 mice over 10 months of age. Lesne et al. (2006) associated extracellular accumulation of a soluble 56 kD beta-amyloid dodecamer (named "beta-amyloid star 56") with spatial memory deficits observed in 6 to 14 month-old Tg2576 mice.

Behavioral Characterization

APP23 mice exhibit spatial learning and memory deficits in the Morris water maze task (Kelly et al., 2003; Lalonde et al., 2002; Van Dam et al., 2003). Evidence of cognitive impairment precedes plaque deposition; heterozygous APP23 mice show significant beta amyloid plaque deposition at 6 months of age, but demonstrate impairment in both the acquisition and retention components of the Morris water maze as early as 3 months of age (Van Dam et al., 2003). Lalonde et al. (2002) noted impairment of 16 month-old mice in the acquisition – but not retention – component of Morris water maze performance. Kelly et al. (2003) obtained similar findings, with 3, 18, and 25 month-old APP23 mice showing progressively increasing latencies in the same task.

Tg2576 mice exhibit a different pattern of behavioral impairment from APP23 mice. Holcomb et al. (1999) found no impairment in sensorimotor tasks, Y-maze entries or alternation, Morris water maze, or visible platform performance of 3 month-old Tg2576 mice. King and Arendash (2002a) also found no impairment in Y-maze entries, Morris water maze, visible platform, and circular platform performance in 3 month-old mice. However, they noted impairment in selected sensorimotor tasks (in-

creased open field activity and poor balance beam performance), as well as Y-maze alternation. Hsiao et al. (1996), by contrast, did not find Y-maze alternation deficits in 3 month-old Tg2576 mice, but noted these impairments at 10 months of age, as well as increased latency of Morris water maze acquisition. Tg2576 mice exhibit impaired novel object habituation and reduced reactivity to spatial novelty as early as 7 months of age (Middei et al., 2006), suggestive of hippocampal-mediated attentional deficits. Circular platform performance deficits were detected in 7 month-old mice, relative to non-transgenic littermates (Pompl et al., 1999). Spatio-temporal context learning impairment has been observed in 10-12 month-old Tg2576 mice (Good et al., 2007), reminiscent of human episodic memory deficits (i.e., "what," "when," and "where" aspects of object presentation). Holcomb et al. (1999) found no impairment in Morris water maze or visual platform performance by 9 months of age, but noted that Tg2576 mice showed decreased Y-maze alternation, compared with non-transgenic animals. By contrast, King and Arendash (2002a) noted a significant increase in visible platform latency at 9 months. Manifold learning and memory impairment (Morris maze spatial acquisition and retention, circular platform escape latency, RAWM working memory) was also found in 9 month-old transgenic mice, relative to age-matched nontransgenics (Arendash et al., 2006). Additionally, transgenic mice displayed significant sensorimotor impairment (balance beam and string agility tasks) by 14 months, as well as impaired Y-maze alternation by 19 months, relative to age-matched nontransgenic animals (King and Arendash, 2002a). Hsiao et al. (1996) observed significant impairment in Morris water maze retention between 9 and 15 months of age. In addition, impaired reversal learning (odor discrimination task) was observed in 6 month-old Tg2576 mice (Zhuo et al., 2007), with animals requiring more trials to reach criterion, relative to age-matched non-transgenic controls.

1.8.3 APPsw+PS1 Transgenic Mouse Model

The APP/PS1 mouse model incorporates two human FAD-associated mutation variants: the Tg2576 APPsw Swedish double-mutation and the presenilin-1 (PS1) M146L mutation (Holcomb et al., 1998).

Pathological Characterization

The most striking feature of the APPsw/PS1 double transgenic mouse model is its accelerated beta amyloid deposition temporal profile, relative to the APPsw mouse. Compact beta amyloid deposits are found as early as 12-16 weeks of age and, by 24-32 weeks of age, plaques surrounded by reactive astrocytes are observed (Holcomb et al., 1998). Takeuchi et al. (2000) found beta amyloid plaque deposits in the neocortex and, to a lesser extent, the hippocampus, of three month-old transgenic mice. By six months, small diffuse plaques and larger compact plaques were ubiquitous in the cortex. No changes were detected in hippocampal CA1 neuron density nor in synaptophysin-associated immunoreactivity, relative to age-matched single transgenic and non-transgenic mice (Takeuchi et al., 2000). These results differ from those of Gordon et al. (2002), however, who found beta amyloid deposits in six month-old APP/PS1 mice, but not in three month-old animals. These deposits were found mainly in the hippocampus and frontal and entorhinal cortices, in association with both reactive astrocytes and dystrophic neurites. By 15 months of age, positive staining for reactive astrocytes increased throughout the brain, particularly within the cerebral cortex and striatum (Gordon et al., 2002). Jensen et al. (2005) also reported significant beta-amyloid immunostaining and Congo red staining within the cortex and hippocampus in 17-month-old mice. Borchelt et al. (1997) found substantial beta amyloid deposition in the hippocampus and cerebral cortex of APP/PS1 mice, with age-related increases in plaque burden from 9 months to 12 months of

age. A significant negative correlation between total beta amyloid burden in the hippocampus and frontal cortex, and T1-T4 acquisition trial error reduction in the radial arm water maze, underscores the parallel trajectories of neuropathological and cognitive deterioration (Gordon et al., 2001). Congophilic staining in the frontal cortex is significantly positively correlated with cognitive impairment (working memory errors in the radial arm water maze task) in APP/PS1 mice (Arendash et al., 2001; Gordon et al., 2001). In addition, non age-dependent cortical electrophysiological anomalies are detected in APPsw/PS1 mice (Wang et al., 2002), manifested as reduced theta (4 to 6 Hz) and enhanced beta (14 to 27 Hz) and gamma (28 to 40 Hz) EEG activity, underscoring the pathophysiological distinctiveness of this genotype.

Behavioral Characterization

APP/PS1 mice up to 9 months of age do not exhibit deficits in sensorimotor tasks (Holcomb et al., 1999). However, by 6 to 9 months of age, these mice show increased activity (i.e., number of entries) and impairment in Y-maze alternation, compared with age-matched nontransgenic mice (Holcomb et al., 1999). Additional studies by Arendash et al. (2001) utilized a more comprehensive behavioral task battery, and identified domain-specific patterns of impairment. At 5-7 months of age, APP/PS1 mice did not differ from age-matched nontransgenics in either sensorimotor- or cognitive-based tasks, including Y-maze alternation, with the exceptions of Y-maze entries (APP/PS1 showed greater activity) and balance beam (APP/PS1 showed impairment). In a later study, however, transgenic animals between 4.5 and 6 months of age exhibited impairment in both Morris maze acquisition and retention, as well as RAWM working memory (Jensen et al., 2005), relative to age-matched nontransgenic mice. This early evidence of cognitive impairment coincides with incipient plaque formation within vulnerable brain regions (cortex and hippocampus). Ten month-old animals exhibit significantly greater duration within

– as well as increased percent-entries into – the open-arms of the elevated plus maze (Pugh et al., 2007). By 15-17 months of age, APP/PS1 mice showed increased activity (both in open field and Y-maze entries) as well as general sensorimotor and cognitive impairment, compared to age-matched nontransgenic animals (Arendash et al., 2001). Ethell et al. (2006) noted working memory impairment (overall number of errors during Trials 4 and 5 in the radial arm water maze task) in eight month-old APP/PS1 mice, compared with age-matched non-transgenic control animals.

1.8.4 APPsw+PS1+Tau Transgenic Mouse Model

The trigenic-AD mouse (LaFerla, 2006) was produced by subcloning an hAPP695 cDNA fragment with the Swedish double mutation (K670N, M671L), and human tau with the P301L mutation, into the murine Thy1.2 expression cassette. These were inserted by co-microinjection into pronuclei of single-cell embryos from PS1 (M146V mutant) knock-in mice. Currently, this mouse model most-closely mirrors the pathological profile of human Alzheimer's disease, with simultaneous expression of human mutant APP and tau proteins and the appearance of their respective histopathological and behavioral manifestations.

Pathological Characterization

Trigenic-AD (3xTgAD) mice develop both amyloid plaques and neurofibrillary tangles in an age-related progressive fashion (Billings et al., 2005). Intraneuronal beta amyloid protein is found in the hippocampus and amygdala of four month-old animals, but neither plaque nor tangle formation is evident at this age (Billings et al., 2005). Beta amyloid deposits initially appear in the cortex, as early as six months of age, and spread to the hippocampus, while tau pathology emerges in the reverse order (Oddo et al., 2003a). Beta amyloid deposition precedes tangle pathology, despite comparable overexpression of the respective human mutant transgenes (Oddo et al.,

2003a), and consistent with the amyloid cascade hypothesis of Alzheimer's disease pathogenesis. Synaptic dysfunction progresses in an age-related manner, with deficits in long-term potentiation preceding plaque and tangle pathology, and deficits in long-term synaptic plasticity related to the accumulation of beta amyloid within neurons (Oddo et al., 2003b; LaFerla and Oddo, 2005). In addition, the trigenic-AD mouse model provides evidence that oligomerization of beta amyloid initially occurs intraneuronally (Oddo et al., 2006).

Behavioral Characterization

At 2.5 months of age, 3xTgAD mice are prepathologic and cognitively unimpaired (Morris water maze), although they exhibit initially decreased vertical open field activity (number of rearings), relative to age-matched nontransgenics (Gimenez-Llort et al., 2007), and by six months of age, increased horizontal open field activity (locomotion) is observed. By four months of age, deficits in long-term memory retention (cued Morris water maze) are found (Billings et al., 2005). Immunotherapy-induced clearance of beta amyloid, which accumulates in the hippocampus and amygdala by four months of age, temporarily alleviates cognitive impairment, however. Additionally, six month-old trigenic-AD mice exhibit poorer acquisition in the Morris water maze, as well as subsequent memory retention deficits (probe trial), relative to nontransgenics (Gimenez-Llort et al., 2007).

1.8.5 Animal Models: A Coda

The most commonly used variants of the Alzheimer's transgenic mouse model were described in this section. However, these examples represent only a sample from the growing diversity of transgenic mouse strains available to researchers for exploring various aspects of Alzheimer pathology. Transgenic mice are currently the best animal model system for investigating the pathogenesis and progression of

Alzheimer's disease, as well as for evaluating experimental diagnostic and therapeutic interventions. However, the Alzheimer's transgenic mouse is not a perfect model, and is incomplete with respect to pathology.

The immuno-neuroinflammatory response, for instance, differs between human AD and the mouse model, possibly due to reduced sensitivity of murine complement factors to human beta-amyloid (Webster et al., 1997). In humans, significant microglial activation and increased levels of complement factors are associated with plaque cores (Schwab et al., 2004). By contrast, a weaker microglial response that is largely confined to the plaque periphery is observed in the Alzheimer's transgenic mouse (Schwab et al., 2004). In addition, plaque-associated neuronal loss occurring in human AD is not consistently observed in transgenic mice. In Tg2576 mice, for example, some studies (e.g., Tomidokoro et al., 2001) report neuronal loss associated with beta-amyloid plaques, albeit "limited," while other studies (e.g., Stein and Johnson, 2002) do not. Finally, it is noteworthy that human APP differs from the murine homolog by only seventeen amino acids (Jankowsky et al., 2007) and that, while wild-type rodents do not exhibit age-associated Alzheimer-like amyloid lesions (Shivers et al., 1988) for unknown reasons (Selkoe, 1989; Cai et al., 2001; Jankowsky et al., 2004), the co-expression of murine APP in Alzheimer's transgenic mice alters the solubility and distribution pattern of the aggregates (Jankowsky et al., 2007). Despite these, and other, potential deficiencies and limitations, each new generation of Alzheimer's transgenic mouse model constitutes an evolutionary step of scientific progress, both reflecting our current knowledge of Alzheimer's disease and promising new insights.

Currently, observing and measuring the overt behavioral (cognitive) manifestations of Alzheimer's disease provide the most direct, and least invasive, means for investigating the disorder. The development of technology for observing, measuring,

recording, classifying, and interpreting animal behavior mirrors advances in psychology, neurology, and associated behavioral sciences. The modern researcher can select from a wide range of experimental apparatus, including arenas, mazes, and operant chambers, and utilize computer-mediated sensors for detecting even the most subtle subject responses. Not surprisingly, the tools available for data analysis reflect the sophistication of these devices, from basic statistics to complex computational algorithms. Only recently, however, have neuroscientists begun to exploit the greater potential of contemporary analytical engines for behavioral research.

1.9 Treatments for Alzheimer's Disease

Several classes of pharmacologic agents (Suh and Checler, 2002; Cummings and Zhong, 2006; Klafki et al., 2006; Silvestrelli et al., 2006) have been developed for the treatment of Alzheimer's disease, including cholinesterase inhibitors, N-methyl-d-aspartate (NMDA) receptor antagonists, and secretase-modulators. In addition, various natural products are under investigation as therapeutic agents (Cox and Balick, 1994).

1.9.1 Pharmaceuticals

Cholinesterase inhibitors (e.g., tacrine, donepezil, galantamine, and rivastigmine) help compensate for diminished availability of the neurotransmitter acetylcholine by blocking the action of acetylcholinesterase, which normally breaks down acetylcholine following release (Gauthier, 2001). The resulting increase in synaptic concentration of acetylcholine modestly ameliorates the depletion of the neurotransmitter caused by progressive loss of cholinergic neurons, but does not alter the underlying pathogenetic mechanism of the dementia. These agents have demonstrated effectiveness in mild to moderate AD (Holmes et al., 2004), albeit for limited duration (1-2 yrs).

Tacrine (Cognex^{TM}) was the first centrally-acting acetylcholinesterase inhibitor approved for the treatment of Alzheimer's disease. Clinical trials in individuals with mild to moderate dementia showed improvements in language skills (production, comprehension, and word recognition) (Raskind et al., 1997), and demonstrated relief of major psychiatric symptoms (anxiety, apathy, and hallucinations) in individuals with moderate dementia (Farlow et al., 1992; Knapp et al., 1994; Kaufer et al., 1996). Tacrine has been replaced by newer medications, due to its poor oral bioavailability and adverse drug reaction issues (incl. gastrointestinal and urinary problems, hepatotoxicity) (Qizilbash et al., 1998). Donepezil (Aricept^{TM}, Eisai), which was introduced in 1997, features 100% oral bioavailability, fewer reported side effects, and a longer half-life (i.e., requires fewer daily doses), compared with tacrine (Scarpini et al., 2003). Clinical trials in mild to moderate AD patients showed improved cognition and activities of daily living (Winblad et al., 2001), and reduced psychiatric symptoms (anxiety, delusions) in moderate to severe AD patients (Rogers et al., 1998; Feldman et al., 2001). Galantamine (Razadyne^{TM}, Ortho-McNeil Neurologics) is an alkaloid originally isolated from flowers and bulbs of the Voronov snowdrop (*Galanthus woronowii*; Amaryllidaceae) (Scott and Goa, 2000). Galantamine is not only a competitive reversible acetylcholinesterase inhibitor, but also increases acetylcholine release by modulating nicotinic cholinergic receptors (Woodruff-Pak et al., 2001). In mild to moderate AD patients, galantamine has been shown to improve cognition and decrease anxiety and hallucinations (Tariot et al., 2000). Rivastigmine (Exelon^{TM}, Novartis) inhibits both acetylcholinesterase and butyrylcholinesterase, and is particularly beneficial in patients exhibiting Alzheimer's- or Parkinson-associated delusions or hallucinations (Burn et al., 2006; Gauthier et al., 2006; Touchon et al., 2006). Rivastigmine became the first product approved for the treatment of mild to moderate dementia associated with Parkinson's disease (Emre et al., 2004) in 2006.

NMDA receptor antagonists (e.g., memantine) bind to NMDA receptors with greater affinity than magnesium ion, thus preventing the prolonged influx of calcium ions underlying neuronal glutamatergic excitotoxicity observed in AD (Cacabelos et al., 1999). These agents can dissociate from the postsynaptic receptors to permit glutamate binding, thus preserving physiological and signal-conductive processes of the receptor. Memantine (Namenda^{TM}, Forest Pharmaceuticals), introduced in 2003, is a non-competitive, low-affinity NMDA receptor antagonist (Lipton, 2005), and the first pharmaceutical therapeutic agent for AD which targets the glutamatergic system (Robinson and Keating, 2006). It is indicated for treating moderate to severe AD (Reisberg et al., 2003) and, when administered with donezepil, showns enhanced effectiveness for improving cognitive function in severe AD (Tariot et al., 2004). Eight month-old Alzheimer's transgenic mice (APP/PS1) receiving daily memantine (30 mg/kg/day p.o.) for 2 to 3 weeks showed less-impairment in Morris water maze acquisition (i.e., reduced escape latency), relative to untreated transgenic animals (Minkeviciene et al., 2004), suggesting improved hippocampal-mediated spatial learning ability. Memantine is also a potent serotonergic (5HT3) receptor antagonist (Rammes et al., 2001), but the clinical relevance to AD remains unknown.

Selective amyloid lowering agents (SALA), such as R-flurbiprofen (Flurizan^{TM}, tarenflurbil, Myriad Genetics), represent a new class of pharmacologic therapies for treating mild AD currently under investigation. These agents are related to NSAIDs, and have been shown to reduce the levels of pathogenic Aβ_{42} both in vitro and in animal models (Morihara et al., 2002; Eriksen et al., 2003) by modulating the action of gamma-secretase. Both R-ibuprofen and R-flurbiprofen are poor cyclooxygenase inhibitors, but effectively reduce Aβ_{42} production in vitro (Morihara et al., 2002). However, an 18-month Phase 3 clinical study of Flurizan^{TM} in patients with mild Alzheimer's disease showed no significant difference between treated and untreated

patients, with respect to cognitive- or daily living-based indices (either cognition- and daily living-indices (Green et al., 2008).

1.9.2 Natural Products

Several natural products, including medicinal plant extracts, are under investigation as nootropics and potential therapeutic agents for AD based on ethnobotanical traditions (Perry et al., 1999; Philipson, 2003). For example, flavonoid and terpene trilactone extracts from the Maidenhair tree (*Ginkgo biloba* L.; Ginkgoaceae) may slow the progression of dementia, and are being studied as potential treatments for cognitive impairment and Alzheimer's disease (Zimmermann et al., 2002; Cohen-Salmon et al., 1997; Kanowski et al., 1996; Le Bars et al., 1997; DeKosky et al., 2006). The antioxidant properties of Ginkgo extracts may reduce Alzheimer's-related oxidative stress (Christen, 2000; Philipson, 2003). Daily oral administration of Ginkgo extract (70 mg/kg/day) in eight month-old Alzheimer's transgenic mice (Tg2576) for six months resulted in Morris water maze probe trial (memory retention) performance comparable to age-matched non-transgenic animals (Stackman et al., 2003). There was no significant spatial learning difference between Ginkgo-treated and untreated non-transgenic animals, however (Stackman et al., 2003). Studies in rats (Shif et al., 2006) suggest that, while oral administration of Ginkgo extract may not affect either reference memory or working memory in the Morris maze and eight-arm radial maze tasks, chronic administration improves the overall rate of spatial learning over time. Oral administration of Celastrus seed (*Celastrus paniculatus* Willd.; Celastraceae) oil improves memory retention performance in rats (Nalini et al., 1995) using a two-compartment passive avoidance paradigm (Bures and Buresova, 1963). Treated animals have significantly lower brain levels of monoamine neurotransmitters (i.e., norepinephrine, dopamine, and serotonin) and their metabolites, suggesting

both that Celastrus oil decreases the turnover of central monoamines and that these neurotransmitters exert an inhibitory influence on learning and memory (Nalini et al., 1995). In addition, water-soluble extracts from Celastrus seeds have been shown to inhibit NMDA receptors in vitro, which may serve a neuroprotective role by preventing glutamate-induced neurotoxicity (Godkar et al., 2004). Huperzine A, an alkaloid isolated from the Chinese medicinal herb Qian Ceng Ta (*Huperzia serrata* [Thunb.] Trev. = *Lycopodium serratum*) and related genera of clubmoss ferns, inhibits acetylcholinesterase activity and increases brain acetylcholine levels for several hours (Cheng and Tang, 1998; Xiao et al., 2000; Bai et al., 2000; Zangara, 2003). Vinpocetine, isolated from Madagascar periwinkle (*Vinca alba*), increases cerebral blood flow in patients with senile cerebrovascular disease (Balestreri et al., 1987). The spice turmeric (*Curcuma longa*), found in curry powder, contains curcumin, which has been shown to inhibit the formation of beta amyloid fibrils and to destabilize preformed fibrils, in vitro (Ono et al., 2004). Daily curcumin reduces both soluble beta amyloid and plaque burden in Tg2576 (APPsw) mice, without altering APP levels (Lim et al., 2001). The reduced age-adjusted prevalence of Alzheimer's disease observed in India (Ganguli et al., 2000) may be related to dietary turmeric. Young (2 month-old) APP/PS1 mice receiving extract of *Bacopa monniera* (traditional anti-aging, memory-enhancing therapy from India) for 8 months show decreased cortical Aβ_{40} and Aβ_{42} levels, relative to untreated control animals (Holcomb et al., 2006). Transgenic mice (Tg2576) receiving orally-administered pomegranate juice showed enhanced water maze performance and decreased beta-amyloid (both soluble and compact deposit forms), relative to untreated controls (Hartman et al., 2006). Cannabinoid derivatives have been shown to block beta amyloid-induced microglial activation in vitro, and may have additional neuroprotective potential (Ramirez et al., 2005).

1.9.3 Behavior-Based Therapies

As an adjunct to pharmacologic treatments, several behavioral and psychosocial intervention and rehabilitation strategies have been suggested for early to moderate Alzheimer's disease (Cohen-Mansfield, 2001; Olazaran et al., 2004; Sitzer et al., 2006). Research with Alzheimer's transgenic mice, which exhibit progressive neuropathology and cognitive impairment reminiscent of human-AD, suggests that physical exercise may alleviate AD-associated cognitive impairment (Nichol et al., 2007). After three weeks of *ad lib* access to an exercise wheel, aged (16-18 month-old) transgenic animals performed comparably to age-matched non-transgenic controls in the radial arm water maze task for both short-term (working) and long-term (reference) memory performance (Nichol et al., 2007); by contrast, sedentary aged transgenic- and non-transgenic animals were readily distinguishable. Studies in Alzheimer's transgenic animal models demonstrate the cognitive-protective benefits of "enriched" (i.e., physically- and cognitively-stimulating) housing conditions. Arendash et al. (2004a), for example, noted that aged (16 month-old) APPsw mice which are placed into enriched housing conditions for four months exhibit cognitive performance (spatial reference learning/memory, object identification) significantly superior to that of litter-mates housed under standard laboratory conditions. No differences were found between enriched- and standard-housed mice in total Aβ load of either hippocampus or parietal cortex (Arendash et al., 2004a), suggesting an alternative mechanism to brain β-amyloid reduction for the observed cognitive effects. Subsequent investigations have found both decreased cortical beta-amyloid deposition (Lazarov et al., 2005; Ambree et al., 2006) and enhanced cognitive performance (Costa et al., 2007) effects of environmentally-enriched housing environments, relative to standard animal housing conditions. Indeed, long-term "complete" en-

richment environments (combining social, physical, and cognitive stimulation) may provide optimal conditions to protect against cognitive and functional impairment, and ameliorate AD-associated neuropathology (Cracchiolo et al., 2007). Furthermore, behavioral interventions involving physical and/or cognitive stimulation have demonstrated benefits for maintenance and even modest improvements in mental function of AD patients. Ambulatory AD patients who participated in a collective exercise program (one hour, twice weekly for one year; walking, strength-, balance-, and flexibility-training regime) showed a slower decline in activities of daily living (ADL; Katz Index) performance measures, relative to nonparticipants (Rolland et al., 2007). Quayhagen and Quayhagen (1989) showed that AD patients who engaged in mentally-stimulating activities (verbal and memory exercises, problem-solving) maintained cognitive function during the treatment period, compared to patients who did not engage in the activities. Indeed, AD patients who receive regular (two 45-min sessions/day, three times per week for five wks) training in daily living activities (e.g., routine kitchen activities, letter-writing, telephone use) showed significant improvements in both cognitive- and daily living functionality (Farina et al., 2002); by three months after training ceased, however, these benefits vanished and patients regressed to pre-training status. In some cases, where AD patients do not respond to physical therapeutic interventions, the cause may involve executive dysfunction or gait disturbances arising from AD-associated cerebrovascular disease (Scherder et al., 2007).

1.10 Statement of Purpose

The investigations described in this dissertation embrace the multidisciplinary character of modern Alzheimer's research by introducing, implementing, and evaluating a novel human-mouse parallel cognitive/behavioral testing paradigm, in addition to complementary analytical protocols based on data mining techniques, for neurobehavioral assessment in Alzheimer's transgenic mice. An interference testing paradigm recently developed for evaluating AD patients was adapted for mouse-based testing using spatial memory-domain elements in place of the original (human) semantic learning components. The effectiveness of this novel cognitive assessment task for distinguishing transgenic and/or treatment effects was examined in several studies. Data mining methods – widely used in engineering, business and industrial applications – are only recently gaining recognition in neuroscience as complementary analytic methodologies to well-established statistical approaches. These techniques were shown to be effective tools for evaluating genotype and/or treatment effects using multimetric behavioral data.

The specific aims of these investigations were to:

1. Evaluate cognitive effects of long-term caffeine administration in nontransgenic mice, using both conventional statistical and complementary data mining analytic methods, in a comprehensive sensorimotor and cognitive task battery.

2. Examine short-term caffeine admininstration in both aged nontransgenic and Alzheimer's transgenic mice, to identify potentially differential sensitivity to cognitive-protective benefits of caffeine.

3. Analyze the human semantic interference task dataset using conventional statistical and complementary data mining techniques, to evaluate the diagnostic utility and differential sensitivity of these research tools for human clinical applications.

4. Demonstrate utility of the novel human-mouse parallel testing paradigm ("interference task") for distinguishing among nontransgenic and Alzheimer's transgenic mice with and without G-protein coupled receptor kinase-5 (GRK5) gene-knockout manipulation.

5. Utilize the novel interference task to evaluate therapeutic efficacy of granulocyte macrophage colony-stimulating factor (GM-CSF) in both nontransgenic and Alzheimer's transgenic mice.

CHAPTER 2

BEHAVIORAL ASSESSMENT IN ALZHEIMER'S TRANSGENIC MICE

Comprehensive testing batteries have been developed for physical and behavioral assessment of animal subjects in laboratory settings. The theoretical and practical necessity of utilizing multiple behavioral tasks and/or measures cannot be overstated (e.g., Rogers et al., 1999; van der Staay and Steckler, 2001; Arendash et al., 2004b; Caeyenberghs et al., 2006; Vekovischeva et al., 2007), with justifications including convergent validation, enhanced discriminative capacity, and broader sampling across the behavioral repertoire. Indeed, as Crabbe and Morris (2004) argue, using the example of a rodent behavioral model of human intoxication, several measures from different tasks are necessary to depict strain-independent sensorimotor manifestations of ethanol ingestion. In addition, complex behavioral phenomena (e.g., memory and learning) have both spatial and temporal dimensions which must be addressed during both measurement and analysis. For example, cognitive impairment is more easily identified, relative to normal cognitive ability, using tasks which manipulate the latency (time delay) between learning and subsequent recall of information (McDonald and Overmier, 1998).

An early assessment battery (Irwin, 1968) consisted of fifty observational categories (e.g., locomotor activity, grip strength, righting reflex), individually scored using a nine-point (0-8) scale. This screen was used in pharmaceutical research for evaluating and characterizing drug reponses in animals. Another testing bat-

tery (Moser et al., 1995), developed by behavioral neurotoxicologists at the U.S. Environmental Protection Agency, classifies behavioral observations into six neurological domains: Activity, autonomic, excitability, neuromuscular, physiological, and sensorimotor function. SmithKline Beecham Pharmaceuticals developed a comprehensive testing protocol (SHIRPA; Rogers et al., 1997) encompassing behavioral observations, sensorimotor and neurophysiological processes, and cognitive functioning. SHIRPA is useful for identifying strain-specific behavioral patterns ("behavioral phenotyping") and for distinguishing among inbred mouse strains (Rogers et al., 1999; Crawley, 1999).

The protocols and measures described below, and summarized in Table 2.1, comprise the behavioral assessment battery developed by Gary Arendash and colleagues at the University of South Florida for examining treatment- and transgenicity-effects in Alzheimer's transgenic mice. The tasks are presented in order of testing sequence and include practical details, as well as historical design perspectives.

2.1 Sensorimotor Tasks and Associated Measures

2.1.1 Open Field Activity

The apparatus consists of an open box, 80 cm by 80 cm square with 28 cm-high walls, painted black. The floor of the box is divided into sixteen regions by a grid of white equispaced lines. Each animal is placed in the center of the floor, and the number of lines crossed during a single five-minute trial is recorded (OF). The early version of this chamber (Broadhurst, 1961) featured a somewhat larger (120 cm x 120 cm x 45 cm, 25-cell floor grid) enclosure in which rodent behavior could be observed for a specified time (e.g., 5 min, 1 hr), and the number of lines crossed or regions

entered is recorded. The open field activity task provides a measure of spontaneous locomotor and exploratory behavior.

Circular arenas (e.g., Hall, 1934) are also used. In addition, automated monitoring (photocell-detector, videotape recording) methods are sometimes used in place of a human observer (e.g., Kafkafi, 2003). The original paradigm of observing naive rodent (and, subsequently, other species) behavior in confined arenas focused on autonomic responses (e.g., defecation, urination) as indices of "emotionality," anxiety and stress (Hall, 1934; Hall, 1936). Indeed, familiarity (habituation) with the open field through repeated testing decreases anxiety-like responses in rats (Ossenkopp et al., 1994). Furthermore, the cognitive-emotional determinants of overt behavior have both situational (environmental) and dispositional (e.g., genetic, physiological) components which remain largely unexplored (Ramos and Mormede, 1998).

2.1.2 Balance Beam

A wooden rod measuring 50 cm long by 1 cm wide is fixed between two columnar supports, suspended 45 cm above a padded table surface. At each end of the rod, mounted atop the supporting column, is a 14 cm by 10 cm escape platform. At the start of each of three 60-second trials, the mouse is placed at the center of the narrow rod, aligned perpendicularly to the rods length. The amount of time the animal remains on the rod before falling, up to 60 sec, is recorded. If the animal reaches either platform (escapes), the maximum time of 60 sec is assigned. The average time across three trials is recorded, as well (BB). A series of progressively-narrower beams (e.g., Carter et al., 1999) can be used for quantifying sensorimotor deficits. The balance beam task reflects overall motor coordination and balance.

Table 2.1. Task-associated behavioral measures used in analyses

Task	**Symbol**	**Description**
Open Field	OF	Number of line crossings
Balance Beam	BB	Average latency, in seconds
String Agility	SA	Performance rating score
Y-Maze	YM-AE	Number of Y-maze arm entries
	YM-PA	Percentage of spontaneous alternations
Elevated Plus Maze	EP-CE	Number of closed-arm entries
	EP-OE	Number of open-arm entries
	EP-TO	Time spent in open arms, in seconds
Morris Water Maze	WM-Avg	Overall average escape latency across days during the acquisition phase, in seconds
	WM-Fin	Average escape latency on the final day of the acquisition phase, in seconds
	WM-Ret	Percentage of time spent in former platform-containing quadrant during retention testing
Circular Platform	CPE-Avg	Overall average number of errors across days
	CPE-Fin	Average number of errors on the final day
	CPL-Avg	Overall average escape latency across days, in sec.
	CPL-Fin	Average escape latency on the final day, in sec.
Platform Recognition	PR-Avg	Overall average latency across days, in sec.
	PR-Fin	Average latency on the final day, in sec.
Radial Arm Water Maze (RAWM)	RME-T4	Average number of errors during Trial #4, across all blocks
	RME-T5	Average number of errors during Trial #5, across all blocks
	RME-FT1	Average number of errors during Trial #1 of final block
	RME-FT4	Average number of errors during Trial #4 of final block
	RME-FT5	Average number of errors during Trial #5 of final block
	RML-T4	Average latency for Trial #4 across all blocks, in sec.
	RML-T5	Average latency for Trial #5 across all blocks, in sec.
	RML-FT1	Average latency for Trial #1 of final block, in sec.
	RML-FT4	Average latency for Trial #4 of final block, in sec.
	RML-FT5	Average latency for Trial #5 of final block, in sec.

2.1.3 String Agility

The apparatus for this task resembles that of the Balance Beam task, with the wooden rod replaced by a taut, cotton string. At the start of a single 60 sec trial, the mouse is positioned midway along the length of the string and allowed to grasp the string with its forepaws, and then released. The animal's performance is assessed using a six-point scoring system, and this measure is recorded (SA): "0" is assigned if the mouse falls from the string, "1" is assigned if the mouse maintains its grip on the string for 60 sec using its forepaws, "2" is assigned if the mouse maintains its grip on the string for 60 sec using its forepaws and either hindlimb, "3" is assigned if the mouse maintains its grip on the string for 60 sec using all four limbs, "4" is assigned if the mouse maintains its grip on the string for 60 sec using all four limbs and its tail, and, "5" is assigned if the mouse escapes by reaching either platform. A second trial is permitted if the animal falls from the string immediately after it is released at the start of the first trial. This task extends an early technique for measuring grip strength which utilized the wire grid cover of the animal's cage: the mouse was allowed to grasp the wire grid, which was immediately inverted, suspending the animal several inches above the cage floor (e.g., Sango et al., 1996). The string agility task provides a measure of forepaw grip capacity and overall physical strength and agility.

2.2 Cognitive Tasks and Associated Measures

2.2.1 Y-Maze

The apparatus consists of a Y-shaped, three-armed (21 cm long by 4 cm wide) maze enclosed by 40 cm-high walls. All visible surfaces of the interior are painted black to provide a uniform, neutral testing environment without spatial orientation

cues. A light source is positioned above the center of the maze where the three arms intersect. The mouse is placed in the center of the maze and the sequence of arm-visits is observed for a single, five-minute trial. The total number of arm entries (YM-AE) is recorded, as a measure of exploratory behavior, as well as the percentage spontaneous alternation (YM-PA; i.e., the ratio of the number of visits of all three arms during three consecutive arm-choices to the total number of arm-entries), which is associated with general mnemonic function (Lamour et al., 1989). The Y-Maze paradigm is a variation of the T-Maze, an earlier design in which working memory and exploratory behavior are examined in moderately food- or water-deprived test subjects introduced into the lower end of a T-shaped maze in which a small amount of food or water has been placed at both end-walls of the longer cross-arm. During the pre-testing habituation period, the animal must learn to alternate arm-visits to obtain the reinforcer, often requiring several weeks of training (Hepler et al., 1985; Markowska et al., 1989). The detrimental effects of disrupting cholinergic circuits of the brain (either by nucleus basalis lesions or acetylcholine receptor antagonists) on working memory have been demonstrated using T-maze delayed alternation measures (Hepler et al., 1985; Mastropaolo et al., 1988). The Y-maze task involves both cognitive (basic mnemonic function) and sensorimotor (exploration) components of behavior.

2.2.2 Elevated Plus Maze

The Elevated Plus Maze is a four-armed maze consisting of two opposite "open" arms and two opposite "closed" arms, each measuring 30 cm long by 5 cm wide. The four arms meet at right angles, forming a 5 cm by 5 cm center region. Additionally, the "closed" arms are surrounded by 15 cm-high walls. A light source is positioned above the center of the maze where the four arms intersect. The entire apparatus

is positioned 80 cm above the ground, and the interior walls are painted black to reduce environmental cues. The mouse is placed in the center of the maze, facing a closed arm, and observed for a single, 5-minute trial. The number of open arm entries (EP-OE) and closed arm entries (EP-CE), as well as the total time spent in the open arms (EP-TO), is recorded.

This paradigm is based on the apparatus and interpretation described in Briley et al. (1986), downscaled from the original dimensions (50 cm long by 10 cm wide arms) suitable for rats. Briley et al. (1986) also recorded the total number of arm entries during a single five-minute trial as a measure of total locomotion and the percentage of open-arm entries as an index of fear response, and found that rats show a marked preference for enclosed arms. Briley's design, in turn, was based on the earlier observation that rats exhibit reduced exploratory behavior and enhanced avoidance of open elevated alleys, compared with covered tunnels (Montgomery, 1955). The effects of anxiolytic (chlordiazepoxide, pentobarbital, ethanol) and anxiogenic (FG-7142, caffeine, picrotoxin) agents have been studied using the elevated plus maze, in conjunction with a holeboard test (Lister, 1987). This task is used to evaluate emotionality and general anxiety, although abilities related to "decision-making" and "risk-assessment" may be represented as well (Rodgers and Johnson, 1995). Confirmatory factor analyses (e.g., Wall and Messier, 2000), however, suggest a two-factor solution for this task. Indeed, the "Long-Term Caffeine Administration in Non-transgenic Mice" study reported in this dissertation illustrates segregation between the arm-residency duration measure (EP-TO) and the two arm-entry frequency measures (EP-OE, EP-CE), underscoring the distinction between anxiety-associated and exploratory-locomotor metrics in the elevated plus maze task.

2.2.3 Morris Water Maze (Submerged Platform)

The apparatus consists of a water-filled, 100 cm diameter circular pool which is divided into four equal-sized quadrants by two perpendicular lines drawn on the floor of the pool. A clear 9 cm diameter circular platform is placed in the center of quadrant #2, submerged 1.5 cm beneath the surface of the water. Assorted visual cues are placed along the external circumference of the pool, to serve as extramaze navigational landmarks (e.g., Suzuki et al., 1980). The training and testing procedure consists of two phases: a ten-day "acquisition" phase, involving four trials per day, followed by a one-day, single-trial "retention" phase, which is videotape-recorded for later analysis. The two phases of the Morris Water Maze task involve spatial learning ("acquisition") and reference memory ("retention") processes. On each day of the acquisition phase, the mouse is placed into each of the four quadrants (randomized order), initially facing the side wall, and the average latency to locate and mount the submerged platform is recorded (maximum of 60 sec per trial). The animal is allowed to remain on the platform (having reached the platform either by swimming or through gentle guidance by the experimenter) for a 30 sec intertrial "stay" period. The maximum latency (60 sec) is recorded for any trial in which the animal requires guidance to reach the platform. The average latency for each day is calculated, and both the final-day average latency (WM-Fin) and overall average latency across all days (WM-Avg) are recorded. On the day following completion of the acquisition phase, the submerged platform is removed and the mouse is placed into the quadrant opposite the formerly platform-containing quadrant for a single, 60 sec "probe" trial. The videotape of the probe trial is subsequently examined to determine performance parameters: percentage of total swim time spent in quadrant

#2 (WM-Ret), quadrant preference, number of annulus crossings, swim path and/or average speed, etc.

The Morris water maze (Morris, 1981; Morris, 1984), originally intended to examine neuroanatomical substrates of spatial learning and memory in rats, is the most common paradigm for cognitive evaluation in rodents (D'Hooge and DeDeyn, 2001). Hippocampal lesions in both rats (Morris et al., 1982; Eichenbaum et al., 1990) and mice (Logue et al., 1997), for example, consistently produce acquisition deficits in this task. Both species- and strain-specific differences in performance have been reported, with rats generally superior to mice in the Morris water maze (Whishaw and Tomie, 1996) and C57BL/6 strain performing better than Swiss Webster mice (Wright et al., 2004). Idiosyncratic or anomalous behaviors, such as circling or preference for sinking over swimming, are commonly reported (e.g., Wahlsten et al., 2005). In addition, video recording equipment is often utilized to increase test accuracy and efficiency. For example, recorded sessions can be viewed and evaluated independently by multiple observers, thus reducing experimenter bias and improving the reliability of behavioral response scoring (Graziano et al., 2003; Tecott and Nestler, 2004).

2.2.4 Circular Platform

The apparatus consists of a walled 70 cm diameter arena with sixteen circular "escape" holes equidistantly-spaced around the circumference. Assorted visual cues are provided both inside (on the encircling wall) and outside (on the enshrouding curtain) of the arena, which serve as visual landmarks. A refuge box containing bedding material is positioned underneath a single escape hole (randomly selected, once per subject across entire testing period). Noxious auditory (high-speed fan 15 cm above the platform) and visual (two 150-watt flood lights located 76 cm above the platform) stimuli are used to elicit escape behavior during the single, five-minute

daily trial for eight consecutive testing days. The mouse is placed in the center of the arena and the noxious stimuli are activated at the beginning of each trial, and the number of errors (i.e., head pokes into non-escape holes) and escape latency time are measured. The number of errors (CPE-Fin) and escape latency time (CPL-Fin) are recorded for the final day of testing, as well as the respective averages across all days (CPE-Avg and CPL-Avg). The platform surface is thoroughly cleaned with deodorizing disinfectant to remove olfactory cues due to stress-induced urination and/or defecation. This paradigm is a refinement of the Barnes maze design (Barnes, 1979) for rats. This task primarily evaluates spatial reference memory, but also includes general anxiety and sensorimotor components (e.g., Leighty et al., 2004).

2.2.5 Platform Recognition

The Platform Recognition task utilizes the same quadrant-labeled, 100 cm diameter circular pool apparatus as the Morris Water Maze task, except in that a prominent target (9 cm diameter circular platform with a 10 cm x 40 cm black ensign) is positioned 0.8 cm above the water surface. On each of the four testing days, the animal is released against the wall of the same quadrant for each of four 60 sec trials. The platform is moved to a different quadrant of the pool for each trial. The time required for the mouse to locate and mount the platform (up to 60 sec per trial) is recorded, followed by a 30 sec intertrial "stay" period, during which the animal is permitted to remain atop the platform. If the mouse fails to locate the platform after 60 sec, it is gently guided to the platform to begin the stay period, and a latency of 60 sec is recorded. The average latency on the final testing day (PR-Fin) and the overall average latency (PR-Avg) are calculated and recorded. This task assesses recognition/identification memory processes.

2.2.6 Radial Arm Water Maze

The apparatus consists of a 100 cm diagonal circular pool containing a six-armed (30.5 cm long x 19 cm wide) aluminum insert with a 40 cm diameter central area. The metal insert extends from the pool floor to 5 cm above the water line, providing a neutral background for the test subjects. The task is divided into three (or up to five) three-day "blocks," for a total of between nine and fifteen days. Each daily session consists of four acquisition trials (T1-T4), followed by a 30 min delay, and ending with a delayed-recall "retention" trial (T5). At the beginning of the session, a submerged clear platform (same target as in Morris Water Maze task) is placed at the end of the designated goal arm (randomly chosen each day, in contrast to the Morris maze procedure), and the animal is placed into one of the remaining non-goal arms (in a randomized sequence), facing the central swim area, for each of the four acquisition trials. The mouse is allowed up to 60 sec to locate the platform and, if successful, remains on the platform for a 30 sec intertrial stay period. If the mouse enters a non-goal arm, it is gently drawn back to the starting arm. Also, if the animal fails to locate the platform after 60 sec, it is gently led to the platform for the stay period. The latency to locate the goal platform is recorded, with trials in which the mouse fails to locate the platform recorded as 60 sec. In addition, the number of errors (entries into non-goal arms) committed during each trial are recorded. Experimenters also record any behavioral anomalies (e.g., perseveration, circling, swimming difficulty). After completing the four acquisition trials, the mouse is removed from the pool and warmed under a heat lamp, then returned to its cage for 30 min. For the retention trial (T5), the mouse is placed into the single remaining unfamiliar (i.e., not used as the start arm in T1-T4) non-goal arm, and the performance latency and errors measures are recorded. Ten aggregate measures are calculated upon completion of

the task: Average late-acquisition performance is calculated for both latencies and errors using the T4 data both across all sessions (RML-T4 and RME-T4) and for the final block only (RML-FT4 and RME-FT4); average initial-acquisition performance (T1) in the final block is calculated for latency and errors (RML-FT1 and RME-FT1); average retention performance is calculated for both latency and errors using the T5 data across all sessions (RML-T5 and RME-T5) and for the final block only (RML-FT5 and RME-FT5). Hence, both T4 and T5 measures represent indices of working memory. This task extends the standard Morris water maze, which measures only spatial reference learning and memory capacity, by coupling the learning-acquisition component (T1-T4) with a recall-retention testing phase (T5). The final block of testing (i.e., last three sessions) are more indicative of overall cognitive functioning, while the earlier blocks largely reflect procedural learning of this relatively complex task. The radial arm water maze (RAWM) paradigm is a down-scaled water-based adaptation (Hyde et al., 1998) of earlier dry-land radial mazes for rats having 8 or 12 arms radiating from a central enclosure. This configuration permitted one or more food- or water-baited arms to serve as the goal(s) (e.g., Olton and Samuelson, 1976; Olton, 1977; Crusio et al., 1995). Studies with multi-armed mazes (Cole and Chappell-Stephenson, 2003) suggest the upper limit of spatial memory in rats is between 24 and 32 locations.

2.3 New Paradigms and Parallels: Tales of Mice and Men

The identification of homologous neural substrates in humans and other animal species (e.g., Spear et al., 1990) associated with specific cognitive domains (e.g., implicit memory, episodic and episodic-like memory) represents an important advance in neuroscience research, enabling investigators to develop, test, and refine animal models for human neurobehavioral disorders, including Alzheimer's disease.

Cross-species comparative neuroanatomical studies in both normal and pathological conditions, for example, emphasize the role of the hippocampus (or non-mammalian homologous structures) in spatial memory and learning, as well as declarative memory in general (e.g., Squire, 1992; Eichenbaum, 1999; Maviel et al., 2004). The persistence of location memory in food-storing birds (Biegler et al., 2001), spatial-based radial maze learning in mice (Crusio and Schwegler, 2005), and object-position memory impairment following temporal lobectomy (with concomitant hippocampectomy) in patients (Nunn et al., 1999) are significant examples. These findings have diverse applications beyond neuroscience; cognitive science and learning theory rely on both structural and functional understanding of animal nervous systems and their behavioral manifestations. The microarchitecture and structural interconnectivity of the hippocampus, for example, has inspired theoretical models of Hebbian learning (associative synaptogenesis) and artificial neural computation (e.g., Kelso et al., 1986; Foster et al., 2000) which are used in advanced computing architectures having autonomous learning capabilities ("artificial intelligence").

Specialized cognitive tasks used to explore human memory and learning abilities, as well as to diagnose clinical neuropathology, have been adapted for neurobehavioral assessment in animals, including mice. Specifically, several forms of learning (e.g., associative, discrimination) as well as both long-term and short-term memory systems have been examined through naturalistic and artificial experimental paradigms, often very elaborate, in which multiple parameters (e.g., stimulus dimension and complexity, presentation latency and duration) are manipulated and responses recorded. In developing these parallel assessment methodologies, however, it is important to recognize that although human-based tasks for assessing learning and memory often rely upon verbal responses, while animal-based tasks generally uti-

lize motor responses (e.g., turn direction, arm-visit latency), both response classes represent valid behavioral indices for measurement, comparison, and analysis.

The serial reaction time task, for instance, involves measuring accuracy and response latency across trials in a sequence-learning problem (Nissen and Bullemer, 1987) to assess procedural (implicit) memory function. Diagnostic imaging studies in normal and task-impaired individuals implicate the role of basal ganglia, neocortex, and cerebellum in this task (Knopman and Nissen, 1991; Willingham and Koroshetz, 1993; Pascual-Leone et al., 1996; Rauch et al., 1997). The role of the basal ganglia in implicit memory is underscored by lesion studies in rats wherein striatal lesions – but not hippocampal lesions – result in procedural memory acquisition impairment in a fixed-sequence arm-opening variant of the eight-armed radial maze task (DeCoteau and Kesner, 2000). Sequence learning impairment is also observed following dorsal caudate lesions (Christie and Dalrymple-Alford, 2004) in rats trained in 4-, 8-, and 12-trial arm choice sequences; performance deficits are not observed in animals following dorsal hippocampal lesions. Christie and Hersch (2004) adapted the apparatus for mice, and used multiple repetitions of four-trial sequences within training sessions. Interference effects (declining performance on subsequent session) resulted from replacing the repeating sequence with a random sequence during a single session (Christie and Hersch, 2004), similar to the pattern observed in humans (e.g., Nissen and Bullemer, 1987; Reed and Johnson, 1994). Another mouse-version of the serial reaction task (Cho et al., 2007) uses an operant chamber with nose poke holes positioned on the walls; mice conditioned (FR-1 schedule) to respond through nose poke of illuminated holes are trained to nose poke a specific repeating sequence of holes, using water access as reinforcement for correct-sequence responses. Evidence of Pavlovian and operant (Skinnerian) conditioning, both examples of associative learning, is well-documented for a broad range of animal species (e.g., Spear et al., 1990).

In addition, non-human animals exhibit a remarkable capacity for learning complex stimuli, both natural and artificial, subject to species-specific sensory constraints. Visual discrimination experiments in pigeons (*Columba livia*) (which have highly-developed visual systems), for example, indicate that animals can learn arbitrary categories of novel stimuli. These processes require pattern recognition, associative memory, and visuospatial memory. Examples of visually-presented stimulus classes successfully distinguished by pigeons include: pictures of natural objects (e.g., trees; Vaughan, 1988), light emitting diode array patterns (Jitsumori et al., 2002), and paintings by Monet and Picasso (Watanabe et al., 1995). Similar complex learning has been observed in rats using odor recognition/discrimination tasks (e.g., Kesner et al., 2002), which also underscore hippocampal involvement in temporal-sequence (episodic-like) learning. Event-related episodic-like memory (Morris, 2001; Hampton and Schwartz, 2004) has been examined in mice using object exploration tasks based upon spatiotemporal ordering schemes (Dere et al., 2005). These studies demonstrate that mice are able to distinguish novel from familiar objects ("what"), the relative recency of encounter with an object ("when"), and the location where objects were previously presented ("where") (Dere et al., 2005). Additionally, episodic-like memory deficits are strongly correlated with beta-amyloid plaque burden in Alzheimer's double-transgenic (APPswe/PS1) mice (Savonenko et al., 2005), underscoring the diagnostic utility of episodic memory tasks in Alzheimer's disease.

Working memory processes, as well, have been examined in both humans and animals through cross-species experimental paradigms. For instance, the "Hebb-Williams" maze (Hebb and Williams, 1946), widely-regarded as the standard instrument for evaluating animal spatial learning, has been implemented as a "virtual environment" computer simulation for human subjects (Shore et al., 2001) wherein the physical features of a maze (e.g., walls, alleys, corners) are rendered with graph-

ical algorithms. Virtual mazes represent an analogous maze-learning paradigm for comparing human and animal learning capacity under experimental conditions, as well as for evaluating neuropathological syndromes which manifest distinct spatial problem solving impairment. For example, comparisons between the average acquisition performance of young humans (university undergraduate students) and mice (C57BL/6J strain) across a battery of twelve distinct maze configurations (virtual simulation, for humans; actual implementation, for mice) showed that, although humans learn the task more quickly, the learning curves are strikingly similar (Shore et al., 2001). Subsequent correlation analysis among these tasks reveal a significant association between mental rotation and probe trial performance (Morris maze), but not between other spatial performance measures (Astur et al., 2004), suggesting differential assessment of spatial problem solving ability by the Morris maze and radial maze in human subjects. Furthermore, human patients with surgically-induced medial temporal lobe lesions exhibit performance deficits in room-based analogues of the Morris maze and radial maze tasks (Bohbot et al., 2002) reminiscent of cognitive impairment observed in rodents having similar lesions. These findings are consistent with the results of comprehensive behavioral assessment in Alzheimer's transgenic mice, in which component measures obtained from spatial learning/memory tasks (Morris maze, radial arm maze, platform recognition) exhibit co-localization within cognitive "domains" (e.g., Leighty et al., 2004; Arendash et al., 2007; Leighty et al., 2008), as well as discriminability based upon transgenicity and/or therapeutic treatment effects (e.g., Arendash et al., 2001; Arendash and King, 2002; King and Arendash, 2002a; Leighty, 2003; Arendash et al., 2004b; Leighty et al., 2004; Jensen et al., 2005; Leighty et al., 2008).

A promising research direction for comparative psychology and behavioral neuropathology involves interference effects on working memory and learning. Interfer-

ence is a form of "forgetting" which occurs when one recall process compromises another (Postman and Underwood, 1973). Two types of interference – proactive and retroactive – are differentiated by the temporal dependency of forgetting. In proactive interference, earlier learning interferes with subsequent learning, as when a professional typist experiences difficulty learning to play the piano. By contrast, retroactive interference occurs when recent learning interferes with recall of earlier learning as, for example, when an individual has trouble remembering his previous address or telephone number after recently relocating to a new city (and learning a new address and phone number). Exposure to a complex, overstimulating environment (e.g., Toffler's "information overload") is a common cause of interference in humans (and other animals, as well), and sustained pressure to adapt to accelerating change is associated with both cognitive impairment (confusion, disorientation) and physiological dysfunction (anxiety, stress) (e.g., Toffler, 1970). Radial maze learning in rats, for example, is subject to intertrial proactive interference (e.g., Cohen et al., 1996; Roberts and Dale, 1981) when the delay interval between successive trials is very brief; the interference effect is likely due to errors in temporal discrimination among events occurring within- vs. between-trials (Roberts and Dale, 1981). Additionally, interference effects observed in aged rats in visual discrimination tasks (e.g., Winocur, 1984) resemble performance deficits in hippocampally-lesioned young animals, suggesting the utility of interference testing for neuropathological assessment. In humans, interference effects have been examined with object recognition and naming tasks in which subjects are asked to recall the items in a group of familiar objects presented either with or without intervening, unrelated "distractor" items (e.g., Fuld, 1981). The Fuld task has been adapted by Loewenstein et al. (2004), for both proactive and retroactive interference effects, as a diagnostic instrument for evaluating probable Alzheimer's disease and mild cognitive impairment in aged indi-

viduals. Individuals with mild Alzheimer's disease can be distinguished from normal aged adults with remarkable accuracy (84.6% sensitivity and 96.2% specificity) using measures of both proactive and retroactive interference (Loewenstein et al., 2004). The interference testing paradigm has been adapted for mice by substituting a spatial learning task (radial arm water maze) for the human-based (verbal) semantic learning task (discussed in Chapter 7).

CHAPTER 3

BEHAVIORAL DATA ANALYSIS

3.1 On Behavioral Classification

Behavioral evaluation is only the information-gathering phase of neurobehavioral investigation, and must necessarily be followed by an analytic process for identifying, quantifying, and interpreting meaningful patterns and trends in the dataset. There are several important reasons for collecting and studying these data, including: behavioral phenotyping (characterizing groups, e.g., by transgenicity or strain, in terms of behavior); therapeutic monitoring (e.g., to evaluate the effects of ongoing pharmacologic treatments, either cross-sectionally by group or longitudinally by individual, to observe trends over time); and treatment-outcome assessment (e.g., looking at the end-result of a therapeutic intervention or manipulation). All of these procedures involve classification, the assignment of individuals to groups on the basis of measured variables. Unfortunately, choosing an appropriate classifier methodology depends upon the data, the hypotheses under investigation, and the available computing resources, among other considerations; just as there is "no free lunch," there is no universally-superior classification algorithm (Wolpert and Macready, 1997). Several broad categories of classifiers exist, however, and there are hybrid and ensemble combinations of methods which may provide even better performance for their particular application (e.g., Sigut et al., 2007). Expert systems, for instance, encapsulate the knowledge and decision-making capabilities of (human) experts within specific prob-

lem domains (Giarratano and Riley, 2004), such as diagnosing bacterial infections based on patient symptoms and microbiology lab results (MYCIN; Shortliffe 1974, 1976).

Behavioral assessment, in practice, generally occurs under conditions and constraints which are far from what is theoretically ideal. The economic and management aspects of animal maintenance and testing, for instance, often impose severe limitations on feasibility. Unfortunately, the discrepancy between the actual and the desirable in many situations cannot be remedied, only accomodated. Consider the problem of adequate sample size (e.g., Raudys and Jain, 1991): a sample must be sufficiently large to be representative of a (theoretically-immense) population of candidate individuals but, moreover, the sample must be manageable; hence, limitation engenders compromise. Balancing across groups is generally more important, however; Chandrasekaran and Jain (1979), for example, noted performance degradation in (parametric) statistical classifiers resulting from unequal sample sizes. Additionally, the practical limits imposed in certain behavioral tests (e.g., radial arm water maze trials have a 60 sec time-limit), while based upon established performance norms (e.g., "typical" response of test subjects), constitute data censorship and require special analytic consideration. Gehan (1965) and Breslow (1970), for example, advocate distribution-free (nonparametric) statistical tests for comparing groups subject to arbitrary right-censorship, based on Wilcoxon (two-group) and Kruskal-Wallis (multiple-group) methods, respectively. A modified Gehan-Breslow analysis was used by Alkon (1974) in studies of photoperiod-mediated associative learning in the nudibranch mollusc *Hermissenda crassicornis*, subject to time-limits (i.e., right-censored). Neurobehavioral researchers, however, generally consider time-limit as a procedural issue and utilize standard (M)ANOVA or Kruskal-Wallis methods for comparing groups (e.g., King et al., 1999; King and Arendash, 2002a).

Expediency in research often has a price; practitioners who take short-cuts with experimental design and data analysis, regardless of motivation, must proceed with extreme caution. Inappropriate selection of repeated-measures ANOVA instead of nonlinear mixed-effects models for evaluating pharmacologic dose-reponse effects in rats (Kristensen and Hansen, 2004), for instance, disregards the nonlinear response characteristics typical of physiological phenomena (e.g., Peek et al., 2002; Kristensen and Hansen, 2004). Additionally, violations of the underlying assumptions of statistical tests can interfere with interpretability or, at worst, compromise the validity of the results (Montgomery, 1953; Boneau, 1960). In practice, however, the fundamental assumptions of analysis of variance – independence of cases, normality of distributions, homoscedasticity (homogeneity of group variances), and sphericity (homogeneity of paired-difference variances) – can sometimes be violated without serious consequences. A series of preliminary tests (and correction methods) can be used to detect (and ameliorate) potential violations, for example: Kolmogorov-Smirnov test (for normality; Lilliefors, 1968), Bartlett test (for homoscedasticity; Bartlett, 1937, 1947), Greenhouse-Geisser correction (for sphericity), and Bonferroni correction (for multiple comparisons). In addition, one of the best, albeit least sophisticated, techniques for identifying potential problems is to inspect an appropriate graphical depiction of the data, such as a histogram or scatterplot; cursory examination of the raw data, in a suitable format, is often the most efficient manner for detecting anomalies and outliers.

Data transformations (e.g., Box and Cox, 1964; Scott and Wild, 1991; Bland and Altman, 1996a) are sometimes performed prior to analysis, particularly when parametric statistical tests are planned, to correct issues which would otherwise complicate or invalidate the results. Transforming data through linear arithmetic scaling (e.g., dividing each value within a column by the variance of the same col-

umn, to produce unit variance), for instance, changes the mean and variance of the column (relative to the original), but leaves the correlation coefficient (Pearson's product moment correlation coefficient) between transformed columns unchanged (Butler, 1982). This intermeasure correlation-preserving transformation is particularly useful for standardizing data for neural networks, which often converge sooner with "zero-mean, unit-variance" data (Haykin, 1999). Logarithmic transforms (e.g., Montgomery, 1953; Martone et al., 1984; Tranel et al., 1994; Bland and Altman, 1996b) are useful for correcting heterogeneity of variance, and often used for right-skewed positive-valued data (large values, e.g., response latencies) and to convert multiplicative main effects into additive effects; this approach will not work with negative values, and very small values (<1) should be pre-multiplied by a large constant. Square root transforms are often used with count data (e.g., bacterial colonies on plates), and when the variance is proportional to the mean; the square root transform is also useful for normalizing Poisson-distributed data. When data are expressed as proportions (percentages) and, in particular, when these data are binomial, the arcsine or arcsine(square-root) transform is often suggested. Results can be back-transformed using the inverse operation, for subsequent interpretation and reporting. Additionally, rank-transformation of data (i.e., replace each value with its rank) followed by ANOVA is sometimes used when the original data neither satisfy the assumptions for standard ANOVA, nor meet the criteria for nonparametric analysis (Conover and Iman, 1981), although this transformation can compromise testing for effect interactions (Seaman et al., 1994).

A classifier represents a mathematical model of empirical data (measured variables) which is intended to split (partition) the data into two (or more) groups. However, partitioning one set of (linearly separable) multidimensional data from another requires a separator with the same number of dimensions (e.g., in one di-

mension, a point; in two dimensions, a line; in three dimensions, a plane; in four or more dimensions, a hyperplane); this is the geometric basis of linear classification. Many sets of multidimensional data are not linearly separable, but can be partitioned using nonlinear classifiers, such as neural networks. Hence, for any given dataset, there are likely to be several alternative classifier architectures to be compared. A common method for evaluating these models involves partitioning the dataset into three subsets by random assignment of individual cases: training data (which will be used to construct the initial classifier model), testing data (used to refine the model to improve generalizability), and evaluation data (for simulating real-world application of the model using unfamiliar data). The relative proportions of these subsets varies with the application and the practitioner. However, in situtations where data are scarce or when critical instances must necessarily be included, cross-validation techniques are used (e.g., Stone, 1977; Goutte, 1997). Threefold cross validation, for example, involves splitting the dataset into three parts – two training sets and one testing set – and repeating the evaluation protocol three times, thus allowing all combinations of testing and training subsets to be assigned and evaluated. Leave-one-out (k-fold) cross-validation ("jackknifing") is a more rigorous approach (computationally-intense), in which a model is constructed using all but a single case, which is subsequently used to evaluate the model, and the construction-evaluation process repeated k-times (i.e., for each of the cases in the dataset). Another approach, called "bootstrapping," generates training and testing sets using sampling with replacement from the dataset (e.g., Efron, 1983). Jackknifing, however, tends to produce relatively unbiased results, compared with resampling techniques, particularly for small sample sizes (Lance et al., 2000).

Comparing the performance of multiple classifiers requires several parameters calculated from the output generated by each model (e.g., Altman and Bland, 1994). A

classification matrix, or "confusion table," for example, often provides useful performance data; the rows of the matrix depict actual (true) categories, while the columns depict the (classifier-) predicted categories. The "true-positive" (TP) and "true-negative" (TN) totals represent, respectively, the number of correctly-categorized "positive" (e.g., transgenic, treated) and "negative" (e.g., non-transgenic, control-group) case instances. The "false-negative" (FN) total indicates how many true-positive individuals are misclassified as "negative" and, correspondingly, the "false-positive" (FP) total shows how many true-negative individuals are categorized as "positive."

The overall success rate, or accuracy, of the classifier is the ratio of correctly-classified cases to the total number of cases. Mathematically,

$$Accuracy \quad = \quad \frac{TP + TN}{TP + TN + FP + FN}$$

The sensitivity is the proportion of true-positive cases which are identified as positive. In diagnostic terms, sensitivity is the probability that the test is positive given that the patient has the disease. Expressed as,

$$Sensitivity \quad = \quad \frac{TP}{TP + FN}$$

By contrast, the specificity is the proportion of true-negative cases correctly recognized as negative or, clinically, the probability that the test is negative given that the patient does not have the disease. Computed as,

$$Specificity \quad = \quad \frac{TN}{TN + FP}$$

Hence, in a study comparing spatial reference memory of transgenic (Tg, treatment group) and non-transgenic (NT, control group) animals: the higher the accuracy, the more Tg and NT animals correctly assigned into their respective groups; the higher the sensitivity, the more Tg animals recognized as Tg (and not misclassified as NT); the higher the specificity, the more NT animals identified as NT (and not misclassified as Tg). In terms of hypothesis testing (Neyman and Pearson, 1933/1967), false-positive outcomes represent type I errors (i.e., the rejection of a true null hypothesis), while false-negative outcomes represent type II errors (i.e., failure to reject a false null hypothesis). Classifier performance is sometimes evaluated using a receiver operating characteristic (ROC) curve (Hanley and McNeil, 1982; Zweig and Campbell, 1993; Obuchowski, 2003; Lasko et al., 2005; Fawcett, 2006), which depicts sensitivity, or "true positive rate," along the ordinate against (1 - specificity), or "false positive rate," along the abcissa, for different values of a discriminative parameter (e.g., the cutoff-value for a classifier, test result criterion). This method is useful for fine-tuning individual classifiers and for testing component classifiers to be combined into ensemble models (Swets, 1988; Swets et al., 2000). An additional statistic, Kappa, representing the proportion of successfully-classified cases corrected for chance performance, is sometimes reported, as well (Cohen, 1960; Fleiss, 1971; Viera and Garrett, 2005) for comparing the performance of multiple independent classifiers.

3.2 Statistical Analysis

Conventional analytic protocols for behavioral experiments generally involve one or more of the following: examining relationships between manipulations (e.g., levels of an independent variable) and observable/measurable phenomena (e.g., dependent variables), for example, dose-response patterns or treatment effects; identifying as-

sociation(s) between/among dependent measures, for example, correlation between behavioral deficits and histopathological markers; characterizing clinical syndromes in terms of behavioral response patterns, for example, motor deficits in Parkinsonism; classifying or categorizing subjects in terms of behavioral response patterns, for example, distinguishing between normal-aged and Alzheimer's patients using cognitive assessment inventories. Multivariate statistical techniques (Tabachnick and Fidell, 2001) are currently the standard approach in these investigations for identifying, testing, measuring, and reporting experimental group differences (e.g., analysis of variance, or ANOVA), association (e.g., correlation), latent patterns in datasets (e.g., factor analysis), and classification (e.g., discriminant analysis). The availability of powerful computing technology and ready-to-run "canned" software (e.g., SYSTAT^{TM}, Statistica^{TM}, R^{TM}) enables any researcher to perform complex analyses easily and efficiently.

This section presents and discusses three of the most important statistical methodologies for neurobehavioral research: correlation analysis, factor analysis, and discriminant analysis.

3.2.1 Correlation Analysis

Correlation analysis measures the linear association between two (or more) variables. This relationship is expressed as the "correlation coefficient" (Pearson's product-moment correlation coefficient), r, which is a real-valued number between -1.0 and +1.0. Mathematically, this value is calculated by dividing the covariance of two variables by the product of their standard deviations. The magnitude (i.e., absolute value) represents the strength of the association, with increasing value indicating greater association between two measures. Negative values denote pairs of variable which change in opposite directions, i.e., one variable increases as the other

decreases; positive values indicate change in the same direction, i.e., both variables either increase or decrease. Independence between two variables is represented with a correlation coefficient of zero. It is important to emphasize that correlation only reflects linear association between two measured variables; correlation does not imply a causal relationship between the variables. In addition, correlations which are statistically significant are not necessarily meaningful, in a practical sense. Finally, when multiple, simultaneous tests of correlation are performed on a dataset, the resultant significance values (p-values) should be adjusted (e.g., Bonferroni's correction) prior to interpretation.

Correlation analyses are frequently used to explore complex relationships among behavioral and/or pathologic measures in experimental paradigms. For example, extensive inter-task and intra-task correlations were identified in a multimetric behavioral assessment battery (Arendash and King, 2002; Leighty et al., 2004). Significant intertask correlations were found between sensorimotor tasks (e.g., Balance Beam and String Agility), cognitive tasks (e.g., Y-Maze and Circular Platform), as well as between sensorimotor and cognitive tasks (e.g., Open Field and Circular Platform) in the age-matched nontransgenic-control and Tg2576 Alzheimer's transgenic mice (Arendash and King, 2002). Similarly, significant intertask correlations between sensorimotor tasks (Open Field and Y-Maze Entries) and cognitive tasks (Morris Water Maze, Circular Platform, Platform Recognition, and RAWM), as well as between sensorimotor and cognitive tasks (Open Field and RAWM, Balance Beam and RAWM) were observed in four Alzheimer's transgenic mouse lines (Leighty et al., 2004). In addition, performance measures showed significant intratask correlation within both the Circular Platform and Morris Water Maze tasks (Arendash and King, 2002) in Tg2576 mice. Likewise, significant intratask correlation was found in the Circular Platform, Morris Water Maze, Platform Recognition, and RAWM tasks completed by

four Alzheimer's transgenic mouse lines (Leighty et al., 2004). Subsequent histological examination of brain tissues from Tg2576 transgenic mice using synaptophysin immunostaining revealed that increased hippocampal staining is associated with both sensorimotor (Balance Beam) and cognitive (Morris Water Maze) performance deficits in the Tg2576 mice (King and Arendash, 2002b). In addition, the Y-Maze entries measure was significantly negatively correlated with Congophilic (compact) beta-amyloid levels in both cortex and hippocampus of Alzheimer's transgenic mice (Leighty et al., 2004); Platform Recognition and RAWM (T4L, T5L) measures were also significantly (positively) correlated with both cortical and hippocampal compact beta-amyloid. Total cortical and hippocampal beta-amyloid (6E10 immunostaining) levels were significantly correlated with Balance Beam (negatively), and Morris Water Maze acquisition, Platform Recognition, and RAWM (all positively) (Leighty et al., 2004).

Significant positive correlation between alternative psychometric batteries (Fillenbaum et al., 1987) used for clinical diagnosis of Alzheimer's disease reflects the similarity of cognitive abilities assessed by these instruments. Indeed, multitask-multimetric cognitive assessment typically results in positive manifold correlation structure, i.e., subjects typically retain rank-ordering across problem-solving tasks, which has been taken as evidence of a single underlying cognitive construct since Spearman's (1904) elucidation of a "general intelligence" factor. Additionally, post-mortem studies in Alzheimer's patients showed significant correlation between regional (frontal cortex, hippocampus) cholinergic activity and cognitive performance (MMSE and MRDS tests) measured within a year of death (Pappas et al., 2000), underscoring the diagnostic utility of behavioral (psychometric) assessment.

3.2.2 Factor Analysis

Factor analysis is a multivariate data reduction technique for identifying underlying associations among measured variables and using these patterns to generate a smaller ensemble of composite variables (called "factors"). Exploratory factor analysis, the most common methodology, attempts to extract the latent associative structure among variables without using prior theory, allowing the practitioner to interpret the resulting factor loadings matrix. Confirmatory factor analysis, by contrast, determines whether the number and composition of factors extracted from a given dataset conform to a pre-established (usually theoretical) expectation. The "extraction" of factors is an iterative optimization procedure in which variance-maximizing linear combinations of the measured variables are generated, the resulting variance is subtracted and a second linear combination generated, and so forth; this approach is called "principal components analysis," and results in orthogonal (i.e., uncorrelated) factors. Alternative extraction algorithms, such as "principal factor analysis," can be used to determine the minimum number of factors which account for the common variance of a dataset (e.g., Velicer and Jackson, 1990). The factor loadings matrix (component loadings, in principal components analysis) represents the correlation coefficients between factors (columns) and component measure variables (rows); hence, these values reflect the amount of each component's variance explained by each factor on which it loads. Frequently, when the initial component loading patterns are difficult to intuit, additional mathematical processing ("rotation") is performed to improve interpretation; Varimax rotation (Kaiser, 1958), for example, generally produces a factor matrix in which each component variable is associated with a single factor (instead of loading on several factors).

Factors extracted are not necessarily factors interpreted; the practitioner is ultimately responsible for making sense of each factor based, in part, upon expectation and past experience (e.g., Fabrigar et al., 1999). Researchers use a collection of heuristics for interpreting factor analytic results, supported by theoretical, empirical, and sometimes anecdotal evidence (Zwick and Velicer, 1986; Lance et al., 2006). The eigenvalue (or "characteristic root") of each factor reflects the proportion of overall variance (of all measured variables) which is accounted for by that factor; eigenvalues greater than unity (1.0), for example, are often interpreted as significant (Guttman, 1954; Kaiser, 1960). Factors having fewer than three high-magnitude component loadings, for instance, should not be interpreted (Velicer and Fava, 1998). In addition, contrary to common practice, larger sample sizes alone are insufficient to improve factor pattern resolution (Guadagnoli and Velicer, 1988; MacCallum et al., 1999), although the number of cases (samples) must always exceed the number of factors. However, Monte Carlo simulations have shown that factor structure and the magnitudes of loadings can influence factor scores, particularly in the case of small samples (Grice, 2001). Finally, factor structure can be altered (sometimes profoundly) by the inclusion of additional behavioral measures within a single task. Rodgers and Johnson (1995), for example, identified two factors ("anxiety" and "locomotor activity") using the standard spatiotemporal measures (i.e., frequency and duration of open- and closed-arm entries) in the elevated plus-maze task. However, a third factor ("decision making") emerged when the duration of time spent in the maze center was included in the analysis. Indeed, these examples underscore the popular maxim that interpreting factor analytic results is as much an art as a science.

The first practical application of factor analysis involved intelligence testing during the early 20th century, when Charles Spearman (1904) postulated a common

underlying "trait" for human intelligence to explain observed individual correlations between cognitive tasks. Indeed, a single construct of general cognitive ability may explain competence across a broad range of areas (academic, vocational, etc.) (Kuncel et al., 2004). Many behavioral researchers, however, have questioned the validity of a single-factor cognitive theory and subsequently developed more complex models comprised of multiple distinct factors, each representing an unique facet of cognition. Thurstone (1938), for instance, identified seven "primary mental abilities" (i.e., associative memory, number facility, perceptual speed, reasoning, spatial visualization, verbal comprehension, and word fluency) by using multiple mental assessment inventories. Human behavioral and psychological characteristics are also identified through factor analysis. For example, four factors were extracted from a comprehensive psychological assessment of Alzheimer's outpatients (N=435), accounting for 57% of total variance (Mirakur et al., 2004): "affect" (e.g., agitation, anxiety, depression); "physical behavior" (e.g., apathy, disturbances in appetite and sleep); "psychosis" (e.g., delusions and hallucinations); and, "hypomania" (e.g., disinhibition, euphoria). Factor analysis is used to validate psychometric inventories, as well. For example, factor analytic studies support the five-factor structure of the MMSE. A large, multisite sample of older adults (N=8556; 50-80 y/o), for example, returned an oblique five-factor solution (Jones and Gallo, 2000). The five factors are: Concentration (serial subtraction, backward word-spelling); Language and Praxis (three-step command, read-follow instructions, sentence writing, polygon copying, object naming); Orientation (time- and place-orientation); Memory (delayed-recall); and, Attention (immediate word repetition). Additionally, Baños and Franklin (2002) administered the MMSE within a more heterogeneous population (N=339; 18-87 y/o; psychiatric patients) and also found an oblique five-factor solution: Orientation, Attention-Working Memory, Comprehension-Praxis, Naming,

and Verbal Recall. Factor analysis also reveals lifespan-developmental trends in psychometric measures. For example, the tendency for human intelligence to become increasingly undifferentiated with increasing age (e.g., Stricker and Rock, 1987); factor analysis of verbal, quantitative, and analytical items of the Graduate Record Examination^{TM} shows that, although cognitive factors remain largely distinct, the intercorrelation of the factors increases with age.

Manifold positive correlation across cognitive tasks (e.g., Morris water maze, Hebb-Williams maze, and water plus-maze) has been found in both rats (e.g., Thorndike, 1935) and mice (Matzel et al., 2003; Galsworthy et al., 2005), as well; unrotated principal components analysis returns a first factor which accounts for between 38 and 61 percent of task variance. Additionally, the strength of correlation between tasks increases with increasing problem-solving complexity (Thorndike, 1935). Mouse-based studies (Kolata et al., 2005) also indicate significant covariance of general cognitive ability with short-term (working) memory, but not with long-term (retention) memory, suggesting the existence of separate memory domains (cf. King et al., 1999; Leighty et al., 2004), as observed in humans. The primary factor in mice, therefore, may represent an overall cognitive performance metric (Plomin, 2001; Galsworthy et al., 2005) which parallels the general cognitive ability construct in humans. Factor analysis has been used for dimensional reduction of multimetric performance assessments in mice; the thirteen or nineteen performance measures obtained from a comprehensive behavioral assessment battery, for example, can be reduced to two or three recognizable factors (e.g., "working memory," "locomotion and exploratory behavior"; Leighty et al., 2004). These factors can be used subsequently to characterize group differences in behavior, for example, sex differences in the "anxiety" factor extracted from elevated plus maze data of rats (Fernandes et al., 1999). The complex interplay between cognitive processes, suggestive of underlying neurophysio-

logical connectivity, is also explored through factor analysis. The interdependence of human working memory (e.g., forward or reverse digit span) and complex cognitive tasks (e.g., arithmetic processing, counting), for example, is revealed through shared component loading of psychometric measures (e.g., Miyake et al., 2001). Similarly, the influence of noncognitive elements of mouse behavioral tasks, such as species-specific swimming patterns in the Morris water maze, can be identified through factor loading patterns (Wolfer et al., 1998). Sensorimotor components can introduce variability, analogous to electronic "noise" (Wolfer et al., 1998), obscuring the subtle cognitive manifestations of genotypic differences. Furthermore, because the relative contributions of each component measure of all factors are calculated (called the factor structure), factor analyses are useful tools for developing parsimonious models of complex behavioral phenomena. In addition, factor "scores" (generated for each individual subject from their corresponding performance measures, for instance) can be regressed onto other variables (e.g., pathologic measures) (Pappas et al., 2000).

Factor analyses have also been used with a nineteen-measure comprehensive behavioral assessment battery in Alzheimer's transgenic mice, both with and without pathologic measures (total and compact beta-amyloid in cortex and hippocampus), to explore patterns of association among measures (e.g., Leighty et al., 2004). Measures from the RAWM, Platform Recognition, and Circular Platform (latency) tasks comprise the first factor, accounting for over 32% of variance. Behavioral measures from the Morris Water Maze, Y-Maze (entries), and Circular Platform (errors) loaded independently as the next three factors, each accounting for about 10% of overall variance. When included, all four pathologic measures loaded with the RAWM, Platform Recognition, and Circular Platform (latency) measures on the first factor, increasing the variance component to approximately 35%. These findings underscore both the diagnostic utility of the three water-based tasks (Morris Maze, Platform Recognition,

and RAWM) which share interrelated cognitive domains, and the strong association between cognitive impairment (reflected as performance deficits) and neuropathology (cortical and hippocampal beta-amyloid) displayed in Alzheimer's transgenic mice (Leighty et al., 2004). Furthermore, exploratory factor analyses of behavioral assessments completed at different time-points may provide important insights on patterns of cognitive development and/or progressive impairment. For example, factor-structure changes were observed in APP/PS1 transgenic mice between 4.5-6 months and 15-16.5 months of age (Jensen et al., 2005). At the later time-point, the primary factor was comprised of measures from the RAWM, Morris Water Maze, and Platform Recognition tasks; at the earlier time-point, only the first two tasks were represented. In addition, five distinct significant factors were identified at the later time-point, while only three factors were returned at the earlier time-point. These results suggest a growing dissociation among cognitive domains with increasing age, perhaps reflecting some progressive decoupling of cognitive processes, concomitant with advancing neuropathology.

3.2.3 Discriminant Analysis

Discriminant analysis is a multivariate statistical technique for identifying essential variables which reliably distinguish between (or among) pre-defined groups and predicting group membership for unclassified individuals on the basis of the variables. The procedure generates a mathematical model, consisting of one or more functions (typically linear) which partition individuals into groups using a combination of the measured variables. The "complete" (or "direct-entry") variant attempts to include all the measured variables in the model, while "stepwise" (e.g., stepwise-forward) variants iteratively select and test each measure for inclusion based on variance-optimization, exhaustively adding (and, possibly, removing) measures. Stepwise

methods often generate more parsimonious models (i.e., having fewer variables) and exhibit superior classificatory ability, relative to complete discrimiminant analyses. For example, direct-entry discriminant analysis of a comprehensive behavioral task battery (consisting of either 15 or 32 sensorimotor and cognitive measures) in APPsw (Tg2576), Tau, and non-transgenic mice failed to distinguish among the three groups (Arendash et al., 2004b), however, a stepwise-forward analysis achieved significant discriminability for both sets of measures ($p<.0001$ and $p<.0005$, respectively). A nineteen-measure comprehensive behavioral assessment was unable to distinguish among four Alzheimer's transgenic mouse lines using complete discriminant analysis, although a stepwise-forward analysis significantly distinguished among the lines using only four measures (YM-Alt, YM-Ent, PR-Avg, and RM-B3T4L) (Leighty et al., 2004). Inclusion of total and Congophilic cortical and hippocampal beta-amyloid measures also resulted in significant discriminability among the four mouse lines using stepwise-forward analysis, generating a four predictor-variable model (YM-Alt, YM-Ent, RM-B3T4L, and HP-Cng) as well (Leighty et al., 2004). Hence, the inclusion of pathologic biomarkers with behavioral performance measures may provide supplemental discriminative information for classification and diagnostic evaluation.

Physical measurements, collected from individuals to be classified, are typically employed in discriminant analysis. For example, Fisher's (1936) discriminant function analysis is the classic demonstration of using morphometric features (length and width of sepal and petal structures) to classify individuals into groups (three *Iris* species: *setosa*, *versicolor*, and *virginica*) using a linear combination of the four physical measures. In addition to structural feature-based classification, function-based taxonomies (e.g., behavioral observations or cognitive performance) are also possible. Behavioral phenotyping of inbred mouse strains using the SHIRPA test battery (Rogers et al., 1999), for example, relies on discriminant analysis to classify

individual animals into their respective strain category on the basis of sensorimotor and cognitive task performance. In addition, ^{18}F-fluorodeoxyglucose positron emission tomography (FDG-PET) was used to distinguish between age-matched groups of AD patients (MMSE criteria) and healthy subjects, yielding sensitivity of 90+% and specificity of 95% (Habeck et al., 2008). Predicting an individual's future performance using historical data (e.g., graduate school admissions, employment candidate screening) is another common application of discriminant analysis (e.g., Neely, 1977). Discriminant analysis is also useful in clinical diagnosis, for generating reduced-length cognitive assessment batteries. For example, discriminant analysis of a fifteen-measure neuropsychological test resulted in a subset of five measures representing overall cognitive ability which accurately distinguished among normal, mildly, and moderately-to-severely impaired geriatric outpatients (Whelihan et al., 1997).

3.3 Data Mining Analysis

Data mining is a multi-step process for exploration and recognition within collections of data (Fayyad et al., 1996b). Driven by the ever-increasing volume of data available today, the goal of data mining is to extract, identify and analyze meaningful (i.e., pertinent) information content (e.g., consistent patterns, systematic rules or trends) in the context of extraneous detail. Some of the important considerations for data mining operations include: selecting information-appropriate representation schemes (data structures; e.g., lists, arrays, or graphs); designing and implementing task-appropriate algorithms (e.g., depth-first search, feed-forward backpropagation of error, entropy minimization); and evaluating accuracy (e.g., Are the analytic results correct?), utility (e.g., Does the analysis reveal novel, nontrivial knowledge?), and performance efficiency (e.g., Is there a faster or less-expensive method for ob-

taining comparable results?). Data mining techniques have a wide range of applications (e.g., Fayyad et al., 1996a; Koh and Tan, 2005), including: finance (e.g., optimal portfolio allocation and value prediction), marketing (e.g., identifying consumer grocery shopping patterns), industry (e.g., monitoring plant operations and quality control), and medicine (e.g., risk management and treatment protocol assessment). Furthermore, through the integration of shared databases (e.g., Amari et al., 2002), contemporary neuroscientists can exchange both research data and analytic tools and establish large-scale collaborative networks for studying human brain disorders. The capacity of data mining methodologies for managing and coordinating immense, heterogeneous collections of information makes these research innovations possible.

Data mining techniques useful for scientific research include both supervised-learning (training involves learning a mapping between input items and their corresponding target output; e.g., classifiers) and unsupervised-learning (constructing a model using only the training items, without *a priori* structural knowledge; e.g., data compression, clustering algorithms) approaches for knowledge discovery, such as decision trees, neural networks, support vector machines, and others. Enhanced classifier performance (e.g., greater accuracy, speed, and/or generalizability), for instance, is often achieved by combining several similar (or different) data mining methods to produce *ensemble models* (e.g., Bauer and Kohavi, 1999; Provost and Fawcett, 2001). Multiple-classifier systems can be used to partition complex problems for distributed processing, or to exploit the respective strengths of each individual methodology (Dietterich, 1997; Wolpert and Macready, 1997; Kiang, 2003). Phillips-Wren et al. (2007), for example, compared several data mining techniques (decision trees, neural networks, logistic regression) for predicting medical oncologist visits based on patient data (demographics, insurance-eligibility, medical history), and showed that neural

network-based classifier performance improves when the data are pre-processed using decision trees.

This section describes three common data mining techniques – decision trees, neural networks, and support vector machines – which have been used separately from, as well as in conjunction with, conventional statistical methods for analyzing biomedical data in both experimental and clinical contexts.

3.3.1 Decision Trees

This data mining technique is based on a graph-theoretic depiction of actions and consequences, using the metaphor of a tree with branches and leaves to represent classification rules and case instances, respectively. Each branch corresponds to a decision event, and the degree of arborization (breadth and depth of the tree) reflects the complexity of the classifier. Ideally, decision events (called "splits") should partition the dataset into groups containing a single dominant class, and splitting should continue until a user-established criterion is reached (e.g., all individuals in a group are of the same class, no sample cases remain). Classification using decision trees involves learning through induction (i.e., rule-inference from examples consisting of attributes and a corresponding conclusion), information theory, and the reduction of entropy (uncertainty). The number of bits (decisions) needed to classify individual cases – similar to the guessing game of "twenty questions" – is related to the depth of the tree; the objective is to choose splits in a manner that significantly reduces uncertainty at each step. Consider the example of a multi-task behavioral evaluation of transgenic and nontransgenic mice, for which we are provided with a table listing each animal's performance measure for each task, along with its genotypic identity (transgenic or nontransgenic). Each behavioral measure represents an attribute, and each corresponding genotypic identity represents a conclusion, for each animal (case).

The goal is to construct a tree with a decision event at each branch point (e.g., "Is the animal's string agility score greater than three?"), such that branches terminate at correct conclusions. After generating the complete tree, one can elucidate overall rules, for example, "If an animal averages fewer than five RAWM Trial 5 errors, and has a balance beam latency of at least 45 sec, then it is nontransgenic" and "in the retention-test phase of the Morris maze task, if an animal spends less than 35% of the time in the goal quadrant, then it is transgenic."

This technique has several advantages: decision trees are easily generated and readily understood; both nominal and categorical data can be included in the dataset, missing cases are easily handled as a separate class; and, there is no "black box" (i.e., all observed splits can be described in terms of Boolean logic, as conjunctive/disjunctive rules, instead of complex mathematical formulae). One significant disadvantage of decision trees, however, is their sensitivity to small variations in the dataset. For example, the use of information-gain (entropy reduction) criteria at each split means that when two or more attributes (e.g., RME-T4 and WM-Ret, in the above example) have nearly the same calculated entropy-reducing effect, the addition or deletion of a single case may determine the shape of the entire tree. Examples of computer algorithms for decision tree learning are ID3 (Quinlan, 1986), which is restricted to categorical values, and C4.5 (Quinlan, 1993), which allows continuous-valued attributes.

Decision tree-based classifiers have been used for analyzing large datasets of neurophysiological data to assist medical diagnosis, as well as in the development of behavior-based screening inventories for Alzheimer's. Additionally, decision trees are used in neuropharmacological research for generating drug screening protocols. For detecting epileptiform EEG spikes, decision tree-based classifiers (J4.8) are comparable to visual analysis by trained human experts (Valenti et al., 2006). A bi-

nary decision tree was used to generate an eleven-question, true-false survey called the "Symptoms of Dementia Screener" (Mundt et al., 2000), intended for use by nonclinical personnel and caregivers to screen for probable-AD. The optimal proportion of correctly-identified probable-AD individuals (i.e., sensitivity) was 90.2%, and the proportion of correctly-identified non-demented individuals (i.e., specificity) was 84.6% (Mundt et al., 2000); this instrument performed better than (conventional) MMSE scoring, which demonstrated 78.7% sensitivity and 92.2% specificity. Finally, pharmacologic research on mouse-based models of depression involves multiple behavioral response measures (e.g., tail suspension test, forced swimming test; Bourin et al., 2005). However, because behavioral responses in mice are influenced by strain-dependent differential drug sensitivity, it is often difficult to interpret and compare results of experiments in which different strains and/or drug dosage levels are examined. A decision tree was constructed to identify an appropriate selection sequence for mouse strain(s) (Swiss NMRI, C57B1/6J, or DBA/2) and behavioral task(s) (tail suspension or forced swimming) to evaluate candidate antidepressant compounds (Bourin et al., 2005), based on mechanism of action.

3.3.2 Neural Networks

Artificial neural networks ("neural networks," ANNs; Abdi, 1994; Haykin, 1999) are a class of learning methods inspired by structural and functional features of biological nervous systems. Neural networks consist of a large number of interconnected, independently-operating computing elements whose aggregate activity results in classification, recognition, feature-extraction, and other higher-order computational capabilities; these "emergent collective computational abilities" (Hopfield, 1982) are the hallmark feature of complex self-organizing systems, such as ecosystems, cities, and brains. The first mathematical model of neural computation

based on threshold-logic and all-or-none response was proposed by McCulloch and Pitts (1943), whose "neuron" received excitatory and inhibitory input information through "synapses," performed arithmetic spatial summation, and produced binary output. Later, Rosenblatt (1958) designed a two-layered trainable computing model for pattern recognition, called the "perceptron," which could only solve simple (i.e., linearly-separable) classification problems. Perceptrons cannot solve more complex (e.g., non-linearly separable) problems such as the exclusive-OR logical relation (e.g., Minsky and Papert, 1969). A multilayered-perceptron architecture, however, with at least a single hidden layer (and sufficient computing elements) becomes a universal function approximator which can describe any continuous function (Kolmogorov, 1957; Cybenko, 1989); an ANN with two hidden layers can describe any arbitrary function. Subsequently, many different neural network computing architectures have been developed, each utilizing its own optimization paradigm (e.g., thermodynamic metaphor in Boltzmann machines; Ackley et al., 1985) and information representational scheme (e.g., self-organizing maps; Kohonen, 1982). The ability of neural networks to construct nonlinear mappings between independent and dependent variables (e.g., between behavioral measures and treatment groups) exemplifies their utility in diverse real-world applications where the relationships between (or among) measured variables are either complex or unclear (Zhang, 2000; Almeida, 2002). Hence, neural networks are useful for constructing dynamical models of ensemble systems having multiple interacting components.

Neural networks are robust (highly fault- and noise-tolerant), and particularly useful for handling incomplete or inexact data. Training ANNs, however, is computationally intense and requires fast processing machines for efficient operation in real-world applications. Untrained neural networks contain no *a priori* knowledge of the learning problem (e.g., the initial matrix of interneuronal connection weights

contains random numbers), hence, trained ANNs sometimes converge to different internal stable weight-matrix configurations – either spontaneously, or in response to small induced perturbations (e.g., input and/or weight noise; Pendharkar, 2002). Consequently, leave-one-out (k-fold) cross-validation is recommended in addition to performing and comparing multiple training sessions (Cunningham et al., 2000). Furthermore, because ANNs make no assumptions concerning the underlying distribution of data (i.e., nonparametric), their performance is relatively immune to departures from statistical normality (of distributions) and linearity (of relationships between variables). However, because no specification of an explicit relationship between predictor variables and outcome is required or – indeed – necessary for constructing and training neural networks, there is currently no prescribed approach for assembling optimal network topologies (e.g., number of layers or hidden-layer units). Instead, practitioners often rely on experience, heuristics, and trial-and-error experimentation for guidance (Walczak and Cerpa, 1999), e.g., using the geometric mean of the numbers of input- and output-layer elements to assign the number of hidden-layer units.

Although ANNs are normally implemented as software (computer programs) executed on high-performance hardware platforms, a geometric metaphor provides a useful means of description. One standard architecture for supervised learning, called a feed-forward network, is comprised of finite sets of computing elements (also called "neurons" or "computing units") arranged into distinct layers, with each layer fully-connected to the preceding and succeeding layers (i.e., all elements are pairwise-connected between layers) by weighted links. The first layer, called the "input layer," receives the stimulus data (e.g., case instances used for training and testing), and the final layer, called the "output layer," returns the final computed result generated by the network; layers of elements, positioned between the input and output layers,

are known as "hidden layers." Hence, information provided to the input layer is processed within the elements of that layer, then passed forward to the second layer for processing within elements of the second layer, and proceeding across all layers until the elements of the output layer receive and process the information. Training a network necessarily requires structural changes, which are represented by adjustments of the connections (i.e., weights) between elements. The magnitude of the weight adjustments, as well as the choice of connections to be adjusted, is determined by comparing the desired output of the network (recall, this is supervised learning) with its actual output and applying an appropriate learning algorithm. The back-propagation training algorithm (Rumelhart et al., 1986), for example, realizes the principle of Hebbian learning (Hebb, 1949), by iteratively adjusting the connection strengths (weights) between successive layers of the network, in retrograde fashion starting from the output layer, to minimize the discrepancy (error) between the desired and actual outputs of the network. Training involves a compromise between memorization and generalization (Almeida, 2002), i.e., between acquiring a perfect mapping between training examples and corresponding output, and acquiring sufficient association between input and output patterns for adequate handling of novel, testing examples; stopping criteria (e.g., number of training epochs, asymptotic error analysis) are used to determine how long to continue training.

Artificial neural networks represent a promising new computational tool for both clinical and experimental data analysis, as well as for medical decision-making (e.g., Baxt, 1995; Papik et al., 1996). Indeed, diagnoses generated by neural networks are generally comparable, and sometimes superior, to those made by (human) experts (e.g., Miller et al., 1992; Tafeit and Reibnegger, 1999). For example, a review of clinical and randomized-control trials (Lisboa and Taktak, 2006) involving neural network adjuncts in cancer diagnosis and prognosis showed a significant benefit (21

out of 27 studies) by including these methods with conventional medical assessment. Similar diagnostic advantages are reported for the prediction of mortality in cardiac surgery patients (Nilsson et al., 2006), dynamic identification of gait pathologies using kinematic data (Schöllhorn, 2004), and evaluation of treatment efficacy of combination HIV therapy (Larder et al., 2007). In experimental settings, neural networks are also used for analyzing real-time data in computer automated studies involving rodents. For example, in rats, to classify behavior using Fourier-transformed, digitized images (Heeren and Cools, 2000; Rousseau et al., 2000) or to identify sleep stage through EEG recordings (Robert et al., 1997). Human- and neural network-based evaluation of rat sleep stage (waking, paradoxical sleep, non-REM) using a single parietal-occipital EEG channel were in agreement over 90% of epochs examined (Robert et al., 1996). In mice, neural networks have been used to distinguish among five different species using both frequency- and time-domain characteristics of recorded vocalizations (Tian and Shang, 2006). Surprisingly, a three-layered network (257 x 89 x 5) which included potentially extraneous signal input outperformed a four-layered topology (33 x 35 x 30 x 5), which was supplied with a smaller subset of power spectrum data, with respect to overall classification (approximately 95% versus 45%) (Tian and Shang, 2006); the seemingly extraneous signal characteristics provided to the first network evidently provided important discriminative details for the classifier. Indeed, the performance of neural network-based classifiers is often strongly influenced by the number of attributes (e.g., behavioral/cognitive measures, recorded signals) supplied as input to the network (Leighty et al., 2008).

In addition to experimental and medical applications, artificial neural networks have demonstrated utility for psychological and behavioral assessment (Price et al., 2000). For example, a neural network-based classifier using coded responses from parent interviews showed high discriminability (92% accuracy during training) and

generalizability (92% correct-assignment of test cases) for autism (Cohen et al., 1993). Nair et al. (1999) used eleven depression-related symptoms to identify clinically-depressed (DSM-III) individuals. Subsequent contribution analysis (which measures the impact of each symptom on overall network-produced diagnosis) of the trained neural network identified different patterns of symptom "importance" between depressed and non-depressed individuals (Nair et al., 1999), although "sadness" was the strongest individual predictor. The prediction of psychiatric treatment outcome, as well, is a practical application of neural networks. Serretti et al. (2007) combined multiple predictors of antidepressant response in patients with major depression, obtaining a correlation of .46 between predicted and observed response (i.e., explaining 21% of variance); this example underscores the ability of neural networks to combine diverse forms of predictor variables (e.g., demographic, individual medical and psychiatric history, known therapeutic response parameters). In addition to cross-sectional studies (i.e., sampling conducted at a single time-point), longitudinal analyses (i.e., sampling at multiple time-points) are useful performance metrics for disorders with nonlinear pathological trajectories, such as Alzheimer's disease (Tandon et al., 2006).

Neural networks have theoretical significance, as functional models of biological nervous systems, for exploring normal and pathological cognitive processes as well as complex systems, machine learning and artificial intelligence. ANNs, as examples of parallel distributed processing, are used to study human learning and problem solving (Rumelhart and McClelland, 1986) and are the foundation of "connectionist" architecture models in cognitive science (e.g., for modeling the Stroop effect; Cohen et al., 1990). Despite the ANN's capacity for fault-tolerance, a sharp decline in operating performance associated with external disruption or disconnection of connections between computing elements may occur when up to 70% of connections are

compromised (Kalampokis et al., 2003); this critical threshold represents the onset of "pathology" in the ANN, reminiscent of incipient neuropathological degeneration of the brain. In addition, computational models of episodic memory have been based on autoassociative neural networks (e.g., Hopfield nets, 1982), and recognition memory has been studied using backpropagation in "encoder problem" learning architectures (Ratcliff, 1990).

3.3.3 Support Vector Machines

Support vector machines (SVMs) are a learning methodology for performing binary classification (Burges, 1998; Cristianini and Shawe-Taylor, 2000; Noble, 2006). SVMs originated in Vladimir Vapnik's (Vapnik and Lerner, 1963; Vapnik and Chervonenkis, 1964; Boser et al., 1992; Cortes and Vapnik, 1995; Vapnik, 1998) research on learning theory, postulating a two-stage process for constructing abstract partitions between two groups of training examples. In the first stage, the input data (represented as numerical vectors) are mathematically mapped into a higher-dimensional "feature space" (extending the conventional two- or three-dimensional space). The mapping operation requires an appropriate transformation (known as a "kernel function"), which must be selected in advance (e.g., polynomial, radial basis function). In the second stage, a decision surface (called a "hyperplane") is calculated which maximizes separation between the two groups. At the end of training, the two groups will occupy distinct regions of the feature space. Like neural networks, SVMs are very robust (noise-tolerant), highly resistant to overfitting, and capable of nonlinear classification.

SVMs have become a standard tool in bioinformatics research for classifying gene expression profiles. The expression patterns of candidate genes are classified through comparison with patterns of known co-regulated groups of genes (Gaasterland and

Bekiranov, 2000). Golub et al. (1999), for example, distinguished between individuals with acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) using SVMs trained with Affymetrix microarray data (6817 genes) from 38 bone marrow samples (ALL, N=27; AML, N=11). Biological sequence analysis is another important bioinformatics application for support vector machines. Prediction of proline *cis*/*trans* isomerization (Song et al., 2006), for example, using adjacent amino acid sequence information from a protein database, along with multiple sequence alignment (PSI-BLAST profiles) data; radial basis function kernels outperformed both polynomial and linear kernel functions in the support vector machines tested.

SVMs also exhibit great facility for handling the voluminous datasets of measurements sampled during physiological and behavioral experiments in rodents. Crisler et al. (2008), for example, used a support vector machine-based classifier for sleep-staging in rats. Electrophysiological data were recorded (three electrodes: frontal cortex, parietal cortex, and temporalis muscle), filtered (low-pass, 90Hz), and digitized to generate time- and frequency-domain datasets, which were subsequent analyzed by support vector machines with a radial basis function kernel. Scoring results between the SVM-based classifier and human experts were in 96% agreement (Crisler et al., 2008). Following treatment with different antidepressants, behavioral recordings (time spent immobile, swimming, and struggling) of rats during the forced swimming test were analyzed by SVMs to distinguish among drug classes (Frohlich et al., 2008). Untreated control animals were distinguishable from rats receiving either tricyclic antidepressants (imipramine, 88% accuracy; desipramine, 83% accuracy) or selective serotonin reuptake inhibitors (fluoxetine, 87% accuracy); in addition, the SVM-based classifier was able to distinguish between animals receiving the two classes of antidepressants (accuracy greater than 83%) (Frohlich et al., 2008).

SVMs are potentially useful adjuncts for medical diagnosis. Übeyli (2008), for example, distinguished among six classes of dermatologic conditions in 358 patients using a support vector machine-based classifier with a radial basis function kernel. The network configuration consisted of 34 inputs (clinical and histopathological features), six outputs (the diagnostic classes), and nine support vectors. The resulting classifier achieved nearly 100% accuracy, sensitivity, and specificity (Übeyli, 2008), outperforming both multilayer-perceptron and recurrent neural networks. Additionally, the use of SVMs in diagnostic image analysis (e.g., detecting microcalcification clusters in digital mammograms; El-Naqa et al., 2002) underscores the remarkable pattern recognition capacity of this methodology.

3.4 Practical Comparisons

Performance comparisons within and between statistical- and data mining-based classifiers involves several practical and theoretical considerations (Duin, 1996). The particular application, for example, imposes constraints on the selection of classifiers; e.g., known/unknown underlying distributions, sample size effects. The prior experience of the practitioner is an important factor, particularly when confronting design heuristics, e.g., neural network topology or support vector machine kernel function. Recalling that there is no "best" classifier methodology for all problem domains (e.g., Duin, 1996; Wolpert and Macready, 1997), comparisons are necessarily context-specific and perhaps, to some extent, user-specific.

Pattern recognition is a typical domain for comparing discriminant analysis and neural networks (e.g., Holmstrom et al., 1997). For instance, handwritten character identification and phoneme recognition are, respectively, space- and time-dependent pattern classification problems. Visual image stimuli (printed characters) may be coded as normalized matrices of pixel-intensity values, for example, and provided as

input to the classifier engine. Similarly, auditory stimuli (sound fragments) may be encoded as frequency spectra, or as signal processing parameters (e.g., calculated by Fourier analysis, wavelet transform). Parsons and Jones (2000), for example, used temporal and spectral features extracted from digitized recordings of echolocation calls from 14 sympatric bat species as classifier input, resulting in 79% accuracy of classification by discriminant functions and 87% accuracy by artificial neural network.

Statistical- and data mining-based classifiers have been compared in studies involving therapeutic interventions in Alzheimer's transgenic mice, to evaluate both transgenic and treatment effects (Leighty et al., 2008). Multimetric behavioral data from two separate investigations – one study examining caffeine administration in young mice, and a second study involving environmentally-enriched housing conditions – were analyzed. Neural network-based classifiers were shown to be superior to discriminant analysis (both complete and optimized, stepwise-forward, methods) for distinguishing between nontransgenic and Alzheimer's transgenic animals in both studies. The two classifiers performed comparably, however, with respect to distinguishing treatment effects in transgenic animals (Tg control vs. Tg treatment) in both studies. In addition, the classifiers performed comparably for the interaction between treatment and transgenicity (i.e., discriminability among all groups). Taken together, these findings suggest that the two classifiers may be suitable complements for neurobehavioral investigation. Furthermore, neural networks may be sensitive to subtle features (e.g., nonlinearity) in datasets which discriminant analysis cannot detect.

Both statistical and data mining methodologies have been applied in Alzheimer's disease research, primarily by psychologists and psychiatrists for developing and refining psychometric screening instruments. For example, Grossi et al. (2007) used four pathological features (neurofibrillary tangle and neuritic plaque measures in

neocortex and hippocampus) to distinguish confirmed-AD individuals (N=26) and normal-aged individuals (N=36) with discriminant analysis and neural networks. The linear discriminant analysis correctly identified 92.3% of individuals, while the neural network correctly identified all the individuals. Subsequent relevance analysis identified the neocortical neurofibrillary tangle measure as the most important predictor variable (Grossi et al., 2007). In addition, French et al. (1997) used the results of eleven neuropsychological tests to distinguish between probable Alzheimer's disease and normal-elderly individuals with linear discriminant analysis and neural networks. The neural network was more accurate than the discriminant analysis (91.1% versus 71.9%) overall, and better able to distinguish levels of severity within the AD individuals (French et al., 1997). Indeed, neural networks are powerful techniques for combining heterogeneous predictor variables in AD diagnosis. For example, Garcia-Perez et al. (1998) trained a three-layered feedforward-backpropagation neural network using 46 measures (demographic, medical history, psychometric assessment, electrophysiological, and diagnostic imaging) from 35 individuals with either clinically-diagnosed vascular dementia (N=19) or AD (N=16). Subsequent testing using a training set of demented individuals (N=23) resulted in 82.6% accuracy of classification. The network configuration consisted of 46 input layer elements, 29 hidden layer units, and a single output unit; the learning rate and momentum parameters were both set to 0.1, connection weights were initialized to 0.3, and the stopping criterion (error tolerance) was set to 0.0000002 (Garcia-Perez et al., 1998).

3.5 "What's Past Is Prologue"

> Whereof what's past is prologue, what to come
> In yours and my discharge.
>
> William Shakespeare (1564-1616)
> *The Tempest*, Act II, Scene i, Lines 253-254.

The preceding chapters were intended to establish the practical foundations for conducting behavioral research using a transgenic mouse model of Alzheimer's disease, including selection of an appropriate animal model and behavioral/cognitive assessment task battery, as well as to introduce statistical and data mining-based approaches for neurobehavioral data analysis.

In contrast, the following chapters illustrate specific applications for examining both transgenicity and therapeutic treatment effects in nontransgenic and Alzheimer's transgenic mice, comparing statistical and data mining-based analytic techniques using a semantic interference testing protocol in human Alzheimer's patients, and for evaluating a novel, mouse-based cognitive assessment paradigm inspired by the semantic interference protocol.

Each application is introduced with a brief statement of background motivation, followed by a complete description of the materials and methods utilized (e.g., subjects, treatment and assessment protocols). Next, a detailed presentation of statistical and data mining analytic results is provided. The presentation is completed by a discussion of key findings, emphasizing significant contributions to Alzheimer's research.

CHAPTER 4

CAFFEINE ADMINISTRATION IN NONTRANSGENIC MICE

4.1 Introduction

Does life-long consumption of caffeine provide any therapeutic benefits in normal, healthy individuals, with respect to cognitive ability? That is, does caffeine consumption significantly prevent, postpone, or ameliorate normal age-associated cognitive impairment?

The effects of long-term caffeine consumption in normal humans have been investigated through retrospective studies on coffee consumption (e.g., Chou, 1992; Rosso et al., 2008). Approximately one-half of the American population consumes coffee on a daily basis, averaging over two cups per day (equivalent to 200 mg of caffeine, or 2.4 mg/kg/day for adults) (Chou, 1992). Despite widespread anecdotal evidence of purported adverse systemic effects (e.g., anxiety, insomnia, hypertension, coronary heart disease, pancreatic cancer, reproductive system-related disorders) associated with chronic caffeine consumption, there is no clear evidence of a link between modest coffee consumption and human disease (Chou, 1992; Smith, 2002). Indeed, self-report studies (Smith, 2002) suggest an association between improved mental functioning (sustained attention, overall alertness) and regular coffee consumption in normal adults. Among elderly women, but not men, increased long-term coffee consumption is associated with better mental performance (MMSE, visual short-term and long-term memory, verbal category fluency) (Rancho Bernardo Study; Johnson-

Kozlow et al., 2002), suggesting a cognitive neuroprotective effect of caffeine (Rosso et al., 2008). Similarly, a community-based French study (Ritchie et al., 2007) reported relatively smaller age-associated declines in verbal retrieval and visuospatial memory ability with increased long-term coffee consumption in elderly women, compared with age-matched non-consumers.

In mice, relatively few studies have addressed the long-term effects of caffeine administration in rodents, despite an abundance of studies on the acute effects of caffeine in mice (e.g., Izquierdo et al., 1979; Angelucci et al., 1999). For example, in a study involving adult rats receiving caffeine orally through their drinking water (0.3 mg/mL) for four weeks, caffeine-treated animals displayed impaired Morris water maze acquisition (i.e., longer latency) during the first trial, but similar latency on subsequent acquisition trials, relative to untreated rats (Hun et al., 2007). Additionally, although all rats performed comparably on a probe trial (Morris water maze retention) administered one week after acquisition, the caffeine-treated animals performed significantly poorer on probe trials administered two- and three-weeks after acquisition (Hun et al., 2007), relative to untreated rats. The performance impairment observed in the hippocampal-dependent Morris water maze task by caffeine-treated rats was underscored by immunohistochemical assays, which showed decreased neurogenesis in the hippocampal subgranular zone (Hun et al., 2007), relative to untreated animals. Hence, long-term caffeine administration in normal rats may be associated with functional damage to vulnerable regions of the hippocampus, expressed behaviorally as domain-specific cognitive impairment (e.g., spatial reference memory). By contrast, a study by Arendash et al. (2009) compared cognitive performance between aged (15-16 month-old) nontransgenic mice which received oral caffeine (0.3 mg/mL) for the previous ten months and age-matched untreated animals. As illustrated in FIgure 4.1, cognitive measures obtained from a comprehensive behavioral assess-

ment battery (Figure 4.1A, Y-maze; Figure 4.1B and C, Morris water maze; Figure 4.1D, Circular Platform; Figure 4.1E, Platform Recognition; and Figure 4.1F, Radial Arm water maze tasks) showed no significant differences between caffeine-treated and untreated nontransgenic mice for any of the 18 behavioral measures evaluated (Arendash et al., 2009).

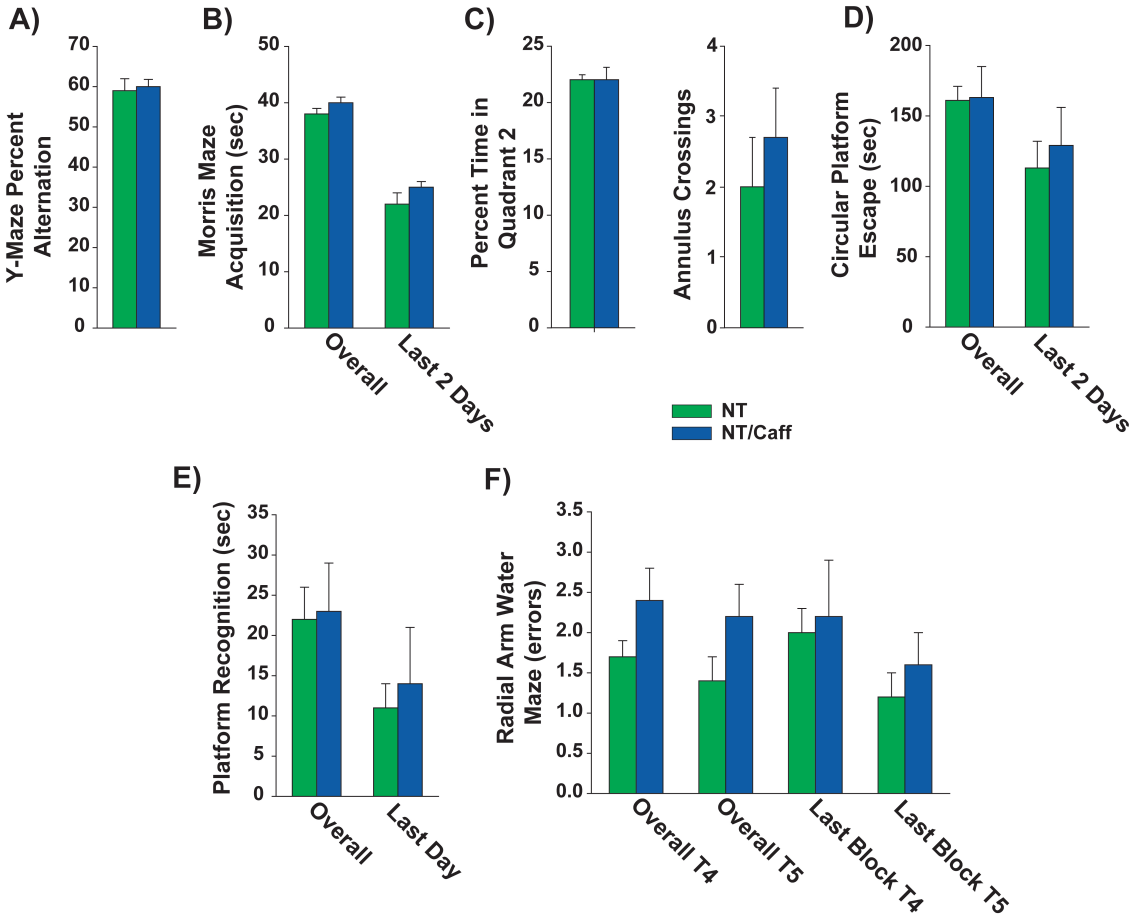

Figure 4.1. Comparisons between untreated and caffeine-treated nontransgenic mice.

Assessing the treatment effect of caffeine in nontransgenic animals provides an important reference benchmark for subsequent evaluation of therapeutic efficacy in Alzheimer's transgenic mice. In addition, the dataset generated from such investigations provides an unique opportunity for comparing conventional statistical analytic

approaches (e.g., ANOVA) with novel data mining-based methods for identifying experimental treatment effects using a multimetric behavioral assessment battery. The purpose of this study is to re-examine the results of the Arendash et al. (2009) study using discriminant analysis and data mining-based classifiers, to determine whether groupwise differences in cognitive performance can be detected through advanced statistical and data mining-based techniques.

4.2 Materials and Methods

A total of nineteen subjects (nontransgenic mice; C57/SJL/SW/B6 hybrid background) were included in the study, randomly-assigned to either of two groups: control (N=11) and caffeine-treated (N=8). Caffeine treatment (provided through drinking water, 0.3 mg/ml) began at five and one-half months of age, and continued until all animals of both groups were tested at 15-16 months of age. Behavioral testing consisted of a nine-task behavioral assessment battery, from which nineteen measures were obtained (refer to Table 2.1 for coding): Open Field (OF), Balance Beam (BB), String Agility (SA), Y-maze (YM-AE, YM-PA), Elevated Plus maze (EP-CE, EP-OE, EP-TO), Morris water maze (WM-Avg, WM-Fin, WM-Ret), Circular Platform (CPE-Avg, CPL-Avg), Platform Recognition (PR-Avg, PR-Fin), and Radial Arm water maze (RAWM; RML-T4, RML-FT4, RML-T5, RML-FT5).

Animal care and use was in accordance with the Guide and Use of Laboratory Animals, National Research Council, 1996, in a program and facilities fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, International, under a protocol approved by the University of South Florida Institutional Animal Care and Use Committee (No. 2729, Gary Arendash, Ph.D., Principal Investigator).

General Analytic Protocol

Pentium^{TM}-class microcomputing hardware platforms were used to execute current versions of both statistical and data mining analytical software. Supplementary software applications were implemented in the C++ programming language (Stroustrup, 1985) for database cross-checking, data preprocessing, and other routine non-analytic tasks. Statistical analyses were performed using the SYSTAT^{TM} (Systat Software, Inc.) software package. Correlation analysis, factor analysis, and discriminant analysis were conducted interactively using default application and system-configured settings, except as indicated. The Pearson product-moment correlation coefficient (r) and its associated significance value (p) was calculated and reported for each significant (i.e., $p<.05$) pairwise comparison between variables. Bonferroni's correction for multiple simultaneous comparisons was applied, as needed, and pairwise-deletion was used to handle missing values. Factor analysis was performed using principal components method, with subsequent Varimax rotation. Discriminant analysis was performed using both complete (direct-entry) method, in which all available predictor variables are used to generate multivariate classification function(s), and stepwise-forward method, wherein a model is iteratively constructed through inclusion/deletion of predictor variables on the basis of individual variance contribution. The "alpha-to-add" parameter for the stepwise-forward approach was 0.15. Classifier performance was evaluated using jackknifing, and the success rate (overall accuracy), sensitivity, and specificity were calculated and reported.

The formulae for calculating accuracy, sensitivity, and specificity are:

$$\begin{aligned} Accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \\ Sensitivity &= \frac{TP}{TP+FN} \\ Specificity &= \frac{TN}{TN+FP} \end{aligned}$$

Where: TP and TN denote the number of true-positive and true-negative cases, respectively, and FP and FN represent the number of false-positive and false-negative cases, respectively. Accuracy, sensitivity, and specificity measures are expressed as percentages. In addition, Wilks's lambda statistic (Wilks, 1932) and associated significance value were reported. Data mining analyses were performed using the Java-based Weka^{TM} (Witten and Frank, 2000) software package. Decision tree-, neural network-, and support vector machine-based classifiers were executed interactively using default application and system-configured settings, except as indicated. Decision tree induction was performed using the J48 algorithm, a Java-based implementation of the C4.5 method (Quinlan, 1993). A multilayer perceptron architecture was used for the neural network-based classifiers, consisting of an input layer (corresponding to the number of predictor/attribute variables), an output layer (corresponding to the number of distinct classification groups), and a single hidden layer. The number of computing elements in the hidden layer was varied between the number of input-layer units and the number of output-layer units, and the corresponding classifier performances (k-fold cross-validation) compared to determine the optimal hidden-layer configuration. The backpropagation learning algorithm was used for training each neural network, using fixed parameters for both learning rate (0.9) and momentum (0.2). The support vector machine implementation used the SMO method and radial basis function (RBF) kernel for classification. Substitution of a normalized

polynomial kernel (NPK), in some cases, resulted in improved classifier performance, where indicated. Multiple executions (program runs) of both neural networks and support vector machines were performed to achieve stable convergence, as well as to avoid spurious (inconsistent) results. The performance of all data mining-based classifiers was evaluated by k-fold (leave-one-out) cross-validation, and the success rate (accuracy), sensitivity, and specificity were reported, in addition to the Kappa statistic (for comparison among classifiers).

4.3 Results of Statistical Analyses

4.3.1 Correlation Analysis

Significant correlations ($p < .05$) observed in pairwise comparisons of all behavioral measures are shown in Table 4.1, wherein marked cells include both the correlation coefficient (r-value, top) and significance (p-value, bottom). Significant intra-task correlations exist within the Elevated Plus maze, Morris water maze, Platform Recognition, and Radial Arm water maze. Significant inter-task correlations are observed between sensorimotor and cognitive tasks: Open Field and Morris water maze (OF / WM-Avg), Open Field and Circular Platform errors (OF / CPE-Avg), and Balance Beam and RAWM latency (BB / RML-FT4, RML-FT5, and RML-T5). Indeed, observed intertask correlations between sensorimotor and cognitive measures underscore the interdependence among sensory, motor, and cognitive behavioral processes. Additionally, significant inter-task correlations exist between cognitive tasks: Y-maze percent alternation and RAWM latency (YM-PA / RML-FT5, RML-T5), Circular Platform latency and Morris water maze (CPL-Avg / WM-Ret), and Platform Recognition and RAWM latency (PR-Fin / RML-T4, RML-T5; PR-Avg / RML-T4, RML-T5). Correlations observed between measures spanning multiple

tasks suggests overlapping or shared cognitive domains across tasks, or dependencies therein. Not surprisingly, for example, spatial memory-biased tasks (Morris water maze, Platform Recognition, and RAWM) exhibit extensive intercorrelation between component measures. The absence of significant correlations between measures from the RAWM and Morris water maze tasks is consistent with factor analyses, which indicated segregation between these tasks. Finally, the multifactorial character of certain tasks was underscored: Significant correlation between the Elevated Plus maze task and both RAWM latency (EP-TO / RML-FT4, RML-T4) and Morris water maze (EP-TO / WM-Ret) suggests an anxiety component to the RAWM and Morris water maze tasks, while an activity component to the Circular Platform task is suggested by the correlation between Y-maze entries and Circular Platform errors (YM-AE / CPE-Avg).

4.3.2 Factor Analysis

An unrotated principal component analysis of the behavioral measures is shown in Table 4.2, with significant factor loadings (absolute value greater than 0.540) indicated. Calculated eigenvalue results (i.e., >1 criterion) were in agreement with observed scree plot (Cattell, 1966) identification of five significant factors. The primary factor (accounting for over 27% of overall variance) was a cognitive-biased structure comprised of the four RAWM latency measures, as well as the Balance Beam, and Platform Recognition measures, consistent with correlation analyses. The second factor was comprised of arm-entry counts (both open- and closed-arms) from the Y-maze, as well as the probe trial (retention) measure from the Morris water maze, reflecting the proportion of time spent in the previously platform-containing quadrant of the pool. Collectively, this factor may reflect elements of persistence or search behavior. Both sensorimotor (open field activity, Y-maze arm entries) and

Table 4.1. Correlations between behavioral measures in the nontransgenic caffeine study

	OF	BB	SA	YM-AE	YM-PA	EP-CE	EP-OE	EP-TO	WM-Avg	WM-Fin	WM-Ret	CPE-Avg	CPL-Avg	PR-Avg	PR-Fin	RML-T4	RML-FT4	RML-T5
BB																		
SA																		
YM-AE																		
YM-PA																		
EP-CE																		
EP-OE						1.00 .000												
EP-TO																		
WM-Avg	.48 .039																	
WM-Fin									.67 .002									
WM-Ret								-.49 .032										
CPE-Avg	.50 .028			.71 .001														
CPL-Avg											.73 .000							
PR-Avg																		
PR-Fin														.92 .000				
RML-T4								.54 .017						.63 .004	.58 .009			
RML-FT4		-.48 .037						.46 .046								.83 .000		
RML-T5		-.54 .018			-.55 .014										.47 .041	.82 .000	.77 .000	
RML-FT5		-.53 .020			-.49 .032											.75 .000	.86 .000	.90 .000

Table 4.2. Unrotated factor component loadings in the nontransgenic caffeine study

	Factor				
Measure	I	II	III	IV	V
RML-T5	0.931				
RML-FT5	0.873				
RML-FT4	0.873				
RML-T4	0.863				
BB	-0.594				
PR-Avg	0.564				
PR-Fin	0.545				
PM-OE		-0.847			
PM-CE		-0.846			
WM-Fin		0.552			
CPE-Avg			0.690		
OF			0.688		
YM-AE			0.599	0.652	
WM-Avg			0.575		
WM-Ret				0.721	
CPL-Avg				0.699	
YM-PA					-0.576
Variance	27.69%	15.14%	14.35%	11.99%	6.87%

cognitive (Circular Platform error average, Morris water maze acquisition) measures are included in the third factor, which may represent a meshing of spatial reference learning and exploratory behavior. The fourth factor, which is comprised of the average Circular Platform latency, number of Y-maze arm entries, and Morris water maze retention measures, may reflect elements of escape behavior. Note that two of these tasks involve escape scenarios – from noxious stimuli to a safe refuge (Circular Platform) or from the water to a dry platform (Morris water maze) – in which the animal must learn to associate experimental cues with safety. The Y-maze arm entries measure (YM-AE) loads on both Factors III and IV, underscoring an exploratory/activity component to other measures in both of these factors. The fifth factor consists solely of the Y-maze percent alternations, which reflects systematic

visitation of all arms in the Y-maze. This factor may therefore represent general mnemonic function as a cognitive domain separate from others (i.e., working memory). Note the absence of Elevated Plus maze open-arm visit latency (PM-TO) and String Agility (SA) measures, indicating relative non-significance of these behavioral components.

4.3.3 Discriminant Analysis

A direct-entry (complete) discriminant analysis did not yield significant discriminability between the two groups (Wilks lambda = 0.084, p = .7703). However, a stepwise-forward analysis produced significant discriminability (Wilks lambda = 0.162, p = .0041) with 79% classification accuracy ("success rate") (Jackknifing method). The classifier exhibited a sensitivity of 75% (i.e., correct indication of treated animals) and a specificity of 82% (i.e., correct indication of untreated animals), with respect to caffeine treatment effect. Indeed, stepwise-forward discriminant analysis surpassed conventional ANOVA-based techniques (Arendash et al., 2009) by disinguishing between untreated and caffeine-treated nontransgenic mice. The eight variables retained by the stepwise-forward analysis included both sensorimotor and cognitive measures: OF, BB, YM-PA, WM-Fin, CPE-Avg, CPL-Avg, PR-Avg, and PR-Fin. The significant discriminability between the two groups achieved through a combination of sensorimotor and cognitive measures suggests that caffeine may exert influence beyond the cognitive domains.

4.4 Results of Data Mining Analyses

4.4.1 Decision Tree Analysis

```
WM-Ret <= 0.117: NT+CAFF (3.0)
WM-Ret > 0.117
|   SA <= 3
|   |   OF <= 98: NT (3.0)
|   |   OF > 98: NT+CAFF (4.0)
|   SA > 3: NT (9.0/1.0)
```

Figure 4.2. Sample decision tree classifier generated in the nontransgenic caffeine study

A decision tree classifier constructed using the J48 algorithm (variant of C4.5) is depicted in Figure 4.2. The classifier attempted to split the dataset using three measures, based on the entropy-reducing (information gain) capacity of these predictors: WM-Ret, SA, and OF. Only one of the measures (Morris water maze retention, WM-Ret) has a strong cognitive bias, while the other two reflect sensorimotor integrative (String Agility) and exploratory (Open Field) abilities. Cross-validation (k-fold) of the optimal tree resulted in only 36% correct classification of cases (Kappa = -0.34). The sensitivity of this classifier was only 13% and the specificity was 55%, with respect to caffeine treatment effect. Therefore, although this was the optimal decision tree-based classifier, it did not successfully distinguish between the two groups of animals.

4.4.2 Neural Network Analysis

A three-layered neural network, trained using the backpropagation learning algorithm (rate = 0.3, momentum = 0.2), was evaluated with the number of hidden-layer units varied between 4 and 12. The optimal performance was obtained for

a network containing nineteen input-layer elements (the behavioral measures), five hidden-layer elements, and two output-layer elements (the two treatment groups). Cross-validation (k-fold) of the optimal network showed 58% correct classification of cases (Kappa = 0.16), with 63% sensitivity and 55% specificity. Although these results are superior to those of the decision tree-based classifier, the discriminability observed is not significant.

4.4.3 Support Vector Machine Analysis

A support vector machine architecture (SMO method) attempted to classify the nineteen cases into their respective groups, resulting in 52% correct classification (Kappa = 0.08) using k-fold cross-validation, with 63% sensitivity and 46% specificity. These results are inferior to those obtained by the neural network-based classifier and, subsequently, represent unsuccessful discrimination between the two groups.

4.5 Discussion

The correlation analysis showed significant intercorrelations between and within behavioral tasks, which is consistent with findings from earlier studies involving multimetric behavioral assessment (e.g., Arendash and King, 2002; Leighty et al., 2004). The factor analysis is consistent with findings from earlier studies, in which the primary factor represents a strong cognitive-loaded component of behavioral performance (e.g., Galsworthy et al., 2005), largely driven by the RAWM task, underscoring the primacy of this task for cognitive assessment (e.g., Leighty et al., 2004; Jensen et al., 2005). The discriminant analysis is also consistent with earlier studies, in which non-significant discriminability observed with direct-entry analysis is contrasted with significant discriminability when the stepwise-forward approach is used (e.g., Arendash et al., 2004b; Leighty et al., 2004). In addition, the presence of

both sensorimotor and cognitive measures in the optimal model suggest that caffeine influences multiple behavioral domains. None of the data mining-based classifiers successfully distinguished between animals which received caffeine and those which did not. Table 4.3 summarizes the performance of all classifiers examined, with respect to sensitivity, specificity, and overall success rate, in distinguishing between caffeine-treated and untreated nontransgenic mice.

Table 4.3. Classifier performance comparison in the nontransgenic caffeine study

Evaluation Criterion	Discriminant Analysis		Decision Tree	Neural Network	Support Vector Machine
	Complete	Step-Fwd			
Accuracy	63%	79%	36%	58%	52%
Sensitivity	75%	75%	13%	63%	63%
Specificity	55%	82%	55%	55%	46%

Although the stepwise-forward variant of discriminant analysis was able to distinguish between treated and untreated animals using a highly-selective subset of the behavioral measures, the consensus among the classifiers examined is that the two groups are indistinguishable. Strong agreement among the data mining-based classifiers underscores the lack of discriminability between the two groups of mice. Indeed, the combination of measures selected by the stepwise-forward discriminant analysis to achieve discriminability does not exhibit face validity. Upon careful inspection, the list of predictor variables appears arbitrary and lacks coherence (i.e., it is not a "meaningful" sampling of variables, based on actual experience with behavioral testing). Furthermore, despite the capacity of neural networks to learn arbitrary mappings between features (e.g., behavioral measures) and corresponding instances (e.g., treatment groups) (Cybenko, 1989), the inability of neural network-based classifiers to successfully distinguish between groups of animals on the basis of behavioral metrics further implies similitude. These results mirror the standard analyses of vari-

ance comparing corresponding performance measures between caffeine-treated and untreated nontransgenic mice in the Y-maze, Morris water maze, Circular Platform, Platform Recognition, and Radial Arm water maze tasks (Arendash et al., 2009); measure for measure, these ANOVAs indicate no significant differences between the two groups of mice.

Collectively, the lack of discriminability between caffeine-treated and untreated animals suggests that chronic caffeine administration in nontransgenic mice, beginning in "early adulthood" and continuing until "late adulthood," does not provide significant benefits with respect to cognitive performance, relative to untreated animals. One implication for humans may be that long-term, daily oral intake of caffeine in moderate doses is unlikely to provide cognitive benefits in normal, healthy individuals.

CHAPTER 5

CAFFEINE ADMINISTRATION IN ALZHEIMER'S TRANSGENIC MICE

5.1 Introduction

Does long-term consumption of caffeine provide cognitive therapeutic benefits in individuals exhibiting Alzheimer's disease or Alzheimer-like neuropathology?

The neuroprotective potential of caffeine in humans has been studied through community-wide retrospective and prospective studies among coffee consumers (e.g., Johnson-Kozlow et al., 2002; Ritchie et al., 2007; van Gelder et al., 2007), wherein increased coffee consumption is associated with both reduced risk for Alzheimer's disease and smaller relative age-associated declines in cognitive abilities. Indeed, a 21 year follow-up study (Eckelinen et al., 2008) suggests that midlife consumption of between three and five cups of coffee substantially reduces the risk of subsequent AD. Another study suggesting an association between long-term caffeine consumption and AD (Maia and de Mendonca, 2002) found that AD patients consumed less caffeine during the 20 years prior to diagnosis, relative to age-matched individuals without AD. However, investigating the potential therapeutic benefits of caffeine within vulnerable populations (e.g., individuals predisposed to Alzheimer's disease) is a more difficult research problem. Indeed, widespread methodological issues (e.g., lack of standard measures for caffeine consumption, AD diagnostic criteria) further complicate efforts to evaluate possible therapeutic benefits of caffeine in AD patients (Rosso et al., 2008).

Studies involving long-term caffeine consumption in adult (four month-old) Alzheimer's transgenic mice (Arendash et al., 2006) show cognitive protection across multiple domains (spatial learning, reference memory, working memory, and object recognition) following five months of oral treatment (1.5 mg/day), relative to age-matched untreated animals. In addition, long-term caffeine administration is associated with decreased hippocampal beta-amyloid deposition and reduced expression of both presenilin (PS1) and beta-secretase (BACE) (Arendash et al., 2006), suggesting reduced production of beta-amyloid in caffeine-treated animals as the mechanism of cognitive protection. In another study (Arendash et al., 2009), the effects of long-term caffeine administration were examined in aged (18-19 month-old) Alzheimer's transgenic (APPsw) mice. Following pre-testing in the RAWM to confirm cognitive (working memory) impairment, aged transgenic animals consumed either caffeine-treated (0.3 mg/mL) drinking water or normal, untreated drinking water for four to five weeks. Age-matched nontransgenic animals, included in the study for subsequent groupwise performance comparisons, consumed untreated drinking water throughout the study. All animals were subsequently evaluated using a multimetric behavioral battery, and the results analyzed to determine whether chronic caffeine dosing influences cognitive performance against a background of Alzheimer-like pathology and cognitive impairment.

Figure 5.1 illustrates caffeine treatment-associated reversal of memory impairment in aged Tg mice (Arendash et al., 2009). Aged Tg mice display significant ($^{*}p < .000005$) working memory impairment in the RAWM task, relative to age-matched NT control animals, prior to caffeine treatment (Figure 5.1-A). Aged Tg mice receiving caffeine (Tg+CAFF) performed comparably to NT mice, with respect to RAWM working memory final-block T4 and T5 trials, after 4-5 weeks of caffeine treatment (Figure 5.1-B), while Tg animals differ significantly ($^{**}p < .025$) from

the other two groups. Untreated transgenic mice (Tg) did not exhibit significant improvement between pre-treatment and during-treatment evaluation in RAWM T5 performance, although both NT animals ($^{*}p < .05$; paired t-test) and caffeine-treated animals (Tg+CAFF) ($^{**}p < .005$; paired t-test) improved markedly (Figure 5.1-C). Comparable performance during the final two days of platform recognition testing (strategy switching) was observed in NT and Tg+CAFF mice, while Tg control animals showed impairment ($^{*}p < .02$; Tg vs. NT).

The purpose of this study was to re-evaluate the results of the aged Alzheimer's transgenic mice experiment involving long-term caffeine administration (Arendash et al., 2009), to compare standard statistical analyses (ANOVA) with advanced statistical and data mining-based methodologies for identifying transgenic- and therapeutic treatment-effects.

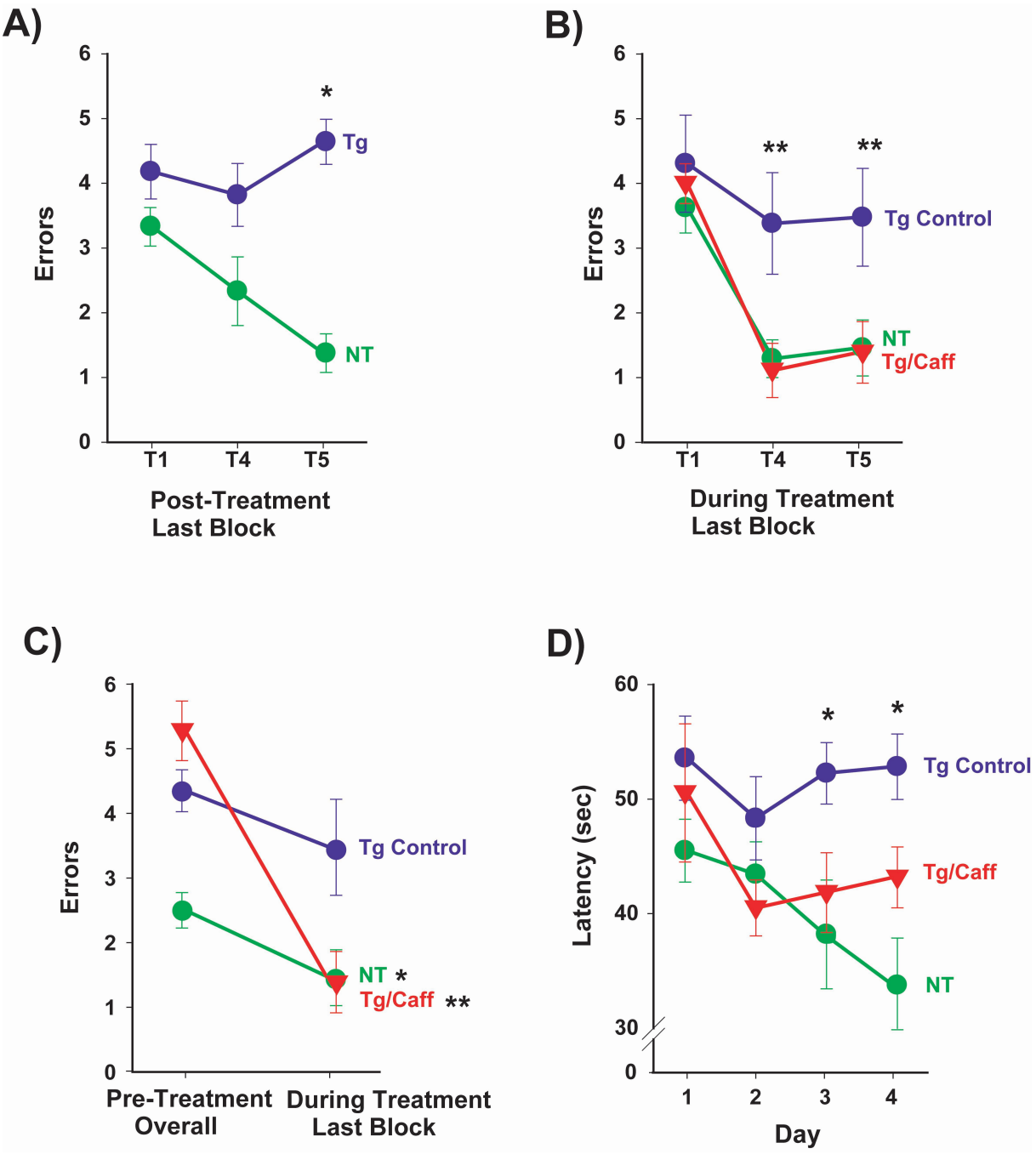

Figure 5.1. Behavioral assessment in NT and Tg mice receiving caffeine

5.2 Materials and Methods

A total of twenty subjects (twelve APPsw Alzheimer's transgenic mice and eight age-matched nontransgenic mice) were included in the study. At 18 to 19 months of age, all animals were pre-treatment tested in the RAWM task for six days to confirm working memory impairment in transgenic mice. After pre-treatment RAWM evaluation, the transgenic animals were randomly assigned to either of two treatment groups (Tg and Tg+CAFF), balanced by RAWM performance. Five of the transgenic animals received caffeine (Tg+CAFF) through their drinking water (0.3 mg/mL), while the other seven (Tg) continued to receive normal drinking water, as did the nontransgenic mice, for a four- to five-week treatment period. Following the treatment period, all animals completed a three-task behavioral assessment battery, from which eight measures were obtained (refer to Table 2.1 for coding): Radial Arm water maze (RAWM; RML-T4, RML-FT4, RML-T5, RML-FT5), Platform Recognition (PR-Avg, PR-Fin), and Y-maze (YM-AE, YM-PA). Detailed descriptions of the computing resources (hardware, software), as well as parameter settings for analytic programs, are provided in the General Analytic Protocol in Section 4.2.

Animal care and use was in accordance with the Guide and Use of Laboratory Animals, National Research Council, 1996, in a program and facilities fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, International, under a protocol approved by the University of South Florida Institutional Animal Care and Use Committee (No. 2729, Gary Arendash, Ph.D., Principal Investigator).

5.3 Results of Statistical Analyses

5.3.1 Correlation Analysis

Table 5.1. Correlations between behavioral measures in the Alzheimer's transgenic caffeine study

YM-PA							
PR-Avg							
PR-Fin			.90 .000				
RML-T4			.65 .002	.52 .020			
RML-FT4			.80 .000	.68 .001	.75 .000		
RML-T5			.69 .001	.71 .000	.74 .000	.61 .004	
RML-FT5			.63 .003	.58 .007	.58 .008	.69 .001	.79 .000
	YM-AE	YM-PA	PR-Avg	PR-Fin	RML-T4	RML-FT4	RML-T5

Significant correlations ($p < .05$) observed in pairwise comparisons of all measures for all three groups (NT, Tg, Tg+CAFF) collectively are shown in Table 5.1, wherein marked cells include both the correlation coefficient (r-value, top) and significance (p-value, bottom). Significant intra-task correlations exist within the Platform Recognition and Radial Arm water maze tasks, suggesting extensive association and/or coordination among cognitive processes within these tasks. Effective overall cognitive performance in RAWM, for example, is dependent upon successful functional integration between spatial learning and both working and reference memory systems.

Widespread significant inter-task correlations were identified between the Platform Recognition and Radial Arm water maze tasks, as well, suggesting the involve-

ment of similar underlying cognitive domains or neurobehavioral substrates across tasks. Observed inter-task correlations, therefore, suggest that the Platform Recognition and RAWM tasks share multiple learning- and memory-related features in common. The correlations between component measures across tasks in mice are analogous to the observed correlations between psychometric testing instruments used for human AD evaluation, resulting from parallel assessment of similar cognitive abilities (e.g., working memory, spatial recognition).

Neither measure obtained in the Y-maze task was significantly correlated with any other measures. This is not surprising for the activity-based Y-maze arm entries (YM-AE) measure. The relative independence of Y-maze alternations (YM-PA) from the other cognitive measures may reflect its general mnemonic foundation as separate from the cognitive domains for RAWM and Platform Recognition performance and/or context-sensitivity of learning and performance, i.e., learning within a land-based task (Y-maze) versus water-based tasks (Platform Recognition, RAWM).

5.3.2 Factor Analysis

Table 5.2. Unrotated factor component loadings in the Alzheimer's transgenic caffeine study

	Factor	
Measure	I	II
PR-Avg	0.902	
RML-T5	0.883	
PR-Fin	0.854	
RML-FT4	0.849	
RML-FT5	0.836	
RML-T4	0.809	
YM-AE		-0.868
YM-PA		0.712
Variance	57.11%	17.96%

An unrotated principal component analysis of the eight measures is shown in Table 5.2, with significant factor loadings (absolute value greater than 0.700) indicated. Calculated eigenvalue results (i.e., >1 criterion) were in agreement with observed scree plot (Cattell, 1966) identification of two significant factors. The primary factor (57% of overall variance) was comprised of all behavioral measures from the Platform Recognition and RAWM tasks, and represents an overall cognitive function component of behavior. Both Y-maze measures together comprised the second factor (accounting for approximately 18% of variance), underscoring their independence from the other behavioral measures, as was evident from correlation analyses as well. The Y-maze task involves both sensorimotor and cognitive elements, and the segregation of this land-based task from the two water-based tasks suggests context-sensitivity in both learning and performance (land vs. water) and/or the general mnemonic nature of the Y-maze task, as mentioned previously.

5.3.3 Discriminant Analysis

The results of the discriminant function analyses, both direct-entry (complete) and stepwise-forward approaches, are presented in Table 5.3 for reference purposes. All discriminant analysis comparisons involve the eight behavioral measures from "during-treatment" testing.

Nontransgenic (NT) vs. Transgenic-Control (Tg) Groups

A direct-entry (complete) discriminant analysis did not demonstrate significant discriminability between the two groups. However, a stepwise-forward analysis returned significant discriminability (Wilks's lambda = 0.456, p = .0017) with 87% accuracy (Jackknifing method). The sensitivity was 86% and the specificity was 88%, with respect to the transgenic-control group. Only the RML-T5 measure was

Table 5.3. Classifier performance comparison in the Alzheimer's transgenic caffeine study

Groups	Evaluation Criterion	Discriminant Analysis		Decision Tree	Neural Network	Support Vector Machine
		Complete	Step-Fwd			
Tg	Accuracy	NS	87%	80%	73%	67%
vs	Sensitivity	NS	86%	100%	71%	43%
NT	Specificity	NS	88%	63%	75%	88%
Tg	Accuracy	NS	75%	NS	67%	75%
vs	Sensitivity	NS	80%	NS	80%	60%
Tg+CAFF	Specificity	NS	71%	NS	57%	86%
NT	Accuracy	NS	100%	NS	77%	85%
vs	Sensitivity	NS	100%	NS	80%	80%
Tg+CAFF	Specificity	NS	100%	NS	75%	88%
All three	Accuracy	60%	80%	45%	70%	65%

retained by the stepwise-forward model, underscoring the sensitivity of the reference memory measure to distinguish transgenicity.

Transgenic-Control (Tg) vs. Transgenic+Treatment (Tg+CAFF) Groups

Nonsignificant discriminability was obtained using direct-entry discriminant analysis. A stepwise-forward analysis, however, returned significant discriminability (Wilks's lambda = 0.233, p = .0065) with 75% accuracy (Jackknifing method). The observed sensitivity was 80% and the specificity was 71%, with respect to the transgenic+treatment group. Three variables were retained by the stepwise-forward analysis: YM-AE, YM-PA, and RML-FT5. This subset of measures includes both sensorimotor (through the Y-maze measures) and cognitive elements, suggesting a more complex behavioral profile may distinguish between treated- and untreated-transgenic animals, reflecting both sensorimotor and cognitive components.

Nontransgenic (NT) vs. Transgenic+Treatment (Tg+CAFF) Groups

Although direct-entry discriminant analysis did not result in significant discriminability between the two groups, stepwise-forward analysis found significant dis-

criminability (Wilks's lambda = 0.164, p = .0007) with 100% accuracy (Jackknifing method). Both the sensitivity and specificity indices were 100%, with respect to the transgenic+treatment group. The stepwise-forward analysis returned a model having three predictor variables: YM-PA, RML-FT5, and RML-T5. This cognitive-biased subset of measures reflects systematic, coordinated search capability (Y-maze percent alternations), as well as spatial reference memory function.

All Three Groups

In contrast to the pairwise group analyses, a direct-entry discriminant analysis of all three groups showed significant discriminability (Wilks's lambda = 0.067, p = .0042), but with only 60% classificatory accuracy (Jackknifing method). Furthermore, a stepwise-forward analysis also returned significant discriminability (Wilks lambda = 0.158, p = .0001) with 80% accuracy (Jackknifing method). The three predictor variables retained by the classifier model were: YM-PA, RML-FT5, and RML-T5. This is the same subset of measures retained by the "Nontransgenic vs. Transgenic+Treatment" classifier model, and suggests that a combination of sensorimotor and cognitive behavioral indices may be necessary for successful discrimination among nontransgenic and transgenic animals (either caffeine-treated or untreated). The canonical scores plot generated from stepwise-forward discriminant analysis is depicted in Figure 5.2. The axes represent the two functions used to distinguish among the three groups in the analysis.

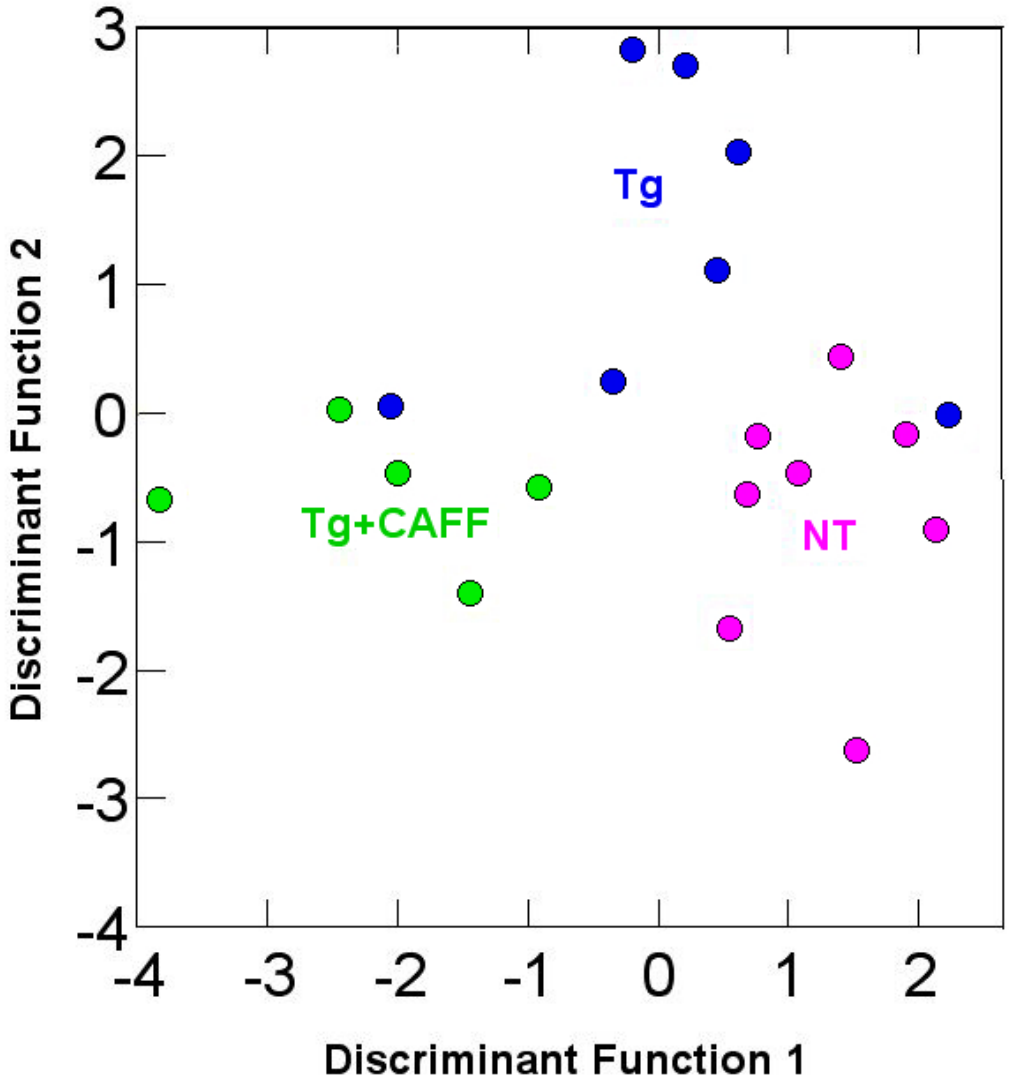

Figure 5.2. Canonical scores plot of discriminant analysis for all three groups

5.4 Results of Data Mining Analyses

5.4.1 Decision Tree Analysis

Nontransgenic (NT) vs. Transgenic-Control (Tg) Groups

A decision tree constructed using the J48 algorithm identified a single measure, PR-Fin, as the best discriminative attribute for splitting the dataset into distinct groups. Subsequent cross-validation (k-fold) of the optimal tree showed in 80% correct classification of cases (Kappa = 0.61), with 100% sensitivity and 63% specificity.

Transgenic-Control (Tg) vs. Transgenic+Treatment (Tg+CAFF) Groups

Although two attribute measures, PR-Avg and YM-PA, were selected by the decision tree-based classifier for providing optimal discriminability between the two groups, only 50% (Kappa = -0.03) of the individual animals were correctly classified. Hence, the performance of the classifier represents chance-level (nonsignificant) discriminability between untreated and caffeine-treated Alzheimer's transgenic mice.

Nontransgenic (NT) vs. Transgenic+Treatment (Tg+CAFF) Groups

Both Platform Recognition task measures (PR-Fin, PR-Avg) were selected by the decision tree classifier as having optimal capacity for splitting the dataset into two distinguishable groups. However, because only 31% of animals were correctly classified (Kappa = -0.52), the performance of the classifier did not exceed a level comparable to random assignment of individuals to groups. This finding suggests that, while the Platform Recognition measures exhibit modest bias between the two groups, the caffeine-treated transgenic mice do not significantly differ from the nontransgenic animals.

All Three Groups

The Y-maze percent alternations (YM-PA) and both Platform Recognition task measures (PR-Fin, PR-Avg) were identified as having sufficient information value to

split the database into three groups. Although only 45% of animals were correctly assigned to their respective groups (Kappa = 0.16), this modest level of classifier performance exceeds that of random-chance assignment (33%).

5.4.2 Neural Network Analysis

Nontransgenic (NT) vs. Transgenic-Control (Tg) Groups

A three-layered neural network, in which the number of hidden-layer units was varied between three and nine, demonstrated optimal performance when configured with eight input-layer elements (the behavioral measures), three hidden-layer elements, and two output-layer elements (the two category groups). Subsequent cross-validation (k-fold) of the optimal network showed 73% correct classification of cases (Kappa = 0.46), with 71% sensitivity and 75% specificity.

Transgenic-Control (Tg) vs. Transgenic+Treatment (Tg+CAFF) Groups

Four hidden-layer units were necessary to achieve optimal classification performance between treated and untreated transgenic mice. The resulting network performed modestly, correctly classifying 67% of cases (Kappa = 0.35), with 80% sensitivity and 57% specificity. Although only two-thirds of individuals were correctly identified by group, the classifier showed superior detection of caffeine-treated Alzheimer's transgenic mice (sensitivity), relative to untreated Tg controls (specificity).

Nontransgenic (NT) vs. Transgenic+Treatment (Tg+CAFF) Groups

The neural network-based classifier utilized four hidden-layer computing units to produce optimal discriminability between treated and untreated nontransgenic animals. Seventy-seven percent of individual animals were correctly assigned to their respective groups (Kappa = 0.53). The classifier displayed 80% sensitivity and 73% specificity. Hence, an identically-configured network was better able to accurately

and reliably distinguish (77% vs. 67%) a caffeine treatment effect in transgenic animals against a nontransgenic standard, relative to a transgenic standard.

All Three Groups

The optimal neural network-based classifier for distinguishing among the three groups contained four computing units in the hidden layer. This classifier performed remarkably well, demonstrating 70% accuracy (Kappa = 0.54), relative to the 33% which would be attributed to random assignment. The neural network's difficulty in distinguishing between untreated and caffeine-treated Alzheimer's transgenic animals (Tg vs. Tg+CAFF), mentioned previously, may have undermined the classifier's overall performance when all three groups were included.

5.4.3 Support Vector Machine Analysis

Nontransgenic (NT) vs. Transgenic-Control (Tg) Groups

A support vector machine-based classifier was trained using the behavioral measures to distinguish between nontransgenic (NT) and untreated transgenic (Tg) animals. Two-thirds of the animals were correctly classified (Kappa = 0.31), with 43% sensitivity and 88% specificity. Hence, although only modest discriminability was observed between the two groups, the classifier was particularly effective in correctly identifying nontransgenic animals (88% specificity), relative to the transgenic mice.

Transgenic-Control (Tg) vs. Transgenic+Treatment (Tg+CAFF) Groups

The classifier correctly distinguished between caffeine-treated and untreated transgenic mice in 75% of the cases (Kappa = 0.47). The sensitivity of the support vector machine was 60% and the specificity was 86%, indicating greater detection accuracy for untreated mice relative to treated animals.

Nontransgenic (NT) vs. Transgenic+Treatment (Tg+CAFF) Groups

The support vector machine-based classifier performed remarkably well, correctly assigning 85% of the mice comprising these two groups (Kappa = 0.68). The classifier displayed 80% sensitivity and 88% specificity. Hence, the classifier showed approximately equal facility in correctly identifying nontransgenic and caffeine-treated transgenic animals.

All Three Groups

When a support vector machine was trained using behavioral data from all animals of the three groups, only 65% of individuals were correctly assigned to their respective groups (Kappa = 0.45). Overall, this level of discriminability is similar to that between treated- and untreated-transgenic animals.

5.5 Discussion

Consistent with findings from earlier studies (e.g., Arendash and King, 2002; Leighty et al., 2004), significant intercorrelations were found both between and within behavioral tasks. The factor analysis was also consistent with prior findings (Arendash and King, 2002; Jensen et al., 2005; Arendash et al., 2006), wherein a cognitive-loaded behavioral component was returned as the primary factor. Additionally, in both the current study and in prior studies examining long-term caffeine administration in mice (Arendash et al., 2006), measures from the Platform Recognition and RAWM tasks loaded separately from the Y-maze task measures, underscoring the distinct sensorimotor/cognitive domain assessment of the Y-maze paradigm. The discriminant analyses are consistent with earlier studies (e.g., Arendash and King, 2002; Leighty et al., 2004), as well, where non-significant discriminability returned through direct-entry analysis is contrasted with significant discriminability achieved by utilizing the stepwise-forward approach.

Among the data mining classifiers, decision trees were particularly sensitive to transgenicity effects, while support vector machines were more sensitive to (caffeine) treatment effects in Alzheimer's transgenic mice. The decision tree method was successful only in distinguishing between nontransgenic and untreated (control) transgenic animals. The reliance of decision tree-based classifiers upon platform recognition-associated measures, however, contrasts with statistical classifiers, which favored Y-maze and RAWM measures. The data mining methods, while inferior to the stepwise-forward variant of discriminant analysis, were collectively superior to the direct-entry (complete) discriminant analysis classifier. These findings support the use of data mining techniques as supplements to – rather than substitutes for – standard statistical analytic methods. Table 5.3 summarizes the performance of all classifiers examined, with respect to sensitivity, specificity, and overall accuracy, in distinguishing between/among caffeine-treated and untreated transgenic, and nontransgenic mice.

Taken together, these results suggest differential sensitivity to treatment and/or transgenicity by predictor variables; some variables are better discriminators of transgenic- or treatment-effects than others. This explains the success of stepwise-forward discriminant analysis (and failure of direct-entry discriminant analysis) in distinguishing between groups, in terms of treatment as well as transgenicity. Furthermore, the stepwise-forward discriminant analyses suggest the three groups examined are pairwise-discriminable, albeit utilizing different predictor variables. A prior study utilizing the same dataset as in the present study (Arendash et al., 2009) employed standard ANOVA-based comparisons between/among nontransgenic, untreated- and caffeine-treated transgenic mice, and found that caffeine-treated transgenic animals differed significantly from untreated transgenics in the Platform Recognition (escape latency) and RAWM (final acquisition trial latency, T4; retention trial la-

tency, T5) tasks (Figure 5.1). In addition, caffeine-treated transgenic mice did not differ significantly from nontransgenics in these tasks (Arendash et al., 2009). Although these results seem to contradict the findings of the current study involving discriminant analysis and data mining-based methods, it is important to recognize that multiple predictor variables (behavioral measures) were utilized within each analysis of the current study, in contrast to the single measure-by-measure comparisons performed through ANOVA in Arendash et al. (2009). For example, measures from both the Y-maze and RAWM tasks – together – were necessary to distinguish between nontransgenic and caffeine-treated animals using discriminant analysis. Hence, significant discriminability, when observed, in the current study reflects distinguishable collective patterns of measures between groups, rather than differences between single measures. Indeed, the ability of both stepwise-forward discriminant analysis and data mining-based classifiers to distinguish between groups using multiple behavioral measures underscores the importance of these advanced techniques as complements to conventional behavioral analytic protocols, such as ANOVA.

Several implications for humans may be inferred from these findings. First, because transgenic mice (both caffeine-treated and untreated) were distinguishable from nontransgenics, it is likely that normal individuals and AD-individuals, regardless of therapeutic influence, will exhibit distinctive patterns of cognitive performance, however subtle, which are discernable to a well-conditioned classifier using a sufficiently rich dataset of measures. This supports the use of comprehensive behavioral and cognitive assessment for accurate, reliable diagnosis of probable Alzheimer's disease. Second, in light of findings by Arendash et al. (2009) that long-term oral caffeine consumption confers cognitive-protective benefits in aged Alzheimer's transgenic mice (based on specific task performance measures), it is reasonable to conclude that certain cognitive domains (e.g., working/reference memory) are highly respon-

sive to the therapeutic effects of caffeine. Finally, that certain tasks are more likely to detect these cognitive benefits, when present, than are other tasks. The Radial Arm water maze, for instance, has demonstrated superiority for identifying cognitive impairment, as well as for evaluating the efficacy of therapeutic interventions, in Alzheimer's transgenic mice (e.g., Arendash et al., 2004; Jensen et al., 2005; Ethell et al., 2006; Cracchiolo et al., 2007). Similarly, the MMSE (Folstein et al., 1975) psychometric inventory is one of the most effective instruments for both diagnostic screening and therapeutic evaluation in Alzheimer's patients. Assessment methodologies for Alzheimer's disease continue to improve, however, and recently-developed paradigms (e.g., semantic interference task; Loewenstein et al., 2004) emphasize specific cognitive domains (e.g., verbal memory) and operations (e.g., proactive and retroactive interference) which exhibit Alzheimer-specific impairment, as explored in the next chapter.

CHAPTER 6

INTERFERENCE TESTING IN HUMANS: A COMPARISON OF STATISTICAL AND DATA MINING METHODS

6.1 Introduction

Memory dysfunction – particularly, impaired delayed recall ability – is an early indicator of probable Alzheimer's disease (e.g., Welsh et al., 1991). Compromised storage and consolidation of new learning in AD patients may be related to increased susceptibility to stimulus intrusions occurring between the learning period and subsequent recall (i.e., interference) which, in turn, is associated with impaired hippocampal function (e.g., Hasselmo and Wyble, 1997). A novel cognitive assessment protocol, developed by Dr. David Loewenstein and colleagues at the University of Miami (Florida) School of Medicine, has demonstrated effectiveness as a psychometric screening instrument for distinguishing among clinically-diagnosed (DALCOG, MMSE) mild probable AD, mild cognitively impaired (MCI), and normal elderly individuals (Loewenstein et al., 2003; Loewenstein et al., 2004). The technique represents an extension of the Fuld object recognition task using two sets of semantically-related stimuli for evaluating both proactive and retroactive interference effects.

The interference testing protocol consists of four tasks. The first task, three-trial recall, is a modified version of the Fuld object memory examination (Fuld, 1981; Loewenstein et al., 2001), in which the subject is presented with ten familiar objects (Bag A) and asked to recall the objects following a brief distraction task, repeated three times. In the second task, proactive interference, the subject is presented

with ten novel objects (Bag B) and asked to recall them, to determine whether previous learning (Bag A objects) intrudes upon present learning (Bag B objects). The third task, short-delay recall, wherein the subject is asked to recall the original set of ten items (Bag A), provides a measure of retroactive interference (difficulty recalling previous learning due to intrusion by present learning). Finally, long-delay recall is evaluated by asking the subject to recall the original set of ten items (Bag A) after a 20-minute delay. The stimulus presentation and response forms were selected both to facilitate administration of the testing instrument as well as to closely mirror familiar, commonly executed cognitive tasks (e.g., object recognition by touch, naming) by human subjects. In addition, the objects comprising the two sets of stimuli were selected to be semantically related (Loewenstein et al., 2003), to increase the likelihood of intrusion (interference) errors. Both Alzheimer's disease patients and individuals with mild cognitive impairment were shown to be vulnerable to proactive interference, relative to normal elderly subjects (Loewenstein et al., 2004).

The purpose of this study was to explore advanced statistical and data mining methodologies for distinguishing among aged-normal, mild Alzheimer's, and mild cognitively impaired individuals using measures from the four tasks described above included in the same dataset as in the original Loewenstein et al. (2004) study. The results provided in the original report (Loewenstein et al., 2004), based on one-way analysis of variance (ANOVA) and logistic regression statistics, were compared with discriminant analyses and data mining-based analyses of the original dataset.

6.2 Materials and Methods

A total of 132 cases representing three classes of age-matched subjects were included in the dataset: Mild Alzheimer's disease (AD, N = 26), mild cognitive impairment (MCI, N = 53), and normal elderly (CON, N = 53). Group membership was assigned by clinical examination using standardized criteria (NINCDS-ADRDA). In addition to behavioral testing, all individuals completed the MMSE examination. The MMSE score was included in the correlation analysis to illustrate convergent validity, but only behavioral measures were used for subsequent analyses. The behavioral evaluation yielded the following measures, in order: Three-trial recall (Modified OME, Fuld) score (Bag-A objects), proactive interference (Bag-B Immediate Recall) score, retroactive interference (Bag-A Short Delay) score, and delayed-recall (Bag-A 20-min Delay) score. A verbal fluency task was used as a distractor between successive trials of the three-trial recall task, as well as immediately preceding the proactive interference task. These data were provided courtesy of Dr. David Loewenstein; complete details of the testing procedure are provided in Loewenstein et al. (2004).

6.3 Results of Statistical Analyses

6.3.1 Standard Analysis

Figure 6.1 depicts significant groupwise contrasts in component behavioral measures of the semantic interference protocol, with respect to memory performance (i.e., average number of correct responses). Using analysis of variance (ANOVA, F-test), all three pairs of groups (i.e., AD vs. MCI, AD vs. CON, and MCI vs. CON) exhibited significant differences in cognitive performance. Both the three-trial recall and proactive interference measures (Loewenstein's FRET3 and SITB measures, re-

spectively) differed significantly between these three pairs of groups (all p = .000). In addition, both the retroactive interference and delayed-recall measures (Loewenstein's SITAB and SITDRE measures, respectively) differed significantly between the CON group and either the AD or MCI group (all p = .000); these two measures also differed between the AD and MCI groups (both p = .020).

Furthermore, significant groupwise differences were observed in the average MMSE scores [Mean ± SEM: AD (23.12 ± 0.53), MCI (27.04 ± 0.26), and CON (28.56 ± 0.19)], wherein all pairwise differences between groups were statistically significant (all p = .000). Hence, for each of the cognitive performance measures, all pairs of groups differ significantly ($p < .05$) and, collectively, portray a continuum of cognitive impairment from relatively-unimpaired performance (CON) to substantial impairment (AD). Normal elderly individuals (CON), for instance, demonstrated superior performance, relative to the other two groups. By contrast, mildly cognitively-impaired (MCI) individuals exhibited intermediate performance between the other two groups.

6.3.2 Correlation Analysis

Table 6.1. Correlations among MMSE scores and behavioral measures in the human semantic interference protocol

Three-Trial Recall	.54 .000			
Proactive Interference	.46 .000	.75 .000		
Retroactive Interference	.41 .000	.68 .000	.51 .000	
Delayed Recall	.41 .000	.78 .000	.60 .000	.75 .000
	MMSE	Three-Trial Recall	Proactive Interference	Retroactive Interference

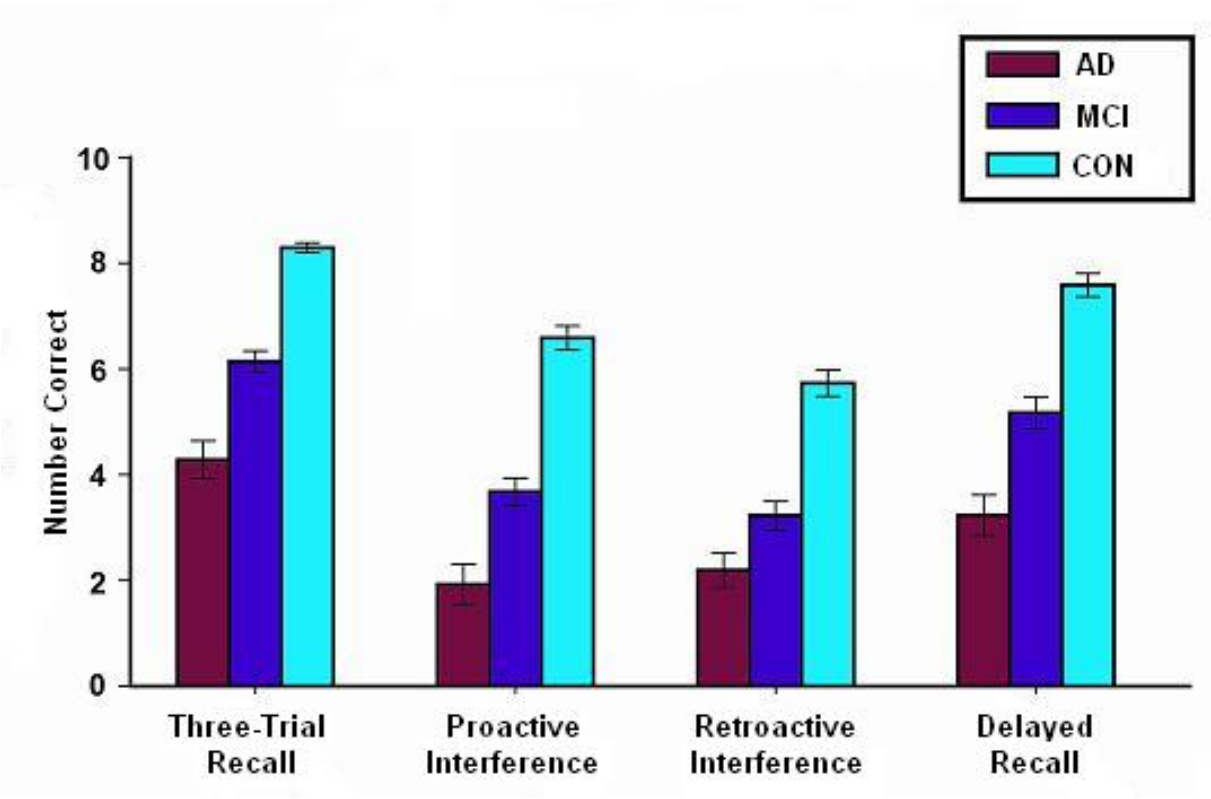

Figure 6.1. Groupwise contrasts for all human interference protocol measures

Significant ($p < .05$) pairwise correlations among all clinical measures (both task-based and MMSE) for all groups of subjects are shown in Table 6.1. Both the Pearson product-moment correlation coefficient (r-value) and its corresponding significance (p-value) are indicated at the top and bottom, respectively. Positive correlation between cognitive performance in the three-trial recall task and both the retroactive interference and delayed-recall measures suggests performance consistency between the original learning component (i.e., the first set of verbal stimuli) and subsequent retroactive interference. In addition, positive correlation between the proactive interference measure and the other three verbal tasks (three-trial recall, retroactive interference, and delayed-recall) suggests comparable learning performance across problem conditions (i.e., different sets of verbal stimuli). Finally, significant positive correlation between the MMSE and all task-based behavioral measures underscores the diagnostic utility of the MMSE psychometric instrument for verbal learning and memory evaluation.

6.3.3 Discriminant Analysis

Table 6.2 compares the performance of the classifiers examined in the present study, including the direct-entry and stepwise-forward discriminant analyses.

Alzheimer's (AD) vs. Mild Cognitive Impairment (MCI) Groups

Direct-entry (complete) discriminant analysis demonstrated significant discriminability between individuals with probable Alzheimer's disease and those with mild cognitive impairment (Wilks's lambda = 0.730, p = .0001), with 70% of subjects correctly assigned to their respective groups, as shown in Table 6.2. The sensitivity of the classifier was 65% and the specificity was 72%, with respect to Alzheimer's identification. Stepwise-forward analysis also returned significant discriminability between the two groups (Wilks's lambda = 0.754, p = .0000) with 76% accuracy,

Table 6.2. Classifier performance comparison in the human semantic interference protocol

Groups	Evaluation Criterion	Discriminant Analysis		Decision Tree	Neural Network	Support Vector Machine
		Complete	Step-Fwd			
AD	Accuracy	70%	76%	73%	68%	73%
vs	Sensitivity	65%	69%	NS	NS	NS
MCI	Specificity	72%	79%	93%	77%	91%
AD	Accuracy	94%	94%	91%	90%	95%
vs	Sensitivity	85%	85%	85%	85%	89%
CON	Specificity	98%	98%	94%	93%	98%
MCI	Accuracy	90%	90%	83%	88%	91%
vs	Sensitivity	87%	87%	83%	87%	93%
CON	Specificity	92%	92%	83%	89%	89%
All three	Accuracy	71%	81%	73%	73%	76%

69% sensitivity and 79% specificity. The stepwise-forward model retained only the three-trial recall score (i.e., Fuld task score) measure. Hence, the three-trial recall task component provides sufficient predictive information to distinguish between mild cognitive impairment and probable Alzheimer's disease. Indeed, the additional behavioral task measures may, in fact, undermine discriminability between these two groups.

Alzheimer's (AD) vs. Normal-Aged (CON) Groups

Successful, and identical, discriminability between normal aged and probable Alzheimer's individuals was achieved using direct-entry and stepwise-forward discriminant analyses (Wilks's lambda = 0.245, p = .0000). Both classifier variants exhibited 94% accuracy, with 85% sensitivity and 98% specificity, with respect to Alzheimer's identification, as reported in Table 6.2. The three-trial recall, proactive interference, and retroactive interference task scores were retained by the stepwise-forward model. Equivalent performance between the direct-entry and stepwise-

forward analyses suggests the delayed-recall measure does not contribute to overall discriminability between normal-aged and probable Alzheimer's individuals.

Mild Cognitive Impairment (MCI) vs. Normal-Aged (CON) Groups

Direct-entry and stepwise-forward discriminant analyses performed identically, distinguishing between normal-aged and mildly cognitively impaired individuals (Wilks's lambda = 0.438, p = .0000) with 90% accuracy, as indicated in Table 6.2. The observed sensitivity was 87% and the specificity was 92%, with respect to the MCI group of subjects. The stepwise-forward model retained the three-trial recall, proactive interference, and retroactive interference task scores. These results once again suggest the delayed-recall measure does not contribute to discriminability between normal-aged and MCI individuals.

All Three Groups

Individuals of all three groups (normal-aged, probable Alzheimer's disease, and mildly cognitively impaired) were successfully distinguished using direct-entry discriminant analysis (Wilks's lambda = 0.337, p = .0000). Seventy-one percent of the subjects were correctly identified by group. The normal-aged individuals were the most reliably distinguished by the classifier. A stepwise-forward analysis also returned significant discriminability among the three groups (Wilks's lambda = 0.220, p = .0000) with 81% accuracy. All four behavioral measures were retained by the stepwise-forward analysis as predictor variables for group membership. These results suggest that all of the cognitive measures may contribute significant, nonredundant discriminative information for determining group membership.

6.4 Results of Data Mining Analyses

6.4.1 Decision Tree Analysis

Alzheimer's (AD) vs. Mild Cognitive Impairment (MCI) Groups

Two behavioral measures (proactive interference and delayed-recall) were selected by the decision tree for splitting the dataset (Kappa = 0.3102). The resulting classifier correctly identified 73% of individuals, with nonsignificant (35%) sensitivity and 93% specificity, as shown in Table 6.2. Most of the subjects were classified as MCI, regardless of actual group.

Alzheimer's (AD) vs. Normal-Aged (CON) Groups

The dataset attributes identified by the decision tree were the three-trial recall and delayed-recall measures. The classifier accurately identified 91% of individuals (Kappa = 0.7974). The sensitivity was 85% and specificity was 94%, with respect to probable Alzheimer's disease. Hence, the normal-aged individuals were somewhat more likely to be correctly identified, relative to Alzheimer's subjects.

Mild Cognitive Impairment (MCI) vs. Normal-Aged (CON) Groups

A decision tree attempted to split the dataset using the three-trial recall and proactive interference measures. The resulting classifier correctly identified 83% (Kappa = 0.6604) of individuals, with 83% sensitivity and 83% specificity, as indicated in Table 6.2.

All Three Groups

The decision tree-based classifier selected three measures (three-trial recall, proactive interference, and delayed-recall) to distinguish among all three groups of individuals. The optimal tree accurately identified 73% of the individuals (Kappa = 0.5678). The retroactive interference measure was not selected by any of the decision tree-based classifiers.

6.4.2 Neural Network Analysis

Alzheimer's (AD) vs. Mild Cognitive Impairment (MCI) Groups

A neural network-based classifier was trained to distinguish between individuals from the probable Alzheimer's disease and mildly cognitively impaired groups using all four behavioral measures from the interference paradigm. Optimal performance was achieved with a three-layered network containing four computing elements within the input layer elements (the behavioral measures), three hidden-layer elements, and two output-layer elements (the two category groups). As indicated in Table 6.2, the classifier correctly identified 68% of cases (Kappa = 0.2763), with nonsignificant (50%) sensitivity and 77% specificity, with respect to Alzheimer's disease. True-AD individuals were assigned equally to either the AD or MCI group.

Alzheimer's (AD) vs. Normal-Aged (CON) Groups

As shown in Table 6.2, a neural network-based classifier successfully distinguished between normal aged individuals and those with probable Alzheimer's disease (Kappa = 0.7707), with 90% accuracy, 85% sensitivity and 93% specificity, with respect to Alzheimer's disease. Although this classifier exhibited remarkably accurate performance, 15% of probable Alzheimer's disease individuals were misclassified as being normal aged.

Mild Cognitive Impairment (MCI) vs. Normal-Aged (CON) Groups

Discriminability between normal aged and mildly cognitively impaired individuals was demonstrated by a neural network-based classifier (Kappa = 0.7547), which accurately assigned 88% of individuals to their respective groups. The classifier showed 87% sensitivity and 89% specificity, with respect to mild cognitive impairment. Thus, excellent discriminability was achieved between MCI and normal-aged individuals through neural network analysis.

All Three Groups

All three groups (normal aged, mildly cognitively impaired, probable Alzheimer's disease) of individuals were successfully distinguished (Kappa = 0.5657), as reported in Table 6.2. Seventy-three percent of individuals were correctly assigned to their respective groups using the classifier.

6.4.3 Support Vector Machine Analysis

Alzheimer's (AD) vs. Mild Cognitive Impairment (MCI) Groups

A support vector machine architecture attempted to distinguish between mildly cognitively impaired individuals and those with probable Alzheimer's disease using the behavioral measures from the interference paradigm. As shown in Table 6.2, the classifier correctly identified 73% of the individuals by group (Kappa = 0.3253), with nonsignificant (39%) sensitivity and 91% specificity. Hence, although almost all of the MCI individuals were correctly matched to their group, most of the AD individuals were incorrectly identified as belonging to the MCI group.

Alzheimer's (AD) vs. Normal-Aged (CON) Groups

Ninety-five percent of normal aged and probable Alzheimer's individuals were correctly assigned to their respective groups using a support vector machine-based classifier (Kappa = 0.8830). The observed sensitivity was 89% and the specificity was 98%, with respect to Alzheimer's disease. This classifier performed remarkably well overall, in that only 11% of probable Alzheimer's individuals were incorrectly identified as being normal aged.

Mild Cognitive Impairment (MCI) vs. Normal-Aged (CON) Groups

Normal aged and mildly cognitively impaired individuals were successfully distinguished (Kappa = 0.8113) using a support vector machine architecture. As reported

in Table 6.2, the classifier accurately identified 91% of the individuals from these two groups, with 93% sensitivity and 89% specificity, with respect to MCI.

All Three Groups

A respectable 76% of individuals from all three groups were correctly assigned to their respective group (Kappa = 0.6110) using classifiers based on support vector machine architectures.

6.5 Discussion

The original report (Loewenstein et al., 2004) presented ANOVA-based statistical results of comparisons between groups of all task measures in the interference protocol. The three groups were found to differ significantly (all $p < .001$), with respect to cognitive performance, in all four tasks (three-trial recall, proactive interference, retroactive interference, and delayed-recall). In addition, pairwise comparisons between groups were performed by logistic regression for all tasks except the three-trial recall component of the protocol. The mild cognitively impaired (MCI) and normal-aged (CON) groups showed significant differences (all $p < .0001$) in the proactive interference, retroactive interference, and delayed-recall measures (Loewenstein et al., 2004). Proactive interference was the best discriminator between these two groups (81.3% accuracy), while delayed-recall measure was the least-effective predictor variable (75.5%). Between the mild AD patients (AD) and normal-aged (CON) individuals, the proactive interference ($p < .0001$), retroactive interference ($p < .001$), and delayed-recall ($p < .0001$) measures differed significantly between the two groups (Loewenstein et al., 2004). However, between the AD and CON groups, delayed-recall provided the best discriminability (89.9%) and retroactive interference showed the poorest discriminability (83.3%), of the four tasks. Additionally, higher accuracy, sensitivity, and specificity were observed when multiple measures

were combined as a single index of cognitive performance (e.g., Total Recognition Memory Score) (Loewenstein et al., 2004).

In this study, the correlation analysis showed significant (all p = .000) positive pairwise correlation among all four behavioral measures, as well as between the standard psychometric screening instrument (MMSE) and each behavioral measure. Because only four behavioral measures were examined, and these exhibited extensive intercorrelation, no factor analyses were performed. The discriminant analyses were consistent with earlier studies (e.g., Arendash and King, 2002; Leighty et al., 2004), in which the stepwise-forward variant exhibits comparable (or, sometimes, enhanced) accuracy, relative to the standard direct-entry (complete) analysis. Indeed, although all discriminant analysis-based classifiers successfully distinguished between pairs of groups, by cognitive impairment, the stepwise-forward approach was superior for distinguishing between AD and MCI groups, as well as among all three groups, as shown in Table 6.2. Overall, all three decision tree-based classifiers successfully distinguished between/among the groups, but the support vector machine's performance was the most similar to the stepwise-forward discriminant analysis result.

Support vector machine-based classifiers demonstrated superior discriminability between the mild cognitively impaired (MCI) and normal-aged (CON) individuals, as depicted in Table 6.2. The stepwise-forward discriminant analysis retained the three-trial recall, proactive interference, and retroactive interference measures for distinguishing between the groups. This finding is consistent with that of Loewenstein et al. (2004), wherein these same three measures provided superior discriminability, relative to delayed-recall, between these two groups. The decision tree, as well, retained the three-trial recall and proactive interference measures, underscoring the relative importance of these cognitive metrics for detecting subtle features of

MCI. Indeed, Loewenstein et al. (2004) identified proactive interference as providing optimal discriminability (81.3%) between the two groups, by logistic regression.

The support vector machine-based classifier also displayed the best discriminability between probable AD and normal-aged individuals, as shown in Table 6.2. As noted earlier, equivalent classifier performance despite the absence of the delayed-recall measure from the stepwise-forward analysis suggests this measure does not contribute significantly to groupwise discriminability. This contrasts with the logistic regression-based results (Loewenstein et al., 2004) which emphasize the delayed-recall measure for distinguishing between these two groups. Indeed, the overall accuracy achieved by discriminant analysis, as shown in Table 6.2, exceeds that of logistic regression between the AD and CON groups. The decision tree-based classifier determined that the three-trial recall and delayed-recall measures together provide the most information bias for distinguishing between these two groups. The observed accuracy of this classifier (91%, from Table 6.2) approximates the logistic regression-based finding (89.9%) for delayed-recall alone, as reported in Loewenstein et al. (2004). The neural network-based classifier performed comparably to the logistic regression method, with respect to discriminability between probable AD and normal-aged (CON) individuals.

As indicated in Table 6.2, the stepwise-forward discriminant analysis displayed the best discriminability between probable AD and mild cognitive impairment, utilizing only the three-trial recall measure. Although Loewenstein et al. (2004) did not report relative discriminability using the Fuld three-trial recall measure, they emphasize the clinical significance of this task for the diagnosis of dementia. Interestingly, the decision tree retained both the proactive interference and delayed-recall measures, and classified almost all individuals as MCI, regardless of true group membership. Indeed, the difficulty of all three data mining techniques to accurately

distinguish between these groups (all data mining-based classifiers displayed non-significant sensitivity, effectively assigning true-AD individuals randomly to either group) underscores the challenge of reliable behavior-based differential diagnosis of AD. The adverse effects of disproportionate sample sizes on training the classifiers cannot be discounted, however. Although the MCI and normal-aged groups had equal sizes (N=53), there were only one-half as many AD individuals (N=26) included in the study. Either doubling the number of AD subjects or halving the numbers of MCI and normal-aged subjects (by Monte Carlo selection, for example) might correct the sample size-bias problem.

Discriminant analyses retained all behavioral measures for distinguishing among the three groups, and displayed 81% overall accuracy, as reported in Table 6.2. As noted earlier, the retention of all four behavioral measures by the stepwise-forward analysis suggests that each measure may contribute unique discriminative information. Indeed, the Loewenstein et al. (2004) findings indicate that multiple measures of cognitive ability provide superior discriminability, relative to individual measures. The decision tree-based classifier retained all measures except retroactive interference, but only identified 73% of the individuals by group, as shown in Table 6.2.

Taken together, these results suggest the interference-based cognitive assessment protocol introduced in Loewenstein et al. (2004) is highly effective in distinguishing between and among mild-Alzheimer's, mild cognitively impaired, and aged-normal individuals. In addition, significant positive correlation between the behavioral measures and the MMSE (standard clinical assessment instrument) supports convergent validity for diagnosis. As portrayed in Table 6.2, there is strong consensus among the classifiers with respect to discriminability between groups, e.g., AD individuals are more easily distinguished from aged-normal than are MCI individuals. The comparisons and contrasts between conventional statistical approaches (ANOVA, logistic

regression), as reported in Loewenstein et al. (2004), and the advanced statistical (discriminant analyses) and data mining-based approaches utilized here further support the use of the latter techniques in conjunction with the former for comprehensive neurobehavioral research. Future studies should address refinements and extensions of both statistical and data mining methods to improve diagnostic utility.

CHAPTER 7

THE INTERFERENCE TASK: A NOVEL ASSESSMENT PARADIGM

7.1 Introduction

As discussed in the previous chapter, the semantic interference task (Loewenstein et al., 2004) is a recently-developed clinical diagnostic protocol for distinguishing among normal-aged, mild cognitively-impaired, and probable Alzheimer's individuals using a verbal-report memory test having proactive and retroactive interference components, as well as a measure of delayed-recall performance. The demonstrated effectiveness of this protocol for detecting and, indeed, discriminating mild forms of cognitive impairment in humans suggested that a similar ensemble of rodent-based behavioral tasks and measures might exhibit comparable utility. The notion of adapting human-based tasks for cognitive assessment in rodents is not without precedent. Indeed, the most common single evaluation instrument is the maze (e.g., Hebb and Williams, 1946), a scaled model of the classic path-search puzzle. Both water- and land-based mazes, featuring two- and three-dimensional configurations, have been developed for studying learning and memory processes (e.g., Spear et al., 1990) in varied contexts. Similarly, the serial reaction time test for procedural (implicit) memory functional evaluation (Nissen and Bullemer, 1987) has also been adapted for mice (e.g., Christie and Hersch, 2004; Cho et al., 2007). Finally, episodic-like memory function has been examined in mice using exploration and object recognition tasks (Dere et al., 2005) to simulate the "where, what, and when" paradigms

commonly used in humans. Mirroring the episodic memory impairment displayed by Alzheimer's patients, Alzheimer's transgenic mice perform poorly in the simulated episodic-like memory task (Savonenko et al., 2005).

The purpose of this study was to implement and evaluate a novel behavioral testing paradigm for mice, adapted from the human-based instrument presented in Loewenstein et al. (2004), in nontransgenic and Alzheimer's transgenic animals having either the wildtype or GRK5-knockout genotype. The mouse-analogue of the instrument, described in Materials and Methods below, substitutes spatial memory-dependent elements for the verbal (semantic) memory-dependent components of the original protocol, while preserving the flexibility of the latter for evaluating proactive and retroactive interference effects on learning. Additionally, in contrast to the Loewenstein et al. (2004) which utilized only error-scores, the mouse-based adaptation provides both error-score and response-latency measures from each task. As shown in prior studies with both nontransgenic and Alzheimer's transgenic animals (e.g., Arendash et al., 2001; Leighty et al., 2004; Arendash et al, 2006), the Radial Arm water maze (RAWM) provides an excellent means for evaluating spatial short-term (working) memory function in mice. Consequently, the RAWM represents a reasonable starting design from which to develop the mouse-based interference testing paradigm. The two alternative configurations of the RAWM apparatus (Pool A and Pool B), featuring unique target placement and peripheral visual cues, for example, are intended to parallel the original protocol's use of semantically-related objects (Bag A and Bag B). Similarly, the intertrial Y-maze exposure period is intended as a brief distractor task, analogous to the intervening verbal fluency (naming) task used by Loewenstein et al. (2004).

G-protein coupled receptor kinase-5 (GRK5) is responsible for selective desensitization of G-protein coupled receptors of the muscarinic acetylcholinergic system in

mice (reviewed in: Suo et al., 2007). Studies *in vitro* show sub-threshold concentrations of soluble beta amyloid reduce functional GRK5 levels and, moreover, that these reductions precede the onset of cognitive impairment in AD transgenic mice (Suo et al., 2004). Aged (19 month-old) GRK5-knockout mice exhibit widespread hippocampal pathology (swollen axonal clusters), as well as muscarinic cholinoceptive system dysfunction (reduced mAChR M1, M2, and M4) (Suo et al., 2007); these animals also display selective impairment in working memory, but not spatial learning or reference memory. The GRK5-knockout mouse has been proposed as a model organism for human AD, as a complement to existing APP-transgenic mice, to evaluate both behavioral (cognitive) and pathologic (e.g., cholinergic) responses to treatment protocols.

In this study, the mouse-based interference paradigm was used to compare conventional statistical (ANOVA) methods, as reported in Suo et al. (2007), with advanced statistical (discriminant function) and data mining-based analytic techniques for evaluating cognitive impairment in mice (both nontransgenic and Alzheimer's transgenic) with/without GRK5 manipulation (i.e., wildtype or knockout). In addition, error-score and response-latency data were examined both separately and together to determine the most effective group of behavioral measures for distinguishing between/among treatment groups of mice. Hence, the interaction between two genetic manipulations (nontransgenic vs. Alzheimer's transgenic, GRK5-wildtype vs. knockout), as reflected in cognitive performance in the interference paradigm, was investigated.

7.2 Materials and Methods

The animals used for this analysis represented an extension of the initial behavioral study (reported in: Suo et al., 2007), involving one-year-old Alzheimer's transgenic (APPsw) Tg2576 mice and nontransgenic littermates, with an additional manipulation of the GRK5 gene (wildtype vs. knockout). Targeted deletion of exons 7 and 8 of the GRK5 gene was used to generate GRK5-KO ("knockout") mice. Hence, there were four groups of animals examined: Nontransgenic-control (NT, N=7), nontransgenic-knockout (NT-KO, N=8), transgenic-control (Tg, N=12), and transgenic-knockout (Tg-KO, N=10). All animals completed the comprehensive behavioral task battery (as used in prior studies, e.g., Jensen et al., 2005; Cracchiolo et al., 2007), prior to the Mouse Interference Paradigm.

All behavioral measures from the comprehensive task battery were analyzed using both statistical (ANOVA, correlation analysis, factor analysis, and discriminant analysis) and data mining-based methods (decision trees, neural networks, and support vector machines). Similarly, the three datasets of interference paradigm-derived behavioral measures obtained in this study were subsequently analyzed using the same analytic suite. A complete description of the computing resources (e.g., hardware platform, software packages) used in the analyses, including parameter settings for the programs, is provided in the General Analytic Protocol of Section 4.2.

Animal care and use was in accordance with the Guide and Use of Laboratory Animals, National Research Council, 1996, in a program and facilities fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, International, under a protocol approved by the University of South Florida Institutional Animal Care and Use Committee (No. 2951, Gary Arendash, Ph.D., Principal Investigator).

Mouse Interference Paradigm

Apparatus. The apparatus consists of two circular pools ("A" and "B"; 100 cm diameter x 25 cm deep), each containing a six-armed (30 cm long x 19 cm wide; radially distributed around a 40 cm diameter central arena) stainless steel insert, and a relocatable, transparent submerged platform (9 cm diameter; e.g., inverted clear glass jar). The pools are filled with warm water (maintained at 24-27 deg C) to a level 1.5 cm above the top of the submerged platform, positioned in the appropriate goal-arm. For each pool, the arms are labeled in a fixed, counter-clockwise order, beginning with the arm located directly in front of the experimenter (#1). A distinct visual cue (unique size and shape) is placed at the end of arms #2 through #6, along the exterior circumference of the pools. In addition, a Y-maze apparatus (three-armed; 21 cm long x 4 cm wide x 40 cm high walls) is located nearby.

Protocol. Behavioral testing typically consists of 4-6 consecutive daily sessions, with each test day's session comprised of: one "Platform A Orientation" trial, three "Platform A Recall" trials, one "Platform B Orientation" trial, one "Platform B Recall" trial, one "Platform A Short Delay" trial, and one "Platform A Long Delay" trial. In addition, each pool is assigned a specific start-arm and goal-arm to be used throughout the day's session. Each pool's daily start-arm and goal-arm assignments (for up to six testing sessions) are: A(2,6)B(3,6); A(5,3)B(1,3); A(4,5)B(2,1); A(4,1)B(3,4); A(6,4)B(2,6); and A(5,2)B(5,3). At the beginning of each session, a submerged platform is positioned near the end of each pool's goal-arm. For the single, unscored "Platform A Orientation" trial, the mouse is introduced at the center of the start-arm of Pool A, facing the central arena, and allowed 60 sec to explore the pool undisturbed. If the animal has not located the submerged platform, it is gently guided to – and allowed to remain atop – the platform for 30 sec. The mouse is then placed into one arm of the Y-maze and allowed to explore for 60 sec. Next,

on each of three recall trials: the mouse is introduced into the start-arm of Pool A, midway along the length and facing the central arena, and allowed up to 60 sec to locate and mount the submerged platform. If the mouse enters a non-goal arm, it is gently withdrawn to the start-arm and an error is recorded. If the mouse fails to locate the platform after 60 sec, it is gently guided there and allowed to remain for 30 sec. After each recall trial, the mouse is placed into the Y-maze and allowed to explore for 60 sec. In addition, for each recall trial, the number of errors (travelling more than 20cm into any non-goal arm) and the latency (time to reach the platform; scored as "60 sec," if unreached) are recorded as "Platform A Recall - Errors" and "Platform A Recall - Latency," respectively. The mouse is then introduced at the center of the start-arm of Pool B, facing the central arena, and allowed 60 sec to explore the pool undisturbed for the single, unscored "Platform B Orientation" trial. If the animal has not located the submerged platform, it is gently guided to – and allowed to remain atop – the platform for 30 sec. The mouse is then placed into one arm of the Y-maze and allowed to explore for 60 sec. Next, for a single trial, the mouse is introduced into the start-arm of Pool B, midway along the length and facing the central arena, and allowed up to 60 sec to locate and mount the submerged platform. Each time the mouse enters a non-goal arm, it is withdrawn to the start-arm. If the mouse fails to locate the platform after 60 sec, it is gently guided there and allowed to remain for 30 sec. The mouse is then placed into the Y-maze for 60 sec. The number of errors and latency are recorded as "Platform B Recall - Errors" and "Platform B Recall - Latency," respectively. The animal is then placed into Pool A for a single trial, as described above, and the "Platform A Short Delay" number of errors and latency are recorded. The mouse is transferred to its home cage for 20 min, after which it is placed into Pool A for a single trial, as described above, and the "Platform A Long Delay" number of errors and latency are recorded.

Dataset. Corresponding behavioral measures obtained from the final two daily testing sessions are averaged as a "block" for subsequent analysis. Hence, eight behavioral measures are determined: Three-trial recall (mean error-score and response-latency, from Platform A Recall); proactive interference (mean error-score and response-latency, from Platform B Recall); retroactive interference (mean error-score and response-latency, from Platform A Short Delay); and, delayed-recall (mean error-score and response-latency, from Platform A Long Delay). Finally, these measures are grouped into three datasets: Error-scores only, response-latencies only, and both error-scores and reponse-latencies.

7.3 Results of Statistical Analyses: Comprehensive Task Battery

7.3.1 Standard Behavioral Analysis

Table 7.1 portrays the standard ANOVA analyses (F-test, $\alpha = .05$) of behavioral measures from the comprehensive task battery for all animals in the study. Significant differences between groups are printed in **bold** and denoted by superscripted characters: (1) The nontransgenic-control (NT) vs. Alzheimer's transgenic-control (Tg) group contrast is represented by an asterisk ($*$); the Alzheimer's transgenic-control vs. transgenic with GRK5-knockout genotype (Tg-KO) contrast is indicated with a dagger ($\dagger$); and, the contrast between nontransgenic-control and nontransgenic with GRK5-knockout genotype (NT-KO) is represented by a double-dagger ($\ddagger$). Significant differences were found between the NT and Tg groups in Y-maze percent alternations (YM-PA, but not YM-AE), elevated plus maze open-arm entries (EP-OE; but not EP-CE or EP-TO), and radial arm water maze retention memory (overall errors and latencies in both T4 and T5; errors in final block T5). The primary cognitive impairment in Tg mice (relative to NT animals), therefore, was

reflected in the RAWM error measures, although a Tg effect on Y-maze alternations (general mnemonic function) was also evident. The Tg mice differed significantly from the Tg-KO animals in both open field activity (OF) and elevated plus maze open-arm entries. Similarly, NT animals and NT-KO mice also differed significantly with respect to open field activity. Hence, no cognitive effects of the GRK5-knockout genotype were observed in either nontransgenic or Alzheimer's transgenic mice. No significant differences were found between the NT and Tg-KO groups for any of the behavioral measures.

Table 7.1. Groupwise contrasts for all behavioral task measures in the GRK5 study

Behavioral Measure	Group mean (± standard error)			
	NT	Tg	Tg-KO	NT-KO
Sensorimotor-based				
OF	136.1 (± 13.5)	156.3 (± 6.7)	**142.0**† (± 11.0)	**87.6**‡ (± 10.5)
BB	53.0 (± 7.0)	39.4 (± 7.2)	35.8 (± 8.6)	39.7 (± 9.0)
SA	4.3 (± 0.7)	4.6 (± 0.4)	3.5 (± 0.8)	3.9 (± 0.7)
EP-CE	12.4 (± 1.4)	15.2 (± 2.8)	11.7 (± 2.4)	11.3 (± 1.8)
EP-OE	0.7 (± 0.5)	**30.3*** (± 14.6)	**10.5**† (± 5.4)	1.4 (± 0.5)
YM-AE	33.3 (± 2.7)	36.9 (± 2.9)	36.7 (± 3.1)	25.1 (± 2.4)
Anxiety-based				
EP-TO	0.7 (± 0.5)	1.1 (± 0.4)	1.0 (± 0.5)	3.8 (± 1.9)
Cognitive-based				
YM-PA	67.0 (± 5.1)	**51.0*** (± 2.6)	50.3 (± 3.3)	62.6 (± 3.0)
WM-Fin	20.1 (± 6.7)	21.1 (± 3.3)	22.6 (± 4.4)	24.5 (± 4.7)
WM-Avg	28.0 (± 4.8)	28.6 (± 3.1)	31.2 (± 3.5)	33.7 (± 3.7)
CPE-Fin	12.2 (± 4.1)	24.3 (± 5.2)	29.8 (± 8.1)	17.6 (± 4.9)
CPE-Avg	18.1 (± 3.5)	23.5 (± 4.2)	30.1 (± 6.2)	17.5 (± 3.0)
CPL-Fin	122.4 (± 37.3)	166.1 (± 35.0)	172.0 (± 29.9)	166.5 (± 29.5)
CPL-Avg	181.3 (± 29.2)	188.4 (± 24.2)	200.4 (± 23.7)	219.6 (± 22.0)
PR-Fin	13.4 (± 6.0)	12.3 (± 4.3)	7.7 (± 1.2)	18.7 (± 7.4)
PR-Avg	19.7 (± 5.8)	19.3 (± 2.6)	16.0 (± 2.0)	27.6 (± 6.4)
RME-FT4	1.0 (± 0.6)	2.1 (± 0.3)	1.5 (± 0.5)	0.4 (± 0.3)
RME-FT5	1.6 (± 0.6)	**3.2*** (± 0.6)	2.5 (± 0.6)	0.8 (± 0.4)
RME-T4	1.3 (± 0.4)	**2.6*** (± 0.2)	2.0 (± 0.3)	1.3 (± 0.1)
RME-T5	1.3 (± 0.4)	**2.7*** (± 0.3)	2.1 (± 0.4)	1.2 (± 0.2)
RML-FT4	17.9 (± 7.6)	32.0 (± 3.8)	21.6 (± 6.6)	10.4 (± 3.7)
RML-FT5	20.8 (± 7.1)	36.0 (± 5.6)	31.3 (± 6.9)	12.4 (± 3.3)
RML-T4	21.9 (± 4.1)	**33.9*** (± 2.6)	29.3 (± 4.9)	21.5 (± 2.0)
RML-T5	21.3 (± 4.4)	**32.9*** (± 2.8)	30.6 (± 4.9)	19.1 (± 2.1)

7.3.2 Correlation Analysis

Table 7.2 portrays significant pairwise correlations ($p < .05$) observed between behavioral measures from the comprehensive task battery. Marked cells include both the correlation coefficient (r-value, top) and significance (p-value, bottom). Widespread intra-task correlations exist within the RAWM task and, to a lesser extent, within the Platform Recognition and Morris water maze tasks. Inter-task correlations were found between sensorimotor and cognitive tasks: Open Field activity and RAWM (all measures except RML-T4), Circular Platform errors, Y-maze arm entries, and Elevated Plus maze closed-arm entries; and, String Agility and Elevated Plus maze open-arm latency. These patterns underscore the interdependence among sensory, motor, and cognitive processes, as well as the role of anxiety-related components on behavior.

In addition, significant inter-task correlations were found between cognitive tasks: (1) Y-maze and all RAWM measures (except between YM-AE and RME-T5); (2) Morris water maze and Platform Recognition; (3) Morris water maze and RAWM latencies (and, to a lesser extent, errors); and, (4) Circular Platform latency and RAWM errors (and, to a lesser extent, latencies, as well). These observed correlations between behavioral measures which span multiple tasks is suggestive of shared, overlapping, or interdependent cognitive domains across tasks. For example, extensive intercorrelation between component measures of the Morris water maze, Platform Recognition, and RAWM tasks reflects the shared spatial memory dependency of these behavioral paradigms.

Finally, patterns in pairwise correlations underscore the multifactorial character of certain tasks. Significant correlation between the Elevated Plus maze closed-arm entries and both Circular Platform measures, as well as between Elevated Plus maze

open-arm entries and RAWM errors, suggests an anxiety/exploratory component within both the Circular Platform and RAWM tasks.

7.3.3 Factor Analysis

A varimax-rotated principal component analysis of all behavioral measures from the comprehensive task battery is shown in Table 7.3, with significant factor loadings (absolute value greater than 0.500) indicated. The eigenvalue results (i.e., >1 criterion) were consistent with scree plot (Cattell, 1966) identification of seven significant factors. The primary factor (accounting for about 28% of overall variance) was a cognitive-biased structure comprised of all RAWM measures, as well as the Y-maze percent alternations, consistent with correlation analysis. The second factor was comprised of Platform Recognition and Morris water maze measures, representing an object recognition/identification-related cognitive domain, distinct from the working memory-associated primary factor. Factor III includes a combination of sensorimotor (Balance Beam latency) and stimulus-avoidance measures (average Circular Platform latency and Elevated Plus maze closed-arm entries), and may reflect an escape/avoidance behavioral component. Similarly, the fourth factor contains measures associated with escape or avoidance behavior (average Circular Platform errors and Elevated Plus maze closed-arm entries), as well as an exploratory/activity measure (Open Field activity), which together may reflect escape/avoidance behaviors distinct from Factor III. The fifth factor includes only a single behavioral measure, Elevated Plus maze open-arm residency time, which is the only Elevated Plus maze measure linked to anxiety; thus, Factor V represents an anxiety factor. Factor VI also includes only a single measure, Elevated Plus maze open-arm entries, which may be related to exploratory behavior, distinct from Factor V. The sixth factor is comprised of a single sensorimotor measure, String Agility, which represents overall physical

Table 7.2. Correlations between behavioral measures of the comprehensive task battery in the GRK5 study

	OF	BB	SA	YM-AE	YM-PA	EP-CE	EP-OE	EP-TO	WM-Avg	WM-Fin	WM-Ret	CPE-Avg	CPL-Avg	PR-Avg	PR-Fin	RML-T4	RML-FT4	RML-T5	RML-FT5	RME-T4	RME-FT4	RME-T5
BB																						
SA																						
YM-AE	.44 .007																					
YM-PA																						
EP-CE	.35 .036																					
EP-OE					-.33 .050																	
EP-TO			-.34 .040																			
WM-Avg																						
WM-Fin									.90 .000													
WM-Ret																						
CPE-Avg	.42 .012					.35 .036																
CPL-Avg						-.39 .019																
PR-Avg									.60 .000	.62 .000												
PR-Fin									.65 .000	.68 .000		-.33 .050		.89 .000								
RML-T4				.45 .005	-.43 .008				.50 .001	.63 .000					.36 .032							
RML-FT4	.40 .014			.41 .011	-.34 .038				.47 .003	.61 .000			-.40 .015			.75 .000						
RML-T5	.37 .023			.35 .036	-.45 .005				.57 .000	.60 .000						.73 .000	.76 .000					
RML-FT5	.42 .010			.49 .002	-.48 .002				.39 .017	.41 .012			-.35 .036			.64 .000	.70 .000	.84 .000				
RME-T4	.37 .022			.50 .002	-.49 .002		.35 .035			.42 .010			-.34 .046			.84 .000	.65 .000	.59 .000	.59 .000			
RME-FT4	.43 .008			.37 .023	-.41 .013				.37 .023	.52 .001			-.41 .012			.66 .000	.96 .000	.68 .000	.67 .000	.65 .000		
RME-T5	.39 .018				-.48 .003		.40 .015		.42 .010	.53 .001			-.39 .019			.62 .000	.68 .000	.90 .000	.74 .000	.60 .000	.65 .000	
RME-FT5	.42 .010			.49 .002	-.51 .001		.44 .006						-.37 .027			.53 .000	.62 .000	.74 .000	.90 .000	.57 .000	.60 .000	.77 .000

Table 7.3. Varimax-rotated factor analysis of comprehensive task battery measures in the GRK5 study

Measure	Factor						
	I	II	III	IV	V	VI	VII
RML-T5	0.877						
RME-T5	0.862						
RML-FT5	0.826						
RME-FT4	0.820						
RML-FT4	0.815						
RME-FT5	0.779						
YM-PA	-0.665						
RML-T4	0.663						
RME-T4	0.629						
PR-Fin		0.911					
PR-Avg		0.882					
WM-Fin		0.812					
WM-Avg		0.810					
CPL-Avg			-0.743				
BB			0.719				
EP-CE			0.511	0.535			
CPE-Avg				0.772			
OF				0.768			
EP-TO					0.883		
EP-OE						0.757	
SA							-0.746
Variance	27.99%	15.61%	8.45%	9.11%	6.26%	5.74%	6.66%

strength, coordination, and grip capacity. Interestingly, the segregation of measures from the Elevated Plus maze task underscores the multifactorial character of this behavioral task, which includes activity and escape/avoidance (anxiety-related).

Table 7.4. Classifier performance comparison using comprehensive task battery measures in the GRK5 study

Groups	Evaluation Criterion	Discriminant Analysis		Decision Tree	Neural Network	Support Vector Machine
		Complete	Step-Fwd			
NT	Accuracy	NS	84%	74%	89%	74%
vs.	Sensitivity	NS	92%	75%	92%	83%
Tg	Specificity	NS	71%	71%	86%	57%
NT	Accuracy	NS	87%	NS	67%	53%
vs.	Sensitivity	NS	100%	NS	63%	NS
NT-KO	Specificity	NS	71%	NS	71%	57%
Tg	Accuracy	NS	81%	71%	52%	62%
vs.	Sensitivity	NS	67%	67%	NS	NS
Tg-KO	Specificity	NS	92%	75%	58%	75%
NT	Accuracy	NS	81%	NS	56%	NS
vs.	Sensitivity	NS	89%	NS	NS	NS
Tg-KO	Specificity	NS	71%	NS	71%	NS
All four	Accuracy	47%	56%	36%	39%	28%

7.3.4 Discriminant Analysis

Table 7.4 compares the performance of classifiers utilizing behavioral measures from the comprehensive task battery, for distinguishing between pairs of groups, as well as among the four groups of animals. The third and fourth columns of the table display the results of discriminant analysis-based classifiers using the direct-entry (complete) and stepwise-forward methods, respectively. Evaluation criteria for classifier performance are provided only when the observed level strictly exceeds expected values associated with random-chance assignment of individuals to groups (i.e., 50% for two-group discriminability, 25% for four-group discriminability).

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

Although the direct-entry (complete) discriminant analysis failed to distinguish between the NT and Tg groups, the stepwise-forward approach returned significant discriminability (Wilks's lambda = 0.191, p = .0003) between the groups. As indicated in Table 7.4, the classifier exhibited 84% accuracy, with 92% sensitivity and 71% specificity, with respect to transgenicity. Five behavioral measures were selected on the basis of variance contribution: SA, YM-PA, CPE-Avg, RME-T4, and RME-T5. The emphasis on cognitive measures, particularly the RAWM working memory T4 and T5 components, underscores the cognitive impairment of Tg mice, relative to NT controls.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

Direct-entry discriminant analysis was unable to distinguish between the NT and NT-KO groups, however the variance-optimizing stepwise-forward approach performed remarkably well (Wilks's lambda = 0.317, p = .0044). Three predictor variables were selected for the model – OF, EP-TO, and RME-FT5 – which represent a cross-section of sensorimotor, anxiety-related, and cognitive measures. Eighty-seven percent of individuals were correctly identified by group, including all of the NT-KO animals and 71% of the NT controls. This suggests that the absence of GRK5 expression in a nontransgenic animal produces a subtle, albeit discernable, characteristic behavioral phenotype. Relative to NT mice, NT-KO animals showed decreased locomotor activity (OF), less anxiety (greater EP-TO latency), and superior working memory (fewer final-block RAWM trial T5 errors).

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

As shown in Table 7.4, the standard direct-entry approach did not successfully distinguish between the two groups, although the stepwise-forward analysis returned significant discriminability (Wilks's lambda = 0.610, p = .0116) using only two be-

havioral measures, BB and RME-T4. The classifier displayed 81% overall accuracy, with 67% sensitivity and 92% specificity, with respect to the Tg-KO group. Hence, the overall performance was largely driven by the classifier's identification of the Tg control group. Interestingly, a sensorimotor measure (balance beam latency) and a robust cognitive measure (RAWM overall T4 errors), together, were found to provide optimal discriminability, although (as shown in Table 7.1) these two measures, individually, do not differ significantly between the groups. The Tg-KO mice displayed better balance/coordination (higher BB) and better cognitive function (fewer overall RAWM trial T4 errors), compared with Alzheimer's transgenic-control animals.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

The direct-entry discriminant analysis did not distinguish between the two groups, although the stepwise-forward approach was successful (Wilks's lambda = 0.199, p = .0028). The classifier displayed 81% accuracy, with 89% sensitivity and 71% specificity, with respect to the Tg-KO group. Five behavioral measures (BB, EP-TO, YM-PA, PR-Fin, and RML-FT4) were included in the model. As described earlier, with reference to Table 7.1, these two groups do not differ significantly with respect to any of these measures. Moreover, the classifier is likely exhibiting sensitivity to a composite behavioral phenotype, which emerges only through sampling multiple sensorimotor, anxiety, and/or cognitive features (utilizing the comprehensive task battery). Relative to nontransgenic-control animals, the Tg-KO mice exhibited poorer balance/coordination, more anxiety, poorer overall mnemonic function (decreased YM-PA), but markedly superior object recognition/identification memory (decreased PR-Fin latency). However, the RML-FT4 measure was slightly lower in Tg-KO mice, suggesting better working memory function, relative to the nontransgenic group.

All Four Groups

Statistically-significant discriminability among the four groups was returned by the direct-entry discriminant analysis (Wilks's lambda = 0.010, p = .0275), with 47% overall accuracy. Individuals of the Tg group were the most likely to be correctly identified (67%), while only 22% of the Tg-KO mice were detected. Similarly, the stepwise-forward approach demonstrated significant and, moreover, superior discriminability (Wilks's lambda = 0.227, p = .0000). Only four behavioral measures were retained by the model (OF, EP-TO, YM-PA, and RME-T4), which accurately classified 56% of animals by group. The NT-KO mice were the most likely to be correctly identified (75%), while only 29% of NT mice were recognized.

7.4 Results of Data Mining Analyses: Comprehensive Task Battery

A comparative summary of the performance of data mining-based classifiers utilizing behavioral measures from the comprehensive task battery is provided in Table 7.4, for reference. The fifth through seventh columns of the table report the performance evaluation (i.e., accuracy, sensitivity, and specificity) for each classifier, with respect to discriminability between pairs of groups, as well as among all four groups.

7.4.1 Decision Tree Analysis

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

The decision tree identified a single behavioral measure, RME-T5, as providing sufficient information bias to distinguish between NT and Tg mice. The resulting classifier correctly identified 74% of animals by group (Kappa = 0.45), with 75% sensitivity and 71% specificity, with respect to transgenicity. Alzheimer's transgenic-control mice exhibited significant cognitive impairment (i.e., increased errors in overall RAWM trial T5), relative to nontransgenic-control animals, as indicated in Table

7.1. This result underscores the primacy of the RAWM task for detecting cognitive impairment and, specifically, overall measures of working memory.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

Decision tree-based classifiers were unable to distinguish between nontransgenic mice with the GRK5-wildtype and animals having the GRK5-knockout genotype. As reported in Table 7.4, none of the behavioral measures of the comprehensive task battery provided sufficient information bias to reliably discriminate between the two groups. This suggests comparable behavioral performance between the two groups, i.e., the absence of GRK5 expression in a nontransgenic mouse does not significantly impact the overall behavioral phenotype of the mouse.

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

The RME-T5 and RME-FT4 measures were selected by the decision tree for their information value in distinguishing between Alzheimer's transgenic mice having the GRK5-wildtype and animals with the GRK5-knockout genotype. The classifier generated by the decision tree exhibited 71% overall accuracy (Kappa = 0.42), 67% sensitivity and 75% specificity, with respect to the GRK5-knockout group. Although the classifier exhibited superior accuracy for identifying Tg mice, relative to Tg-KO animals, the choice of two robust cognitive measures of RAWM working memory (indeed, both final-block T4 errors and overall T5 errors) highlights the importance of the RAWM task for evaluating cognitive impairment. Interestingly, although these two groups differ significantly in only one of these two measures (RME-T5), as shown in Table 7.1, both measures together provide remarkable discriminative potential. Alzheimer's transgenic-control mice (Tg) exhibit significant working memory deficits, relative to transgenics which do not express GRK5. This results suggests that expression of GRK5, superimposed against an Alzheimer's transgenic back-

ground, may ameliorate working memory dysfunction associated with the APPsw genotype.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

As shown in Table 7.4, individuals from the nontransgenic and transgenic-knockout groups could not be distinguished using any combination of behavioral measures from the comprehensive task battery by decision tree-based classifiers. No individual measure, or combination of measures, provided sufficient information bias to reliably split the database into the two groups. These findings suggest that the absence of GRK5 expression in an Alzheimer's transgenic mouse may sufficiently mitigate the behavioral manifestations of Alzheimer's-like neuropathology, as to render these individuals comparable to nontransgenic animals, using the comprehensive task battery assessment.

All Four Groups

Eight measures from the comprehensive task battery, representing a diverse sampling of the mouse behavioral repertoire, were identified by the decision tree for distinguishing among the four groups of animals. These measures were: OF, RME-FT4, RME-FT5, WM-Avg, EP-OE, YM-PA, BB, and RML-FT4. This cognitively-biased subset of measures accurately identified 36% of all animals by group (Kappa = 0.14), as indicated in Table 7.4.

7.4.2 Neural Network Analysis

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

The neural network-based classifier performed remarkably well, accurately identifying 89% of animals by group (Kappa = 0.77). The sensitivity of the classifier was 92% and the specificity was 86%, with respect to transgenicity. This level of perfor-

mance underscores the relative ease of identifying impaired cognitive performance in the Alzheimer's transgenic animals, relative to the nontransgenic mice.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

As shown in Table 7.4, only two-thirds of all nontransgenic animals were distinguishable on the basis of GRK5 genotype (Kappa = 0.34). Animals with the GRK5-knockout genotype were more likely to be misclassified, relative to mice having the GRK5-wildtype (63% sensitivity vs. 71% specificity). Although these two groups do not differ significantly with respect to individual cognitive measures, nontransgenic mice expressing GRK5 were significantly more active than were GRK5-knockout animals, as measured by the Open Field task (i.e., NT-KO mice displayed significantly fewer line-crossings in the open arena).

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

The performance of neural network-based classifiers for distinguishing between Alzheimer's transgenic mice on the basis of GRK5 expression was only slightly superior to random-chance assignment (52% overall accuracy, Kappa = 0.03). Indeed, detection of Tg-KO animals failed to reach criterion (50%).

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

The level of discriminability between nontransgenic animals and mice having both genetic manipulations (Alzheimer's transgenic with GRK5-knockout) was only 56% overall (Kappa = 0.15), as reported in Table 7.4. This performance was largely driven by accuracy in identifying the NT mice (specificity = 71%), because an insufficient proportion of Tg-KO animals were identified by the classifier to reach criterion.

All Four Groups

Thirty-nine percent of all animals were correctly assigned to their respective groups by the neural network-based classifier using behavioral measures from the comprehensive task battery (Kappa = 0.17).

7.4.3 Support Vector Machine Analysis

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

A support vector machine-based classifier trained using all the behavioral measures from the comprehensive task battery successfully distinguished between nontransgenic and Alzheimer's transgenic mice, with 74% overall accuracy (Kappa = 0.42). The sensitivity was 83% and specificity was 57%, as shown in Table 7.4. The observed level of discriminability between these two groups is not surprising, because Alzheimer's transgenic mice differ significantly from nontransgenics in both sensorimotor and cognitive measures. Compared with NT mice, Tg animals explore the open-arms of the Y-maze more frequently, but exhibit poorer systematic search of the Y-maze (relatively fewer arm-visit alternations; YM-PA). The Alzheimer's transgenic mice also display poorer working memory, relative to nontransgenics, as indicated by higher error-scores (RME-T4, and both RME-T5 and RME-FT5) and longer response-latencies (both RML-T4 and RML-T5) in the RAWM trials T4 and T5.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

The support vector machine-based classifier obtained only 53% overall accuracy (Kappa = 0.07) in distinguishing between nontransgenic animals on the basis of GRK5 expression. Classifier-based assignment of NT-KO mice was comparable to random-chance performance, however.

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

As indicated in Table 7.7, sixty-two percent of Alzheimer's transgenic mice were correctly identified on the basis of GRK5-genotype. Although 75% of Tg animals expressing GRK5 were recognized, fewer than half of the mice lacking GRK5 expression were successfully detected.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

Support vector machine-based classifiers failed to distinguish between NT and Tg-KO animals using the behavioral measures from the comprehensive task battery. Classifier performance never exceeded the accuracy expected by random assignment of individuals to groups.

All Four Groups

As reported in Table 7.4, only 28% of all animals were accurately identified by group by support vector machines (Kappa = 0.03). This level of performance only slightly exceeds the accuracy expected by random assignment of individuals to groups.

7.5 Results of Statistical Analyses: Interference Paradigm

7.5.1 Standard Behavioral Analysis

The behavioral performance of all four groups in each measure of the interference paradigm (both error-scores and response-latencies) are shown in Figure 7.1, on the next page, with group means and associated standard errors depicted for each measure. Based upon standard statistical tests (ANOVA, F-test), the Tg group differed significantly from the NT group, denoted by an asterisk ($*$), with respect to both errors and latencies in the three-trial recall, retroactive interference, and delayed-recall tasks (all $p \leq .05$). No significant differences were found between the NT and NT-KO groups, with respect to any of the behavioral measures in the interference paradigm. Similarly, although the Tg and Tg-KO groups differed with respect to the three-trial recall error-score ($p = .075$), this contrast did not meet the significance criterion ($p \leq .05$). Finally, NT and Tg-KO animals exhibit significant ($p \leq .05$) groupwise differences in cognitive performance, as revealed by the retroactive

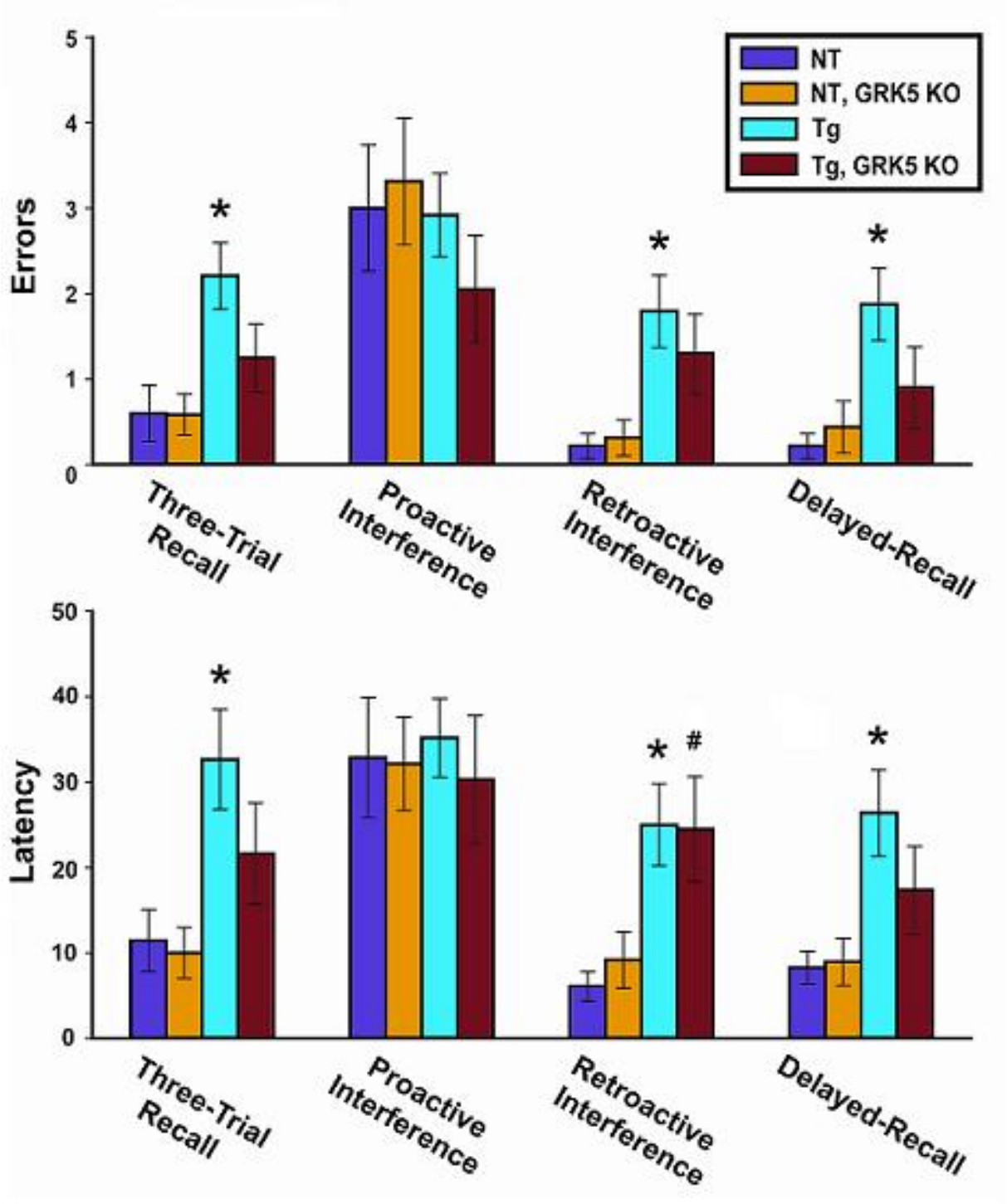

Figure 7.1. Groupwise contrasts for all interference paradigm measures in the GRK5 study

interference task (latencies, but not errors), and indicated with a pound-sign symbol (#). This finding suggests that normal (i.e., comparable to nontransgenic animals) cognitive performance was observed in Alzheimer's transgenic animals which do not express GRK5 in all tasks/measures except retroactive interference, wherein the Tg-KO mice showed impaired context-switching from the interference-context (Pool B) back to the original learning condition (Pool A), relative to nontransgenics. Moreover, it was the response-latency measure of the retroactive interference task which demonstrated greater sensitivity to the Tg-KO behavioral phenotype.

7.5.2 Correlation Analysis

Table 7.5. Correlations between behavioral measures of the interference paradigm in the GRK5 study

TT-L	.72 .001						
PI-E							
PI-L			.95 .000				
RI-E	.65 .003	.68 .001					
RI-L	.63 .004	.75 .000			.97 .000		
DR-E	.61 .005	.73 .000			.87 .000	.85 .000	
DR-L	.54 .017	.77 .000			.89 .000	.93 .000	.95 .000
	TT-E	TT-L	PI-E	PI-L	RI-E	RI-L	DR-E

Significant ($p < .05$) correlations observed between pairs of measures, for all groups, are indicated in Table 7.5. Marked cells in the table include both the correlation coefficient (r-value, top) and significance (p-value, bottom). The abbreviations for the four tasks of the mouse-based interference paradigm are: Three-trial recall,

errors and latency (TT-E, TT-L); proactive interference, errors and latency (PI-E, PI-L); retroactive interference, errors and latency (RI-E, RI-L); and, delayed-recall, errors and latency (DR-E, DR-L).

There were positive correlations between the error-score measures of the three-trial recall and both retroactive interference and delayed-recall tasks. A similar pattern of association was observed among the corresponding latency measures. This is consistent with acquisition of the initial learning component (Pool A) being predictive of later performance in the same spatial environment (i.e., Pool A). The error-score and response-latency measures of the proactive interference task were significantly intercorrelated. However, neither measures from the proactive interference task was significantly correlated with any measures from the other three tasks, suggesting independence of performance between the two learning contexts (Pool A vs. Pool B). Significant correlations were found between corresponding error-score and response-latency measures for the other three tasks, as well (e.g., three-trial recall errors and latency), which underscores the comparability of these two indices of performance within tasks of the mouse-based interference paradigm.

Additional intercorrelations across cognitive components, as well as between error and latency measures, were detected. Both measures from the three-trial recall task were significantly intercorrelated with both measures of retroactive interference, suggesting consistent/stable acquisition and retention of the initial learning component (Pool A), despite the influence of both a distractor task (Y-maze exposure) and the interference task (Pool B exposure). Similarly, complete intercorrelation between all measures of the retroactive interference and delayed-recall tasks suggests temporal stability of learning. This permanence is further underscored by significant pairwise intercorrelation among all measures of the three-trial recall and delayed-recall tasks.

7.5.3 Factor Analysis

Table 7.6. Varimax-rotated factor analysis of interference paradigm measures in the GRK5 study

	Factor	
Measure	I	II
Retroactive Interference (Latency)	0.956	
Delayed-Recall (Latency)	0.948	
Retroactive Interference (Errors)	0.944	
Delayed-Recall (Errors)	0.934	
Three-Trial Recall (Latency)	0.855	
Three-Trial Recall (Errors)	0.744	
Proactive Interference (Errors)		0.992
Proactive Interference (Latency)		0.981
Variance	60.99%	24.61%

A varimax-rotated principal component analysis of the eight measures is shown in Table 7.6. Significant (absolute value greater than 0.700) component loadings are indicated for the two factors returned. Additionally, the calculated eigenvalue results (i.e., >1 criterion) were in agreement with scree plot (Cattell, 1966) identification of two significant factors. The primary factor (accounting for approximately 60% of overall variance) consisted of error-score and response-latency measures from the Three-Trial Recall, Retroactive Interference, and Delayed-Recall tasks, and represents spatial learning and memory function associated with the initial learning condition (Pool A). By contrast, the second factor (approximately 25% of variance) includes only the Proactive Interference measures, and likely represents spatial learning and memory related to the interference condition (Pool B).

Table 7.7. Classifier performance comparison using interference paradigm measures in the GRK5 study

Groups	Dataset	Evaluation Criterion	Discriminant Analysis		Decision Tree	Neural Network	SVM
			Complete	Step-Fwd			
NT vs. Tg	Errors	Accuracy	NS	74%	63%	68%	53%
		Sensitivity	NS	67%	67%	75%	75%
		Specificity	NS	86%	57%	57%	14%
	Latency	Accuracy	NS	79%	68%	53%	NS
		Sensitivity	NS	75%	67%	58%	NS
		Specificity	NS	86%	71%	NS	NS
	Err+Lat	Accuracy	NS	74%	58%	63%	63%
		Sensitivity	NS	67%	58%	75%	67%
		Specificity	NS	86%	57%	57%	57%
NT vs. NT-KO	Errors	Accuracy	NS	NS	NS	NS	NS
		Sensitivity	NS	NS	NS	NS	NS
		Specificity	NS	NS	NS	NS	NS
	Latency	Accuracy	NS	NS	NS	NS	NS
		Sensitivity	NS	NS	NS	NS	NS
		Specificity	NS	NS	NS	NS	NS
	Err+Lat	Accuracy	NS	NS	NS	NS	NS
		Sensitivity	NS	NS	NS	NS	NS
		Specificity	NS	NS	NS	NS	NS
Tg vs. Tg-KO	Errors	Accuracy	NS	NS	73%	NS	NS
		Sensitivity	NS	NS	80%	NS	NS
		Specificity	NS	NS	67%	NS	NS
	Latency	Accuracy	NS	NS	NS	NS	NS
		Sensitivity	NS	NS	NS	NS	NS
		Specificity	NS	NS	NS	NS	NS
	Err+Lat	Accuracy	NS	NS	55%	NS	NS
		Sensitivity	NS	NS	NS	NS	NS
		Specificity	NS	NS	58%	NS	NS
NT vs. Tg-KO	Errors	Accuracy	NS	77%	53%	53%	53%
		Sensitivity	NS	80%	60%	NS	90%
		Specificity	NS	71%	NS	58%	NS
	Latency	Accuracy	NS	88%	88%	71%	NS
		Sensitivity	NS	80%	100%	70%	NS
		Specificity	NS	100%	71%	71%	NS
	Err+Lat	Accuracy	NS	88%	88%	65%	NS
		Sensitivity	NS	90%	100%	60%	NS
		Specificity	NS	86%	71%	71%	NS
All four	Errors	Accuracy	30%	32%	35%	30%	27%
	Latency	Accuracy	NS	46%	32%	NS	NS
	Err+Lat	Accuracy	NS	32%	43%	30%	NS

7.5.4 Discriminant Analysis

A summary table comparing classifier performance is shown in Table 7.7. The fourth and fifth columns of the table display the results of discriminant analysis-based classifiers using the direct-entry (complete) and stepwise-forward methods, respectively. Measures of classifier performance (i.e., accuracy, sensitivity, and specificity) are depicted only when the observed level strictly exceeds expected values associated with random-chance assignment of individuals to groups (i.e., 50% for two-group discriminability, 25% for four-group discriminability).

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

A direct-entry discriminant analysis did not return significant discriminability between the two groups (Wilks' lambda = 0.585, p = .0913) using only error-score data. By contrast, a stepwise-forward analysis of error-scores was successful (Wilks lambda = 0.643, p = .0069) with 74% overall accuracy (sensitivity = 67%, specificity = 86%), as shown in Table 7.7, wherein only the delayed-recall error measure was retained by the stepwise-forward model. Direct-entry discriminant analysis using only response-latency data failed to distinguish between the two groups, although stepwise-forward analysis returned significant discriminability (Wilks lambda = 0.677, p = .0111) with 79% accuracy (sensitivity = 75%, specificity = 86%), retaining only the three-trial recall latency measure as the predictor variable. When both error-scores and response-latencies were included, only the stepwise-forward analysis successfully distinguished individuals by transgenicity (Wilks lambda = 0.643, p = .0069) with 74% accuracy (sensitivity = 67%, specificity = 86%), as depicted in Table 7.7. The delayed-recall error-score was the only behavioral measure (of the eight provided) to be retained by the stepwise-forward model. Hence, stepwise-forward, but not direct-entry, discriminant analyses successfully distinguished between nontransgenic-

control and transgenic-control mice, utilizing either the delayed-recall error-score or the three-trial recall latency measure. Between these two measures, however, the former was shown to be superior when all measures were available for selection. Animals in the Tg group exhibited higher (i.e., poorer) values for both of these measures, relative to mice in the NT group.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

Neither direct-entry nor stepwise-forward discriminant analyses showed significant discriminability between the two groups, using either error-scores or response-latencies, or both. The stepwise-forward analysis indicated that none of the candidate predictor variables met the criterion for model inclusion (alpha-to-enter = 0.15). Thus, removal of GRK5 expression had no effect on behavioral performance in normal (nontransgenic) mice.

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

Both direct-entry and stepwise-forward discriminant analyses failed to distinguish between the two groups using error-scores, response-latencies, or both behavioral measures. None of the candidate predictor variables met the statistical criterion (alpha-to-enter = 0.15) to be added to the stepwise-forward model. Similar to the results obtained for nontransgenic animals (GRK5-wildtype vs. knockout), neither discriminant analysis-based classifier distinguished between GRK5-wildtype and knockout genotype in Alzheimer's transgenic mice. Hence, removal of GRK5 expression did not enhance the behavioral impairment of Tg mice.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

The direct-entry discriminant analysis was unable to distinguish significantly between the two groups using error-scores and/or response-latencies. By contrast, the stepwise-forward method returned significant groupwise discriminability for all three dataset conditions. Using only error-scores, the stepwise-forward classifier retained

the proactive interference and retroactive interference measures, and demonstrated 77% overall accuracy (Wilks's lambda = 0.627, p = .0379) with 80% sensitivity and 71% specificity. Similarly, when response-latencies were used, only the proactive interference and retroactive interference measures were retained by the model, exhibiting 88% accuracy (Wilks's lambda = 0.606, p = .0299) with 80% sensitivity and 100% specificity; all nontransgenic animals were correctly identified by the stepwise-forward classifier using latency data alone. When both error-scores and response-latencies were provided, the stepwise-forward analysis retained the proactive interference error-score and the retroactive interference latency measures, and showed 88% accuracy (Wilks's lambda = 0.549, p = .0151), 90% sensitivity and 86% specificity. Hence, the NT and Tg-KO groups were more easily distinguishable than NT vs. Tg, suggesting additional cognitive impairment, beyond the APP transgenic effect, may be attributed to the absence of GRK5 expression.

All Four Groups

Direct-entry discriminant analysis reported significant discriminability among the four groups (Wilks' lambda = 0.515, p = .0475), using only error-score data. However, the overall classification accuracy was 30%. A stepwise-forward analysis also returned significant discriminability (Wilks lambda = 0.675, p = .0043) with modest 32% overall accuracy, as shown in Table 7.7. Only the delayed-recall error measure was retained by the stepwise-forward model. By contrast, using only response-latency data, only the stepwise-forward discriminant analysis significantly distinguished among the four groups (Wilks lambda = 0.681, p = .0049) with moderate 46% accuracy, as reported in Table 7.7. The only predictor variable included in the stepwise-forward model was the three-trial recall latency measure. Similarly, when both errors and latencies were provided, only the stepwise-forward approach returned significant discriminability (Wilks lambda = 0.675, p = .0043) with 32% accuracy,

as portrayed in Table 7.7. The delayed-recall error measure was the only predictor variable retained by the model for distinguishing among the four groups. Hence,the delayed-recall error measure was identified as the best discriminator among the four groups of mice, although the three-trial recall latency measure also demonstrated potential for distinguishing among the groups. It should be noted, however, that the overall levels of accuracy observed (i.e., 30% to 46%) are only modestly superior to the random-chance performance criterion for four-group discriminability (25%).

7.6 Results of Data Mining Analyses: Interference Paradigm

Table 7.7 compares the performance of classifiers based on advanced statistical techniques (columns four and five) and data mining-based methods (columns six through eight). Performance measures (i.e., accuracy, sensitivity, and specificity) are reported only when the criterion for random-chance assignment (i.e., 50% for two-group discriminability, 25% for four-group discriminability) is exceeded.

7.6.1 Decision Tree Analysis

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

Using only error-score data, the decision tree attempted to split the dataset using three measures (delayed-recall, proactive interference, and three-trial recall), resulting in 63% correct classification of cases (Kappa = 0.23), as shown in Table 7.7. The sensitivity was 67% and specificity was 57%, with respect to transgenicity. When only response-latency data were provided, the decision tree-based classifier used two measures (three-trial recall and proactive interference) to distinguish between the two groups, resulting in 68% overall accuracy (Kappa = 0.36). For the latency-based classification, the classifier sensitivity was 67% and specificity was 71%. Including both error-scores and response-latencies resulted in a decision tree-based classifier

with three measures (delayed-recall errors, proactive interference errors, and three-trial recall errors) selected to distinguish between the two groups, and displaying 58% correct classification of cases (Kappa = 0.15). The sensitivity was 58% and specificity was 57%, as indicated in Table 7.7. These analyses suggest that measures from the three-trial recall and proactive interference tasks most reliably distinguish animals by transgenicity (Alzheimer's transgenic vs. nontransgenic), and the delayed-recall error measure also contributes to the discriminability. However, the level of discriminability (i.e., 58% to 68%) observed was only slightly better than random-chance assignment expectation for two-group classification (50%).

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

Decision tree-based classifiers were unable to distinguish between GRK5-wildtype and knockout genotype nontransgenic animals using either error-score, response-latency, or both sets of behavioral measures. None of the candidate predictor variables met the information gain criterion required by the decision trees for splitting the dataset into groups.

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

Using only the error-score measures, decision tree-based classifiers were able to distinguish between the two groups using only the delayed-recall measure, resulting in 73% overall accuracy (Kappa = 0.46). The sensitivity was 80% and specificity was 67%, as indicated in Table 7.7. However, the classifier was unable to generate a decision tree to distinguish between the two groups using only response-latencies. When provided with both error-scores and reponse-latencies, the decision tree-based classifier selected two measures (delayed-recall errors and retroactive interference latency) to distinguish between the two groups, resulting in 55% correct classification of cases (Kappa = 0.08), as reported in Table 7.7. The classifier's sensitivity was only 50% (comparable to random-chance assignment) and specificity was 58%. Taken together,

these results suggest the significant discriminability observed for errors+latencies is driven by the error-score data, since: (1) Response-latency measures alone do not provide discriminability, and (2) the level of discriminability (accuracy) was much higher for error-scores alone, relative to errors+latencies combined.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

The decision tree-based classifier was only modestly capable of distinguishing between the two groups using error-scores alone, displaying only 53% accuracy (Kappa = 0.03) with 60% sensitivity, utilizing the proactive interference and retroactive interference measures. By contrast, 88% accuracy (Kappa = 0.75) was achieved using only response-latencies, wherein only the retroactive interference measure was selected for its information-bias capacity. All of the Tg-KO animals were correctly identified (100% sensitivity) using latency data alone. Similarly, when both error-score and response-latency measures were available, the overall accuracy was 88% (Kappa = 0.75), with 100% sensitivity and 71% specificity, using only the retroactive interference response-latency measure to split the dataset. Hence, groupwise differences in a single behavioral measure (retroactive interference latency) are sufficient to distinguish between NT and Tg-KO animals.

All Four Groups

The decision tree attempted to distinguish among the four groups using only error-score data, and identified two behavioral measures (retroactive interference and three-trial recall) which optimally classified individuals into their respective groups. As shown in Table 7.7, 35% correct classification of cases (Kappa = 0.13) was observed. When only response-latency data were provided, the decision tree accurately distinguished 32% of the individual animals (Kappa = 0.07) using all four latency measures. None of the nontransgenic-control (NT) animals were correctly identified using only error-score or response-latency data, however. Providing the decision

tree with both error-score and response-latency data for the four groups resulted in selection of four measures (retroactive interference errors, three-trial recall errors, delayed-recall errors, and retroactive interference latency), and overall accuracy of 43% (Kappa = 0.24), as indicated in Table 7.7. Hence, discriminability among groups by decision tree-based classifiers appears significant statistically, but not appreciably above random-chance levels (i.e., 25%).

7.6.2 Neural Network Analysis

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

When either error-scores or response-latencies were used, the optimal architecture of the neural network-based classifier consisted of four input-layer computing units (the behavioral measures), four hidden-layer units, and two output-layer units (the two groups). For errors-only, the classifier showed 68% overall accuracy (Kappa = 0.32), with 75% sensitivity and 57% specificity, with respect to transgenicity, as indicated in Table 7.7. Using only latency data, the overall accuracy was 53% (Kappa = 0.01), sensitivity was 58% and specificity was 43%. Hence, error-scores distinguish between nontransgenic and Alzheimer's transgenic animals more accurately than do response-latencies; three out of four transgenic mice were correctly identified. Using both errors and latencies, the optimal neural network architecture included eight input-layer computing elements (the behavioral measures), four hidden-layer units, and two output-layer units (the two groups). The resulting classifier correctly identified 63% of cases (animals) by group (Kappa = 0.23). The sensitivity was 75% and specificity was 57%. The inclusion of latency measures with the error-score data, therefore, undermined overall discriminability between the groups.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

None of the neural network-based classifiers successfully distinguished between GRK5-wildtype and knockout genotype in nontransgenic mice using either error-scores, response-latencies, or both. The resulting neural networks did not reliably classify individuals by group, based on results from multiple program executions.

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

Neural network-based classifiers trained using error-scores, response-latencies, or both, failed to distinguish reliably between GRK5-wildtype and knockout genotype, superimposed on an Alzheimer's transgenic mouse background. Indeed, significant classifier performance instability was observed across repreated program runs.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

The neural network showed only marginal performance, relative to random-chance assignment, with only 53% of animals correctly identified (Kappa = 0.07) using only error-score measures. Indeed, only 58% of nontransgenics were correctly identified (specificity), while the sensitivity was nonsignificant. Latency measures alone, by contrast, accurately distinguished 71% of the animals by group (Kappa = 0.41), with 70% sensitivity and 71% specificity. When both error-scores and response-latencies were provided, however, the performance declined to 65% accuracy (Kappa = 0.30), with only 60% sensitivity and 71% specificity. These results suggest that response-latencies drive discriminability between the two groups, with respect to neural network-based classifiers.

All Four Groups

Using only error-score data, the optimal neural network-based classifier consisted of four input-layer computing units (the behavioral measures), three hidden-layer units, and four output-layer units (the four groups). The classifier demonstrated 30% overall accuracy (Kappa = 0.05), as reported in Table 7.7, although none of

the control animals (NT or Tg) were correctly identified. Classifiers trained using only response-latency data did not successfully distinguish among the four groups. The optimal classifier utilizing both error-scores and response-latencies, however, was comprised of eight input-layer computing elements (the behavioral measures), three hidden-layer units, and four output-layer units (the four groups), and correctly identified 30% of cases (animals) by group (Kappa = 0.05). Interestingly, none of the nontransgenic animals (NT and NT-KO) were correctly identified. Hence, the inclusion of latency measures with error measures neither improved nor compromised overall discriminability among the four groups, relative to exclusive use of error-scores as predictor variables. Moreover, the observed discriminability among the four groups (30%) was only slightly better than classification performance expected by random-assignment (25%).

7.6.3 Support Vector Machine Analysis

Nontransgenic-Control (NT) vs. Transgenic-Control (Tg) Groups

The support vector machine-based classifier obtained only 53% overall accuracy (Kappa = -0.12) using only error-score measures, with 75% sensitivity and 14% specificity, as shown in Table 7.7. Most animals were classified as Tg, regardless of actual genotype. By contrast, the support vector machine was unable to distinguish between groups solely on the basis of response-latency measures; all NT animals were misclassified as transgenic. By including both error-scores and response-latencies, however, the classifier demonstrated 63% overall accuracy (Kappa = 0.23), as indicated in Table 7.7. The sensitivity was 67% and specificity was 57%.

Nontransgenic-Control (NT) vs. Nontransgenic-Knockout (NT-KO) Groups

Support vector machine-based classifiers were unable to distinguish between GRK5-wildtype and knockout genotype nontransgenic animals using behavioral error-scores,

response-latencies, or both. Classifier performance never exceeded the accuracy expected by random assignment of individuals to groups (i.e., 50%).

Transgenic-Control (Tg) vs. Transgenic-Knockout (Tg-KO) Groups

Alzheimer's transgenic animals, with or without the GRK5-knockout genotype, were not distinguishable by support vector machine-based classifiers using either error-scores, response-latencies, or both. The performance accuracy never exceeded 50%, the expected value based on random assignment of individuals.

Nontransgenic-Control (NT) vs. Transgenic-Knockout (Tg-KO) Groups

Only the support vector machine-based classifier trained using error-score measures alone was able to distinguish between these two groups, and performed only slightly better than random-chance assignment, with 53% accuracy (Kappa = -0.12). Indeed, although 90% of the Tg-KO animals were correctly identified (sensitivity), all of the NT animals were incorrectly identified as Tg-KO. Hence, this classifier performed similarly to the NT vs. Tg classifier, with respect to discriminability and assignment profile.

All Four Groups

Using only error-scores for training, the classifier correctly identified 27% of all animals by group (Kappa = 0.001), as reported in Table 7.7. Transgenic-control (Tg) mice were the most likely to be correctly identified (sensitivity = 67%), however none of the nontransgenic-control (NT) or transgenic-knockout (Tg-KO) were recognized by the classifier. By contrast, the classifier was unable to distinguish among the four groups using only response-latency data. When both error-scores and response-latencies were used for training the classifier, the overall discriminability among groups was not significant (i.e., less than expected level attributed to random assignment). Hence, the inclusion of response-latency data undermined the modest potential of error-score measures for distinguishing among the mice by group.

Reminiscent of the four-group classification performance of decision trees and neural networks, the observed discriminability was only marginally superior to random-chance assignment (using error-scores alone).

7.7 Discussion

The cognitive deficits noted by Suo et al. (2007) in aged nontransgenic GRK5-knockout mice were confined to specific domains. Indeed, basic mnemonic function, reference memory and spatial reference learning, as well as object identification ability were comparable to age-matched GRK5-wildtype animals (Suo et al., 2007). At 17 to 19 months of age, no deficits were observed in the Y-maze (percent alternations), Morris water maze (both acquisition and retention), circular platform, and platform recognition tasks (Suo et al., 2007) between GRK5-wildtype and knockout mice. However, short-term (working) memory impairment was observed in GRK5-knockout animals (i.e., increased RAWM T4 and T5 errors), relative to wildtype (Suo et al., 2007), based on one-way ANOVA across days.

In the present study, a different group of GRK5-knockout mice was investigated, along with combined GRK5-knockout and Tg2576 (APP) mice. The purpose was to determine if any greater cognitive impairment was evident in the combined-genotype mice, relative to either GRK5-knockout or Tg2576 alone. All animals completed both the comprehensive behavioral task battery and the novel mouse-based interference paradigm. As discussed earlier, neurobehavioral investigations of group-differences (e.g., transgenicity, therapeutic efficacy) typically utilize a relatively small subset of multivariate statistical approaches. Moreover, analysis of variance (ANOVA, F-test) and/or regression are frequently the only methods considered, while more-appropriate techniques may be discounted (or, indeed, ignored), despite the widespread availability of powerful computing software and reference doc-

umentation. For this reason, the present study utilized both standard and advanced statistical methods for analyzing the behavioral measures obtained from both behavioral assessment regimens (i.e., comprehensive task battery, mouse-based interference paradigm). In addition to statistical methodologies, data mining-based approaches were also utilized for determining groupwise discriminability, both between pairs of groups and among all groups of mice.

Standard, ANOVA-based groupwise comparisons were performed using all measures of the comprehensive task battery, to identify significant pairwise differences in behavioral measures. Significant differences between the NT and Tg, NT and NT-KO, and Tg and Tg-KO groups are indicated, as well. Nontransgenic-control and Alzheimer's transgenic-control animals were found to differ in both sensorimotor (EP-OE) and cognitive measures (YM-PA, RAWM overall T4 and T5 errors and latencies, and RAWM final-block T5 errors). The primacy of RAWM measures underscores the importance of working memory assessment for identifying Alzheimer's-associated cognitive impairment, as well as the utility of the RAWM task for behavioral phenotyping in mice. Only two significant differences were found between the Tg and Tg-KO animals (both sensorimotor, OF and EP-OE), suggesting the removal of GRK5 expression does not further exacerbate cognitive impairment observed in Alzheimer's transgenic animals. Similarly, nontransgenic animals expressing GRK5 do not differ significantly from GRK5-knockout mice with respect to any cognitive measure, although these two groups differ in Open Field activity. Hence, on the basis of standard statistical methods, we may conclude that GRK5 does not significantly impact cognitive performance in mice, regardless of Alzheimer's transgenic background. However, no significant differences were found for any of the behavioral measures between the NT and Tg-KO groups of mice. Taken together with the NT vs. Tg contrast, this finding suggests that the absence of GRK5 expression may, in

fact, ameliorate cognitive impairment normally observed in Alzheimer's transgenic animals.

The correlation analysis results were consistent with prior studies (e.g., Leighty, et al., 2004; Caffeine Administration in Nontransgenic Mice study, this dissertation) involving behavioral assessment in mice using the comprehensive task battery, wherein widespread significant correlations are found among cognitive tasks (specifically, Platform Recognition, Morris water maze, and RAWM) and, to a lesser extent, between sensorimotor and cognitive tasks (e.g., Open Field activity and RAWM). Additionally, the multifactorial nature of certain tasks (e.g., Elevated Plus maze, Y-maze) was revealed through segregation of component measure correlations, such as between exploratory/activity elements of the Elevated Plus maze and measures of general locomotor function (EP-CE / OF) or systematic search capacity (EP-OE / YM-PA). The aggregation of cognitive-based performance measures across tasks was revealed through exploratory factor analysis, wherein the primary factor, representing general mnemonic function, was comprised of all RAWM measures, as well as Y-maze percent alternations, while a separate factor was comprised of both Platform Recognition and Morris water maze task measures. Segregation of task components, as well, was found through factor analysis (e.g., the three Elevated Plus measures load on separate factors).

Also utilizing data from the comprehensive task battery, discriminant analysis (both direct-entry and stepwise-forward) was used to construct linear classifiers for distinguishing individuals between groups, as well as among all four groups. Significant discriminability among the groups was obtained using stepwise-forward discriminant analysis, although the standard direct-entry approach was unsuccessful. These findings were consistent with earlier studies, wherein stepwise-forward analyses achieve superior discriminability relative to the standard direct-entry (complete)

approach (e.g., Arendash and King, 2002; Leighty et al., 2004). Optimal discriminability was observed between the nontransgenic animal groups on the basis of GRK5 expression, in which 87% of mice were correctly identified using only three of the behavioral measures (OF, EP-TO, and RME-FT5), representing a combination of sensorimotor, anxiety, and cognitive components of behavior. In addition, the stepwise-forward analysis demonstrated the highest overall discriminability among the four groups of all classifiers examined, correctly identifying 56% of individuals, yet requiring only four of the comprehensive task battery measures. The primacy of cognitive measures from the RAWM task, emphasized by both standard ANOVA and advanced statistical methods, further underscores the importance of working memory evaluation in behavioral phenotyping and classification. Taken together, these results demonstrate the remarkable capacity of stepwise-forward discriminant analysis for identifying groupwise differences, often using combinations of predictor variables which may not appear meaningful or coherent (to an experienced animal behaviorist, for example). Although an iterative search through a relatively large array of empirical measures may be acceptable for exploratory data analysis, this is not necessarily appropriate when prior experience (or other guidance) exists for identifying candidate predictor variables. Indeed, the informed selection of individual behavioral measures (e.g., only RAWM error-scores) followed by the application of standard statistics, including ANOVA, is a common protocol for neurobehavioral research, particularly when reliable measures for specific syndromes are known in advance (e.g., use of RAWM trials T4 and T5 to assess working memory impairment associated with Alzheimer's-like neuropathology).

The data mining-based classifiers exhibited diverse, idiosyncratic behavior across both groups and methodologies, in contrast to the relatively consistent performance of the stepwise-forward analyses. Indeed, the only consistent result across all data

mining techniques was the discriminability between the NT and Tg groups. Decision trees exhibited the poorest performance in terms of pairwise group discriminability, only distinguishing between two pairs of groups (NT vs. Tg, Tg vs. Tg-KO), while neural networks demonstrated the best overall performance among the data mining techniques (for both two-way and four-way classification). In fact, the neural network-based classifier outperformed the stepwise-forward analysis for distinguishing between NT and Tg animals (89% vs. 84%). Neural networks, however, performed only slightly better than random-chance level with respect to the Tg vs. Tg-KO and NT vs. Tg-KO comparisons. The support vector machine-based classifiers showed the poorest four-way discriminability of all classifiers examined (28%), relative to the 25% expected through random assignment of individuals to groups. Measures from the RAWM were preferentially selected by the decision trees, underscoring the diagnostic utility and information richness of this cognitive assessment task. Hence, despite the availability of diverse behavioral measures from the comprehensive task battery, data mining-based approaches were generally weaker than advanced statistical methods, with respect to groupwise discriminability involving two genetic manipulations (i.e., APP, GRK5) in mice.

In contrast to the comprehensive task battery, the mouse-based interference paradigm consists only of cognitive-based behavioral measures. Both error-scores and response-latencies are reported for all tasks (three-trial recall, proactive interference, retroactive interference, and delayed-recall). Standard statistical analyses (ANOVA) indicated significant ($p < .05$) differences between the Tg and NT groups in the three-trial recall, retroactive interference, and delayed-recall measures, with respect to both errors and latencies. All three of these measures reflect acquisition and retention of the initial learning task (Pool A), by contrast with the proactive interference task (Pool B) which represents a distinct, albeit cognitively-related

(spatial working memory), learning condition. This context-sensitivity effect was underscored by patterns of correlation and factor component loadings observed in subsequent analyses of the interference paradigm measures. Neither nontransgenics nor Alzheimer's transgenic mice could be distinguished on the basis of GRK5 expression (i.e., NT vs. NT-KO, Tg vs. Tg-KO), using any of the tasks/measures of the interference paradigm. Furthermore, there were no significant groupwise differences observed in either error-scores or response-latencies in the proactive interference task. Indeed, the comparatively higher scores in the proactive interference task (indicative of poorer task performance, relative to the other tasks) exhibited by all groups underscores the difficulty of switching between Pool A- and Pool B-learning experienced by all animals tested. Additionally, the Tg-KO mice only differed significantly from NT mice with respect to retroactive interference response-latency, which involves context-switching between the interference-condition (Pool B) and the original learning condition (Pool A). Hence, generally comparable cognitive performance was displayed by NT and Tg-KO animals, consistent with standard statistical analysis (ANOVA), but contrary to advanced statistical examination (discriminant function) of behavioral measures from the comprehensive task battery wherein a classifier model used sensorimotor and cognitive measures to reliably distinguish between these two groups.

The correlation analysis of behavioral measures from the interference paradigm revealed significant intercorrelation among both error and latency measures of the three-trial recall, retroactive interference, and delayed-recall tasks, reminiscent of correlation patterns observed both within and between cognitive-based tasks of the comprehensive task battery (e.g., Leighty et al., 2004). In addition, corresponding error and latency measures for each task were also strongly correlated, underscoring the comparability of these cognitive performance indices. Neither of the proactive

interference measures was significantly correlated with any other behavioral measure, which suggests independence between learning occurring within proactive interference and the other measures evaluated (from Pool A). Indeed, the two significant factors returned by exploratory factor analysis clearly segregated measures associated with Pool A (Factor I; three-trial recall, retroactive interference, and delayed-recall) from the measures associated with Pool B (Factor II; proactive interference).

The results of the discriminant analyses of measures from the cognitive interference paradigm were generally consistent with earlier studies, wherein stepwise-forward analyses achieve superior discriminability relative to the standard direct-entry (complete) approach (e.g., Arendash and King, 2002; Leighty et al., 2004). Indeed, the only successful "statistical" classifier utilized stepwise-forward discriminant analysis (for both Tg vs. NT, and NT vs. Tg-KO, discriminability using error-scores and/or response-latency data). Alzheimer's transgenic animals were not distinguishable with respect to GRK5 expression, nor were NT mice significantly distinguishable from NT-KO animals (consistent with Suo et al., 2007). Moreover, comparatively superior discriminability between the NT and Tg-KO groups was demonstrated using the two interference-related measures (proactive and retroactive), while the NT vs. Tg contrast emphasized the delayed-recall error score. Hence, the differences in cognitive performance which distinguish NT and Tg-KO individuals differ from those which distinguish NT and Tg individuals. Indeed, these results would be predictable from the performance of mice, as depicted in the standard bar-graph representation (Figure 7.1).

As portrayed in Table 7.7 for the interference task, both the neural network- and support vector machine-based classifiers were inferior to advanced statistical (stepwise-forward discriminant analysis) methods and, moreover, exhibited similarly idiosyncratic performance to the analyses reported earlier for the comprehensive task

battery. The decision tree, however, was the only classifier examined which was capable of distinguishing between Tg and Tg-KO mice (using error-score data), utilizing the delayed-recall measure. In addition, the decision tree was comparable to stepwise-forward discriminant analysis for distinguishing between NT and Tg-KO animals, using response-latency data (with or without error-score data). The retroactive interference latency was shown to be the best predictor variable for group membership, consistent with the findings for NT vs. Tg-KO contrast using stepwise-forward discriminant analysis discussed earlier. The performance profiles of the neural networks and support vector machines are unsatisfactory, relative to the decision trees, and interpretation is further complicated by marginal performance (e.g., often only slightly above random-chance assignment levels), as well as ubiquitous nonsignificant sensitivity and/or specificity.

The goals of this study were to examine the effectiveness of a novel interference-based paradigm for cognitive assessment in mice, through comparison with an established comprehensive behavioral task battery, and to utilize advanced statistical and computational (data mining-based) analytic techniques along with conventional (ANOVA-based statistics) approaches for neurobehavioral research. Two genotypic components (Alzheimer's transgenicity and GRK5) were manipulated, and the consequential behavioral (cognitive) manifestations were evaluated and analyzed. A significant main effect for the Alzheimer's transgenic genotype (APPsw) was identified using the comprehensive task battery, through standard statistical techniques for individual cognitive measures, as well as by advanced statistical and data mining-based classifiers, consistent with prior findings (e.g., Leighty et al., 2004; Leighty et al., 2008). Similarly, behavioral measures from the interference paradigm successfully distinguished between NT and Tg mice using standard statistics, advanced statistics and data mining techniques.

The effect of GRK5 expression against a nontransgenic (NT) background was revealed with the comprehensive task battery, wherein a single sensorimotor measure (Open Field activity) was identified by standard statistics, and both sensorimotor and cognitive measures were returned by stepwise-forward discriminant analysis. By contrast, none of the classifiers examined were able to distinguish between NT and NT-KO individuals using behavioral measures from the interference paradigm. The influence of the GRK5 genotype against an Alzheimer's transgenic background (APPsw) was indicated by two sensorimotor measures (Open Field activity, Elevated Plus maze open-arm entries) through standard statistics, as well as by stepwise-forward discriminant analysis and decision trees using behavioral data from the comprehensive task battery. Using measures from the interference paradigm, the Tg and Tg-KO groups were largely indistinguishable through standard statistics and decision tree-based classifiers (except using error-score data) only. Additionally, the NT and Tg-KO groups were distinguishable using standard statistics (only for retroactive interference response-latency in the interference paradigm), as well as by stepwise-forward discriminant analyses of both the comprehensive task battery (requiring sensorimotor, anxiety, and cognitive-based measures) and interference paradigm measures.

This study also demonstrated that the effectiveness of group classification was dependent upon whether the comprehensive task battery or the interference paradigm was used as the source for the behavioral measures. By comparing Tables 7.4 and 7.7, for example, we find that behavioral measures from the comprehensive task battery provided superior discriminability by all classifiers (both advanced statistical and data mining-based), relative to interference paradigm-based measures, for distinguishing between nontransgenic-control and Alzheimer's transgenic-control animals, as well as among the four groups of mice. This may be attributable to the broader sampling from the behavioral repertoire (sensorimotor, anxiety, and cognitive com-

ponents) offered by the comprehensive battery, relative to the cognitive-based interference paradigm. As reported earlier, successful groupwise discriminability is often based upon a combination of sensorimotor-based, anxiety-based, and/or cognitive-based measures. Indeed, discriminant analysis, neural networks, and support vector machines were able to distinguish between the NT and NT-KO groups, as well as between the Tg and Tg-KO groups, using behavioral measures from the comprehensive task battery; however, none of these three types of classifiers succeeded using interference paradigm measures instead. Here again, the availability of a large number of candidate predictor variables increased the likelihood of distinguishing between groups. By contrast, interference paradigm measures (specifically, response-latencies) provided superior discriminability between the NT and Tg-KO groups, for all classifiers except support vector machines (which showed comparable performance). The choice of behavioral metrics (errors vs. latencies) may also determine the optimal classifier model. For detecting the effect of GRK5 expression on an Alzheimer's transgenic background (i.e., Tg vs. Tg-KO), for instance, a decision tree provided with behavioral measures from the comprehensive task battery will outperform a decision tree using only interference paradigm-based response-latency measures, but underperform a decision tree using only error-score interference measures. Finally, neither decision tree-based classifier (using measures from the comprehensive battery or the interference paradigm) reliably distinguished between nontransgenic mice on the basis of GRK5 expression (i.e., NT vs. NT-KO).

These results have several important implications for transgenic mouse models of Alzheimer's disease, as well as behavioral evaluation and analytic methodology:

First, the APP genotype undoubtedly exerts a very powerful influence on the cognitive phenotype of mice within the mixed background of our colony. Indeed, as previously discussed, even a single behavioral measure (e.g., RAWM T4 errors, delayed-

recall errors) may be sufficient to distinguish between nontransgenic and Alzheimer's transgenic animals.

Second, that classifiers exhibit differential sensitivity to behavioral features. In some cases, error-scores are better discriminators (e.g., Tg vs. NT using neural networks), while in other cases, response-latencies are more effective (e.g., Tg vs. NT using stepwise-forward discriminant analysis). For this reason, it is extremely important for practitioners to examine and compare alternative classifier designs, to evaluate applicability for a specific research problem. Moreover, the importance of diverse sampling across the behavioral repertoire (sensorimotor-based, anxiety-based, and cognitive-based tasks/measures) cannot be overstated, to maximize the chances of detecting groupwise differences otherwise overlooked.

Third, that additional response data may not necessarily improve groupwise discriminability, nor interpretability of analytic results. For behavioral measures of the interference paradigm, the availability of both error-scores and response-latencies did not necessarily improve the performance of trained classifiers, compared to using either of the two response measures alone. In addition, decision tree performance declines, support vector machines improve slightly, and stepwise-forward discriminant analyses often remain the same when latency measures are included, relative to using error-scores alone. And,

Finally, although the comprehensive task battery provides a richer sampling of the mouse behavioral repertoire, thus generally improving diagnostic effectiveness (e.g., distinguishing between treatment groups), the mouse-based interference paradigm developed and refined in this Study represents a powerful new instrument for cognitive assessment in Alzheimer's disease research with transgenic animals.

CHAPTER 8

INTERFERENCE TASK-BASED THERAPEUTIC EVALUATION OF GM-CSF

8.1 Introduction

Granulocyte macrophage colony-stimulating factor (GM-CSF) is an inflammatogenic cytokine component of the immune-inflammatory cascade, which promotes production of both granulocytes (eosinophils, basophils, and neutrophils) and monocytes (which, in turn, mature into GM-CSF-secreting macrophages) by stem cells. In humans, GM-CSF is used to stimulate white blood cell production following chemotherapy. However, when administered to Alzheimer's transgenic mice in conjunction with IL-4, as an adjuvant in beta-amyloid immunotherapy, fewer cortical Congophilic plaques and reduced plaque-associated microgliosis were reported (DaSilva et al., 2006). Coadministration of GM-CSF and IL-r promotes an attenuated Th2 response to beta-amyloid immunization, including antibodies which, in turn, reduce plaque burden (DaSilva et al., 2006). GM-CSF levels, measured in brain slice cultures, are positively correlated with beta-amyloid (both Aβ_{40} and Aβ_{42}) in Alzheimer's transgenic mice (Patel et al., 2005), underscoring the role of cytokines in the neuroinflammatory response to beta-amyloid. Successful attenuation of plaque-associated structural and functional impairment may constitute a therapeutic role for GM-CSF, albeit indirect, to preserve or restore cognitive function in AD.

The purpose of this study was to investigate the cognitive effects of GM-CSF administration in both nontransgenic and Alzheimer's transgenic mice using a novel,

mouse-based interference learning paradigm. This behavioral paradigm was adapted from a semantic interference testing protocol (Loewenstein et al., 2004) for AD diagnosis and screening in clinical settings, by substituting rodent-appropriate test conditions (RAWM apparatus, visual cues) and spatial memory domain in place of the original protocol's verbal stimuli and semantic components, as well as including both error-score and response-latency measures from each behavioral task. Details of the behavioral tasks in the original protocol were presented earlier (refer to the "Interference Testing in Humans: A Comparison of Statistical and Data Mining Methods" Study, Chapter 6). The mouse-based behavioral paradigm consists of four interrelated tasks, as described in the previous chapter ("The Interference Task: A Novel Assessment Paradigm", Chapter 7): Three-trial recall, proactive interference, retroactive interference, and delayed-recall. Behavioral measures (error-scores and/or response-latencies) from one or more of these tasks have successfully distinguished between nontransgenic and Alzheimer's transgenic mice, as well as between Alzheimer's transgenic animals with an additional genetic modification (GRK5-knockout) and transgenics having the GRK5-wildtype genotype (Chapter 7). The applicability of the interference paradigm for evaluating GM-CSF therapeutic benefits was examined using both advanced statistical (discriminant analyses) and data mining-based methods (decision trees, neural networks, support vector machines), and discussed in light of standard statistical (analysis of variance; ANOVA) findings by Arendash et al. (unpublished data).

8.2 Materials and Methods

The subjects consisted of age-matched mice generated from a cross between heterozygous APPsw (APP mutation K670N and M671L) and heterozygous PS1 (Tg line 6.2) mice, from which 16 Alzheimer's transgenic (APPsw) and 16 nontransgenic

(NT) mice were selected. At one-year of age, these 32 animals were cognitively pre-tested for eight days using the RAWM working memory task to assess cognitive performance. The APPsw mice were then separated into two groups by random assignment, balanced for cognitive performance and blood beta-amyloid levels: Alzheimer's transgenic-control (Tg; N=8) and transgenic-treatment (Tg+GMCSF; N=8). Similarly, the NT mice were randomly assigned to either of two groups, balanced for cognitive performance: Nontransgenic-control (NT; N=8) and nontransgenic-treatment (NT+GMCSF; N=8). After two weeks, the mice began receiving daily subcutaneous injections of either GMCSF (5 mcg) (NT+GMCSF and Tg+GMCSF groups) or plain vehicle (NT and Tg groups) for ten days. All animals were then cognitively post-tested for four days using the RAWM, followed by two days without testing, and concluded by the interference paradigm for four days. Daily injections were continued throughout the cognitive evaluation period.

Behavioral measures of all animals from the final two-day block of the four-day RAWM post-test (i.e., both errors and latencies for working memory, trials T4 and T5), as well as from the final two-day block of the four-day interference paradigm (i.e., both errors and latencies for three-trial recall, proactive interference, retroactive interference, and delayed-recall), were compiled for subsequent analyses. Standard statistical methods (ANOVA, F-test) were used to identify group differences in the final-block post-test RAWM dataset, as well as the final-block interference dataset. In addition, performance measures from the final-block of the interference paradigm were used to generate additional datasets for further analysis, as described in the Mouse Interference Paradigm (in Section 7.2) of the "The Interference Task: A Novel Assessment Paradigm" Study (Chapter 7). The eight cognitive measures, consisting of mean error-score and response-latency from each of the four tasks of the interference paradigm, were grouped into three separate datasets (errors-only, latencies-only,

both errors and latencies) for analysis using both advanced statistical (correlation analysis, factor analysis, and discriminant analysis) and data mining-based methods (decision trees, neural networks, and support vector machines). Complete specifications of the computing hardware and software utilized in the analyses, including program parameter settings, are provided in the General Analytic Protocol of Section 4.2 in the "Caffeine Administration in Nontransgenic Mice" Study (Chapter 4).

Animal care and use was in accordance with the Guide and Use of Laboratory Animals, National Research Council, 1996, in a program and facilities fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, International, under a protocol approved by the University of South Florida Institutional Animal Care and Use Committee (No. 3183, Huntington Potter, Ph.D., Principal Investigator).

8.3 Results of Statistical Analyses

8.3.1 Standard Statistical Analysis

Figure 8.1 (next page) shows the cognitive performance (error-scores) of all groups in the post-test RAWM working memory trials T4 and T5 for both two-day blocks. Groupwise means and standard-error are indicated in the figure. The double-asterisk ($**$) denotes significant difference between the Tg group and the other three groups of animals in the RAWM working memory trial T5 measure ($p < .05$). The dagger symbol (†) indicates significant contrast ($p < .05$) with NT+GMCSF and Tg+GMCSF. Measured ten days into the GMCSF treatment period, Tg-control mice exhibited impairment in RAWM working memory trials T4 and T5, compared with NT animals. Cognitive impairment of Tg-control mice was evident within two-day blocks (Figure 8.1, upper), as well as across all four days of RAWM testing (Figure 8.1, lower). Alzheimer's transgenic animals receiving GM-CSF (Tg+GMCSF), by contrast, performed comparably (or, indeed, superior) to NT on RAWM working memory trials T4 and T5, both within blocks and overall. Hence, GM-CSF was shown to reverse working memory impairment (as evaluated by the RAWM task) in Alzheimer's transgenic mice.

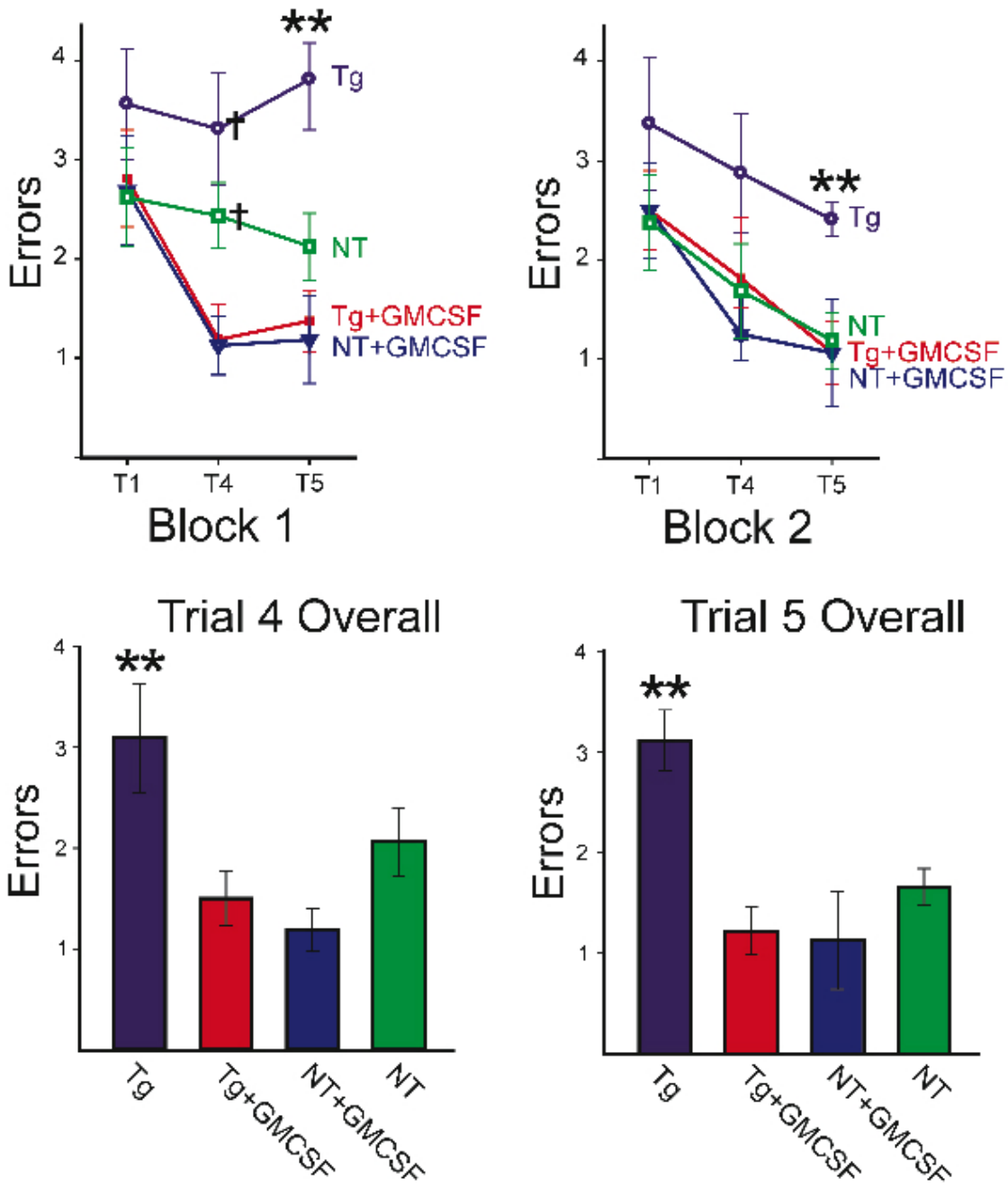

Figure 8.1. Groupwise comparison of error-scores in RAWM post-test trials T4 and T5, by blocks, in the GMCSF study

Figure 8.2 (next page) portrays groupwise cognitive performance on all measures of the final-block of the interference paradigm, indicating the mean and standard error associated with each measure for all groups. Standard statistical tests (ANOVA, F-test) reveal a significant difference ($p < .05$) between the Tg-control group and the other three groups, denoted by a double-asterisk ($**$), as well as a significant difference ($p < .05$) between the Tg-control and NT+GMCSF groups (indicated by a single asterisk, $*$). With respect to both error-score and response-latency measures, Tg-control mice showed significant impairment in both the three-trial recall and delayed-recall tasks of the interference paradigm, relative to GMCSF-treated Alzheimer's transgenic mice (Tg+GMCSF) and both nontransgenic groups (NT, NT+GMCSF). Indeed, GMCSF treatment in Tg mice dramatically improved (normalized) their cognitive performance in both the three-trial recall and delayed-recall tasks to levels comparable to NT animals. Additionally, the Tg-control group was significantly impaired compared with the NT+GMCSF group, with respect to retroactive interference (both errors and latencies). Finally, GMCSF-treated nontransgenic animals (NT+GMCSF) performed better (i.e., both decreased errors and latencies) in the proactive interference and retroactive interference tasks, relative to untreated nontransgenics (NT), although these differences did not meet the criterion for significance (both p's between .05 and .20).

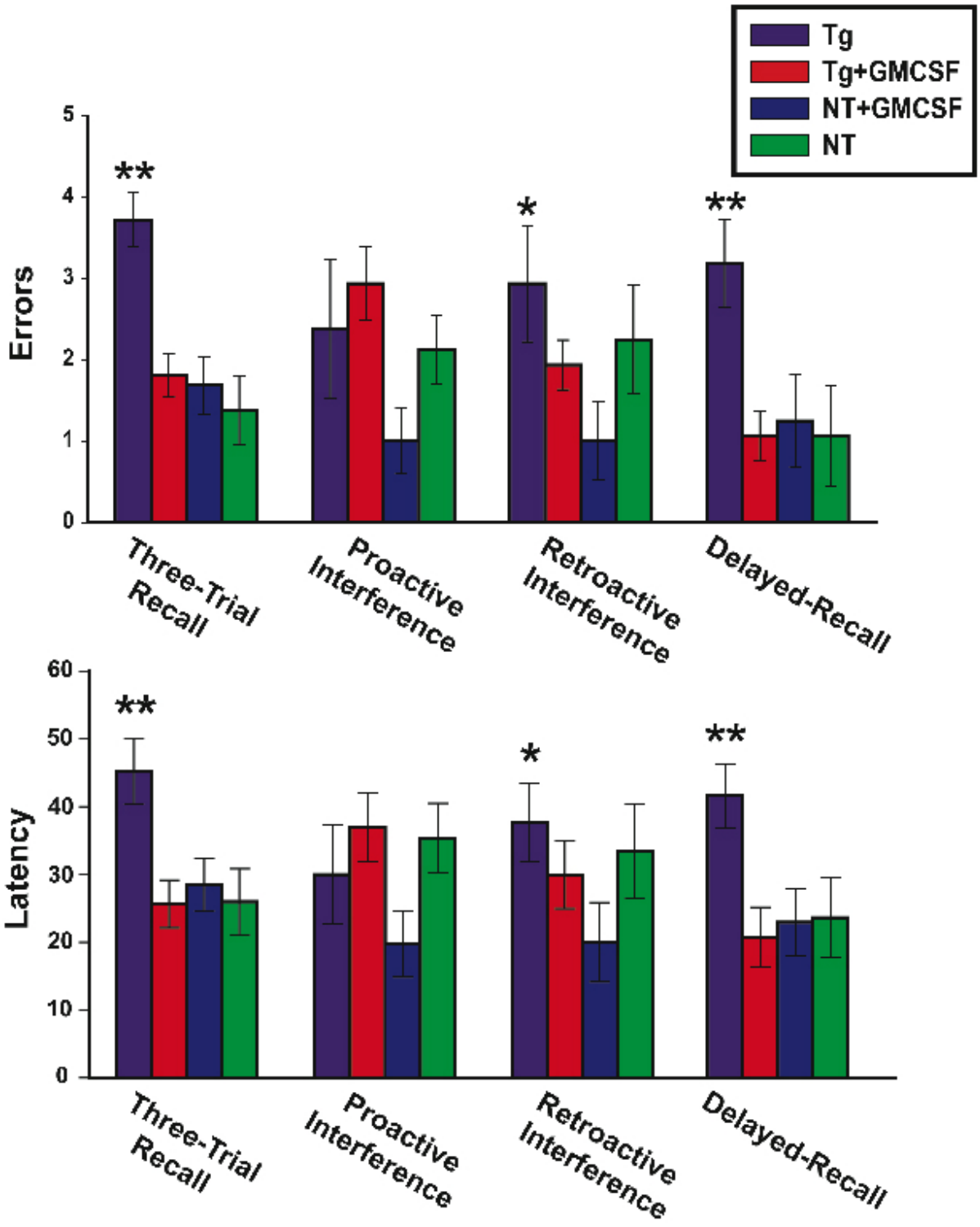

Figure 8.2. Groupwise contrasts for all interference paradigm measures in the GMCSF study

8.3.2 Correlation Analysis

Table 8.1. Correlations between behavioral measures of the interference paradigm in the GMCSF study

TT-L	.93						
	.000						
PI-E							
PI-L			.93				
			.000				
RI-E	.62	.56	.73	.69			
	.000	.001	.000	.000			
RI-L	.55	.58	.67	.74	.92		
	.001	.000	.000	.000	.000		
DR-E	.77	.64	.47	.36	.74	.64	
	.000	.000	.007	.044	.000	.000	
DR-L	.79	.77	.48	.44	.74	.74	.93
	.000	.000	.005	.011	.000	.000	.000
	TT-E	TT-L	PI-E	PI-L	RI-E	RI-L	DR-E

Table 8.1 shows significant ($p < .05$) pairwise correlations observed between behavioral measures for all groups. The correlation coefficient (r-value) and corresponding significance (p-value) are indicated at the top and bottom, respectively, within marked cells of the table. The following abbreviations are used to represent the four tasks of the mouse-based interference paradigm: Three-trial recall, errors and latency (TT-E, TT-L); proactive interference, errors and latency (PI-E, PI-L); retroactive interference, errors and latency (RI-E, RI-L); and, delayed-recall, errors and latency (DR-E, DR-L).

Significant positive correlations were observed between the error-score measures of the three-trial recall and both retroactive interference and delayed-recall tasks. A similar pattern of association was found among the corresponding latency measures of these three tasks. These findings indicate close correspondence between tasks involving Pool A of the behavioral paradigm. The error-score and response-latency

measures of each of the four tasks were significantly correlated, as well, underscoring the comparability of both indices of cognitive performance for the tasks comprising the mouse-based interference paradigm. Additionally, both measures from the three-trial recall task were significantly and positively intercorrelated with both measures of retroactive interference, suggesting successful acquisition and retention of the initial learning component (Pool A), despite the brief intervening distractor task (Y-maze exposure) and interference task (Pool B exposure). Similarly, complete intercorrelation between all measures of the retroactive interference and delayed-recall tasks suggests long-term stability of learning. This interpretation is supported by significant pairwise intercorrelation among all measures of the three-trial recall and delayed-recall tasks. In sharp contrast, both behavioral measures from the proactive interference task did not correlate with both measures from the initial three-trial recall. This indicates independent performance between Pool A (the learning condition for three-trial recall) and Pool B (the learning condition for proactive interference). However, both behavioral measures from proactive interference were significantly correlated with all measures from both the retroactive interference and delayed-recall tasks. For animals exhibiting low errors/latencies across all three of these tasks, this would indicate flexibility of learning across different problem contexts (Pool A vs. Pool B) – that is, the animals were able to recall the original learning (Pool A), despite interference/distractor elements and a time delay, as well as to rapidly switch between the two learning conditions.

The correlation matrix obtained in the present study differs from that obtained in the previous study (Chapter #7), in that significant pairwise correlations were found in the present study between measures from the Proactive Interference task and measures from both the Retroactive Interference and Delayed-Recall Tasks. This impor-

tant difference has implications for the resulting factor structure, to be discussed below.

8.3.3 Factor Analysis

Table 8.2. Varimax-rotated factor analysis of interference paradigm measures in the GMCSF study

	Factor	
Measure	I	II
Three-Trial Recall (Errors)	0.941	
Three-Trial Recall (Latency)	0.897	
Delayed-Recall (Latency)	0.870	
Delayed-Recall (Errors)	0.837	
Retroactive Interference (Errors)		0.736
Retroactive Interference (Latency)		0.749
Proactive Interference (Latency)		0.959
Proactive Interference (Errors)		0.932
Variance	47.51%	39.70%

Table 8.2 depicts a varimax-rotated principal component analysis of the eight behavioral measures of the mouse-based interference task. Significant (absolute value greater than 0.700) component loadings are shown for the two factors returned. The calculated eigenvalue results (i.e., >1 criterion) were consistent with scree plot (Cattell, 1966) identification of two significant factors. Measures from the three-trial recall and delayed-recall tasks comprise the primary factor, accounting for approximately 48% of overall variance. These two tasks involve spatial learning within the initial context (Pool A) and long-term (reference) memory for this same information. Hence, this factor may represent successful (or unsuccessful) acquisition and consolidation of the initial spatial learning problem (Pool A). By contrast, the second factor (representing approximately 40% of variance) consists of all measures from the interference components (both proactive and retroactive) of the paradigm. This second

factor may reflect the subjects' facility of context-switching (Pool A vs. Pool B) or the ability to rapidly adapt to changing learning conditions.

As noted previously, differences in correlation patterns among behavioral measures, observed between this study and the preceding one (Chapter #7), are reflected in the corresponding factor structures. In contrast to the preceding study, Factor II of the present study also includes both behavioral measures from the Retroactive Interference task, which is consistent with the observed pattern of correlation (Table 8.1). Interestingly, while both measures from the Delayed-Recall task loaded on a separate factor from both measures of the Proactive Interference task, in both studies, significant pairwise correlations between measures of the Proactive Interference and Delayed-Recall tasks were only found in the present study. The cause of this discrepancy is unclear, although treatment-bias effects and/or disproportionate sample sizes (unlike the preceding study, all treatment groups of the present study consisted of identical numbers of mice) are potential sources.

Table 8.3. Classifier performance comparison using interference paradigm measures in the GMCSF study

Groups	Dataset	Evaluation Criterion	Discriminant Analysis		Decision Tree	Neural Network	SVM
			Complete	Step-Fwd			
Tg vs. NT	Errors	Accuracy	88%	94%	88%	75%	88%
		Sensitivity	88%	100%	100%	75%	88%
		Specificity	88%	88%	75%	75%	88%
	Latency	Accuracy	81%	81%	56%	81%	81%
		Sensitivity	88%	75%	NS	75%	75%
		Specificity	75%	88%	75%	88%	88%
	Err+Lat	Accuracy	88%	94%	88%	69%	75%
		Sensitivity	100%	100%	100%	63%	63%
		Specificity	75%	88%	75%	75%	88%
NT vs. NT+ GMCSF	Errors	Accuracy	NS	NS	56%	63%	63%
		Sensitivity	NS	NS	63%	63%	63%
		Specificity	NS	NS	NS	63%	63%
	Latency	Accuracy	NS	69%	NS	56%	69%
		Sensitivity	NS	63%	NS	63%	63%
		Specificity	NS	75%	NS	NS	75%
	Err+Lat	Accuracy	NS	69%	NS	56%	56%
		Sensitivity	NS	63%	NS	NS	63%
		Specificity	NS	75%	NS	63%	NS
Tg vs. Tg+ GMCSF	Errors	Accuracy	81%	94%	81%	100%	75%
		Sensitivity	100%	100%	88%	100%	100%
		Specificity	63%	88%	75%	100%	NS
	Latency	Accuracy	94%	94%	NS	88%	75%
		Sensitivity	88%	88%	NS	88%	75%
		Specificity	100%	100%	NS	88%	75%
	Err+Lat	Accuracy	88%	94%	81%	94%	82%
		Sensitivity	88%	88%	88%	100%	100%
		Specificity	88%	100%	75%	88%	63%
All four	Errors	Accuracy	47%	53%	50%	48%	NS
	Latency	Accuracy	41%	38%	NS	NS	28%
	Err+Lat	Accuracy	47%	59%	41%	41%	NS

8.3.4 Discriminant Analysis

A summary comparison of groupwise discriminability is shown in Table 8.3, for reference. The table reports the classifier performance evaluation (i.e., accuracy, sensitivity, and specificity) for advanced statistical techniques (direct-entry and stepwise-forward discriminant analyses) in the fourth and fifth columns.

Nontransgenic (NT) vs. Transgenic (Tg) Groups

Error-Scores Only. As reported in Table 8.3, direct-entry (complete) discriminant analysis returned excellent discriminability (Wilks' lambda = .261, p = .0031) between transgenic- and nontransgenic-control animals (accuracy = 88%, sensitivity = 88%, specificity = 88%) on the basis of the four predictor variables. The stepwise-forward approach generated a more parsimonious model, consisting of only two predictor variables (three-trial recall and retroactive interference), with significant discriminability (Wilks' lambda = .271, p = .0002). The simpler (two-variable) model demonstrated outstanding (94%) accuracy, 100% sensitivity, and 88% specificity for transgenicity (i.e., all transgenic animals were correctly identified).

Response-Latencies Only. Using only latency data from the four tasks, direct-entry discriminant analysis successfully distinguished between the two groups (Wilks's lambda = .383, p = .0225), with 81% accuracy (sensitivity = 88%, specificity = 75%), as shown in Table 8.3. The stepwise-forward approach also demonstrated significant discriminability (Wilks' lambda = .452, p = ..0058) between transgenic- and nontransgenic-control animals (accuracy = 81%, sensitivity = 75%, specificity = 88%), although the overall accuracy was comparable to the direct-entry classifier result. The stepwise-forward model retained only the three-trial recall and proactive interference latency measures.

Errors and Latencies. A direct-entry discriminant analysis showed significant discriminability between groups using both error-scores and reponse-latencies (Wilks' lambda = .079, p = .0031) between transgenic- and nontransgenic-control animals (accuracy = 88%, sensitivity = 100%, specificity = 75%) using all eight measures as predictor variables. By contrast, the stepwise-forward approach retained only three predictor variables (three-trial recall errors, retroactive interference errors, and three-trial recall latency), and displayed significant discriminability (Wilks' lambda = .143, p = .0000) between the two groups, as reported in Table 8.3. The simpler (three-variable) model demonstrated superior (94%) accuracy, 100% sensitivity, and 88% specificity for transgenicity.

Summary: Nontransgenic and Alzheimer's transgenic animals were reliably distinguished using either/both error-scores and response-latencies obtained from behavioral tasks in the interference paradigm. The error measures from the three-trial recall and retroactive interference tasks were emphasized by the discriminant analyses, regardless of the availability of response-latency metrics.

Nontransgenic (NT) vs. Nontransgenic-Treatment (NT+GMCSF) Groups

Error-Scores Only. Neither direct-entry (complete) nor stepwise-forward discriminant analysis successfully distinguished between GMCSF-treated and untreated nontransgenic mice using only error-score measures from the interference paradigm. None of the candidate predictor variables met the statistical criterion (alpha-to-enter) for inclusion in the model.

Response-Latencies Only. Although significant discriminability was not demonstrated using direct-entry discriminant analysis with response-latency data, as indicated in Table 8.3, the stepwise-forward approach returned significant discriminability (Wilks' lambda = .622, p = .0457) between the two groups (accuracy = 69%, sensitivity = 63%, specificity = 75%), with respect to GMCSF treatment. The

stepwise-forward model retained both the three-trial recall and proactive interference measures.

Errors and Latencies. Direct-entry discriminant analysis was unable to distinguish between the two groups using both error-scores and reponse-latencies. However, the stepwise-forward approach returned significant discriminability (Wilks' lambda = .622, p = .0457) between the two groups using a two-variable model (retaining the latency measures from the three-trial recall and proactive interference tasks), as shown in Table 8.3. The two-variable model showed 69% accuracy, 63% sensitivity, and 75% specificity, with respect to treatment.

Summary: The difficulty in distinguishing between untreated and GMCSF-treated nontransgenic animals was underscored by the inability of direct-entry discriminant analysis to identify individuals by group using behavioral error and/or latency measures. Although stepwise-forward discriminant analyses consistently identified the three-trial recall and proactive interference error-score measures as the best predictor variables for distinguishing between GMCSF-treated and untreated nontransgenic mice, the relatively poor discriminability between these two groups (using either/both errors and latencies) suggests that GMCSF treatment does not affect normal mice.

Transgenic (Tg) vs. Transgenic-Treatment (Tg+GMCSF) Groups

Error-Scores Only. Direct-entry (complete) discriminant analysis successfully distinguished (Wilks' lambda = .276, p = .0042) between the two groups (accuracy = 81%, sensitivity = 100%, specificity = 63%), as indicated in Table 8.3. The model generated by the stepwise-forward approach included the three-trial recall and proactive interference error-scores as predictor variables. This two-variable model also showed significant discriminability between the two groups (Wilks' lambda =

.335, p = .0008), with high (94%) accuracy, 100% sensitivity, and 88% specificity, with respect to treatment.

Response-Latencies Only. Significant discriminability was demonstrated using direct-entry discriminant analysis with response-latency data (Wilks' lambda = .236, p = .0018), as reported in Table 8.3. The resulting classifier achieved 94% overall accuracy, 88% sensitivity, and 100% specificity, with respect to treatment. The stepwise-forward approach also returned significant discriminability (Wilks' lambda = .273, p = .0002), and used only the proactive interference and delayed-recall latency measures as predictor variables. This parsimonious model demonstrated 94% accuracy, 88% sensitivity, and 100% specificity, with respect to treatment. All transgenic-control animals were correctly identified by both classifiers.

Errors and Latencies. Direct-entry discriminant analysis using both error-scores and response-latencies successfully distinguished between the two groups (Wilks' lambda = .066, p = .0017) between the two groups (accuracy = 88%, sensitivity = 88%, specificity = 88%), as depicted in Table 8.3. Three variables were retained by the stepwise-forward approach (proactive interference errors and latency, delayed-recall latency). The three-variable model also distinguished between GMCSF-treated and untreated Alzheimer's transgenic mice (Wilks' lambda = .191, p = .0001), with excellent (94%) accuracy, 88% sensitivity, and 100% specificity, with respect to treatment.

Summary: Excellent (94%) discriminability was achieved using either/both error-scores and response-latencies, consistent with significant differences in cognitive performance between the two groups of mice. Error-score data alone, however, was more effective for correctly detecting the GMCSF treatment effect in Alzheimer's transgenic mice (i.e., greater sensitivity).

All Four Groups

Error-Scores Only. Significant discriminability among the four groups was returned by direct-entry (complete) discriminant analysis (Wilks' lambda = .274, p = .0005), with 47% overall accuracy, as indicated in Table 8.3. The GMCSF-treated nontransgenic animals were the most-frequently misclassified individuals, with only 25% correctly identified (random-chance level performance). By contrast, the stepwise-forward approach demonstrated higher discriminability (Wilks' lambda = .383, p = .0002) and accuracy (53%) using only two error measures: Three-trial recall and proactive interference.

Response-Latencies Only. Direct-entry discriminant analysis successfully distinguished among the four groups using only reponse-latency data (Wilks's lambda = .391, p = .0143), and exhibited 41% overall accuracy. Interestingly, although the stepwise-forward approach also returned significant discriminability among the groups (Wilks's lambda = .473, p = .0019), its overall accuracy was only 38%. The stepwise-forward model selected both the three-trial recall and proactive interference latency measures. The GMCSF-treated Alzheimer's transgenic mice were the most likely to be misclassified; only 25% of these animals were correctly identified, as if "normalized" by contrasting with the NT and NT+GMCSF animals.

Errors and Latencies. A direct-entry discriminant analysis returned significant discriminability (Wilks's lambda = .119, p = .0007) among the four groups (accuracy = 47%), as shown in Table 8.3, using the eight behavioral measures as predictor variables. The stepwise-forward approach retained only three measures (three-trial recall errors and latency, proactive interference latency), and showed significant discriminability (Wilks's lambda = .287, p = .0001) and relatively high (59%) overall accuracy among the groups. Only 30% of the GMCSF-treated nontransgenic mice were correctly identified; these animals were the most likely to be misclassified, reflec-

tive of "normalized" performance (the NT, NT+GMCSF, and Tg+GMCSF animals performed comparably).

Summary: Variance-optimized (stepwise-forward) discriminant analysis-based classifiers using both error-score and response-latency measures exhibit better groupwise discriminability than classifiers based on either errors or latencies alone. Moreover, the relatively poor overall accuracy may be attributed to the difficulty in distinguishing among NT, NT+GMCSF, and Tg+GMCSF groups, due to similarities in cognitive performance.

8.4 Results of Data Mining Analyses

Table 8.3 compares groupwise discriminability for each classifier in the present study. Performance evaluation is reported (i.e., accuracy, sensitivity, and specificity) when significant (exceeding random-chance level assignment) for data mining-based methods (decision trees, neural networks, and support vector machines) in the sixth through eighth columns.

8.4.1 Decision Tree Analysis

Nontransgenic (NT) vs. Transgenic (Tg) Groups

Error-Scores Only. When provided with only error-score behavioral data from all tasks of the interference paradigm, the decision tree-based classifier selected a single attribute, the three-trial recall measure, as the criterion for splitting the dataset into two groups. As shown in Table 8.3, the resulting tree correctly identified 88% (Kappa = 0.75) of individuals, with 100% sensitivity (i.e., all Tg mice were correctly identified) and 75% specificity (i.e., all but two NT animals were correctly identified).

Response-Latencies Only. The decision tree identified three behavioral latency measures – delayed-recall, proactive interference, and three-trial recall – for splitting

the dataset into two groups. The classifier displayed a very modest 56% accuracy (Kappa = 0.13), with 38% sensitivity (not significant) and 75% specificity. Hence, this classifier was more appropriate for correctly identifying nontransgenic individuals.

Errors and Latencies. When both error-scores and response-latencies were used to construct the decision tree-based classifier, only the three-trial recall error measure was selected as having sufficient information-bias to split the dataset into the two groups. The classifier correctly identified 88% of individuals (Kappa = 0.75), with 100% sensitivity and 75% specificity, as indicated in Table 8.3. Thus, the inclusion of latency measures with error measures did not change the overall classification accuracy.

Summary: Error-scores, either alone or together with response-latency measures, provide excellent discriminability between nontransgenic and Alzheimer's transgenic mice, using the interference paradigm. This result underscores the cognitive deficits present in Tg mice, relative to nontransgenics, based on conventional protocols for evaluation (e.g., RAWM).

Nontransgenic (NT) vs. Nontransgenic-Treatment (NT+GMCSF) Groups

Error-Scores Only. The decision tree selected a single behavioral measure, retroactive interference errors, as the best attribute for splitting the dataset into two groups, using only behavioral error-score data. As shown in Table 8.3, the classifier demonstrated 56% accuracy (Kappa = 0.13), 63% sensitivity, and 50% specificity (non-significant), with respect to treatment. One-half of the untreated nontransgenic animals were misclassified as "GMCSF-treated."

Response-Latencies Only. None of the response-latency measures provided sufficient information for constructing a decision tree-based classifier to distinguish between GMCSF-treated and untreated nontransgenic mice.

Errors and Latencies. When both error-scores and response-latencies were provided, the decision tree was unable to identify any attributes which reliably distinguish between the two groups.

Summary: Although the retroactive interference error-score measure provided modest discriminability between these two groups, these results generally suggest that GMCSF-treated and untreated nontransgenic mice do not differ significantly, with respect to performance in the interference paradigm.

Transgenic (Tg) vs. Transgenic-Treatment (Tg+GMCSF) Groups

Error-Scores Only. The decision tree identified a single attribute, three-trial recall errors, for splitting the dataset into two groups, corresponding to GMCSF-treated and untreated-control Alzheimer's transgenic mice. As reported in Table 8.3, the classifier accurately identified 81% of the individuals (Kappa = 0.63), with 88% sensitivity and 75% specificity.

Response-Latencies Only. None of the response-latency measures provided sufficient information-bias to accurately distinguish between the two groups.

Errors and Latencies. A single attribute, three-trial recall error, was identified as providing sufficient information to distinguish between the two groups, when all behavioral measures (both errors and latencies) were available to the classifier. The resulting tree exhibited 81% accuracy (Kappa = 0.63), with 88% sensitivity and 75% specificity. This is identical to the result obtained for Error-Scores Only.

Summary: Error-score data, either alone or together with response-latency measures, accurately (81%) distinguishes between Alzheimer's transgenic mice which received GM-CSF treatment and Tg animals which did not. This result suggests a GM-CSF treatment effect, with respect to cognitive performance in the interference paradigm, and a primacy of error-based performance criteria over latency-based measures for detecting GM-CSF treatment efficacy.

All Four Groups

Error-Scores Only. The error-scores from two behavioral tasks, three-trial recall and retroactive interference, were selected by the decision tree as providing optimal discriminabiliity among the four groups of animals. The overall accuracy, as reported in Table 8.3, was 50% (Kappa = 0.33). Nontransgenic-control individuals were the least-likely to be correctly identified, with only 25% of NT mice recognized.

Response-Latencies Only. The decision tree-based classifier was unable to successfully distinguish among the four groups using only response-latency data. Indeed, none of the nontransgenic-control animals were correctly identified.

Errors and Latencies. The inclusion of both error-score and response-latency measures for constructing decision trees resulted in 41% accuracy (Kappa = 0.21) using four attributes: Three-trial recall errors and latency, retroactive interference errors, and delayed-recall latency. However, only 13% of nontransgenic-control animals were correctly identified by this classifier.

Summary: Although the overall discriminability among the four groups was relatively poor, this may be due to difficulty with distinguishing among NT, NT+GMCSF, and Tg+GMCSF animals. However, as reported in Table 8.3, error-scores alone provided the best four-way discriminability (50%) among the decision tree-based classifiers.

8.4.2 Neural Network Analysis

Nontransgenic (NT) vs. Transgenic (Tg) Groups

Error-Scores Only. When provided with error-score measures from all tasks of the interference paradigm, the neural network-based classifier correctly assigned 75% of the individuals into their respective groups (Kappa = 0.50), with optimal performance obtained with three computing elements in the hidden layer of the network.

The sensitivity was 75% and specificity was 75% for this classifier, as indicated in Table 8.3. Indeed, the same proportion of animals from each group were misclassified.

Response-Latencies Only. Using only behavioral response-latency data, the neural network classifier exhibited 81% accuracy (Kappa = 0.63), with 75% sensitivity and 88% specificity. A slightly greater proportion of nontransgenic animals were correctly identified, relative to the proportion of transgenics.

Errors and Latencies. Only 69% of individuals (Kappa = 0.38) were correctly identified by a neural network-based classifier trained with both error-scores and response-latency data. As reported in Table 8.3, the sensitivity of the classifier was 63% and the specificity was 75%.

Summary: The optimal neural network-based classifier for distinguishing between nontransgenic and Alzheimer's transgenic mice used only response-latency data, and showed a slight performance bias favoring nontransgenic animals, with respect to correct identification by genotype.

Nontransgenic (NT) vs. Nontransgenic-Treatment (NT+GMCSF) Groups

Error-Scores Only. Using only behavioral error-score data, the neural network-based classifier correctly assigned only 63% of the individuals into their respective groups (Kappa = 0.25), with 63% sensitivity and 63% specificity. Hence, approximately two-thirds of the animals in each group were correctly identified.

Response-Latencies Only. Only 56% accuracy (Kappa = 0.13) was achieved by neural network classifiers trained using only behavioral response-latency data from the interference paradigm. The observed sensitivity of the classifier was 63%, and its specificity was 50%. Hence, random-chance level classification was observed for the untreated nontransgenic animals.

Errors and Latencies. The neural network-based classifier correctly assigned only 56% of the individuals (Kappa = 0.13) into their respective groups, with 50% sensi-

tivity and 63% specificity, as shown in Table 8.3. Optimal performance was obtained with a neural network architecture containing four computing elements in the hidden layer.

Summary: There was difficulty distinguishing between GMCSF-treated and untreated nontransgenic mice, although modest discriminability was observed in neural network-based classifiers trained using behavioral error-score data from the interference paradigm.

Transgenic (Tg) vs. Transgenic-Treatment (Tg+GMCSF) Groups

Error-Scores Only. A neural network-based classifier trained using only behavioral error-scores correctly assigned 100% of the individuals to their respective groups (Kappa = 1.00), with both sensitivity and specificity of 100%, as shown in Table 8.3. Optimal performance of this classifier was achieved with three computing elements in the hidden layer of the network.

Response-Latencies Only. Using only behavioral response-latency data to train a neural network-based classifier resulted in a lower (88%) accuracy (Kappa = 0.75), with 88% sensitivity and 88% specificity. This neural network required four hidden layer elements for optimal performance.

Errors and Latencies. The neural network-based classifier correctly assigned 94% of the individuals into their respective groups (Kappa = 0.88), with optimal performance obtained with four computing elements in the hidden layer of the network. The sensitivity was 100% and specificity was 88%. Hence, all GMCSF-treated Alzheimer's transgenic animals were correctly identified.

Summary: Neural network-based classifiers trained with behavioral data from the interference paradigm are extremely effective for distinguishing between GMCSF-treated and untreated Alzheimer's transgenic mice. Additionally, the effect of network topology (number of computing elements within the hidden layer) on classifier

performance suggests different underlying complexity associated with latency data, relative to error data.

All Four Groups

Error-Scores Only. As indicated in Table 8.3, the neural network-based classifier demonstrated 48% overall accuracy (Kappa = 0.29) in assigning animals to their respective groups on the basis of behavioral error-scores. A network architecture with four computing elements in the hidden layer was necessary for optimal performance. The poorest discriminability was observed for both nontransgenic groups, wherein only 38% of individuals from each group (NT, NT+GMCSF) were correctly identified, underscoring the similarity of these groups with respect to cognitive performance in the interference paradigm.

Response-Latencies Only. The neural network-based classifier was unable to distinguish among the four groups of mice using only behavioral response-latency data. Indeed, none of the nontransgenic-control (NT) animals were correctly identified.

Errors and Latencies. Using both error-scores and response-latencies, 41% of the animals were correctly assigned to their respective groups (Kappa = 0.21). The GMCSF-treated Alzheimer's transgenic mice were the most likely individuals to be correctly identified (50%), while only 38% of individuals in each of the other three groups were correctly assigned to their respective group. Four computing elements were present in the hidden layer of the optimal network.

Summary: The optimal neural network-based classifier for distinguishing among the four groups used only behavioral error-score data, provided poor overall discriminability (48%), and showed a performance bias favoring the transgenic (both GMCSF-treated and untreated) subjects. The inclusion of response-latency data compromised overall discriminability among groups. Additionally, the optimal network topologies required four computing elements in the hidden layer, regardless of

how many candidate predictor variables (behavioral measures) were supplied to the classifier.

8.4.3 Support Vector Machine Analysis

Nontransgenic (NT) vs. Transgenic (Tg) Groups

Error-Scores Only. As shown in Table 8.3, the support vector machine-based classifier correctly assigned 88% (Kappa = 0.75) of control nontransgenic and Alzheimer's transgenic mice to their respective groups, using only behavioral error-scores from all tasks of the interference paradigm. Both the sensitivity and specificity of the classifier were 88%. Hence, equal proportions of each group were correctly identified by the classifier.

Response-Latencies Only. Using only behavioral response-latency data, the classifier demonstrated 81% accuracy (Kappa = 0.63), with 75% sensitivity and 88% specificity, as reported in Table 8.3.

Errors and Latencies. When both error-scores and response-latencies were provided to the classifier, 75% of the animals were correctly assigned to their respective groups (Kappa = 0.50). The sensitivity was 63% and the specificity was 88%. This classifier exhibited a performance bias favoring nontransgenic mice, with greater likelihood of misclassifying Alzheimer's transgenics.

Summary: The support vector machine-based classifiers successfully distinguished between nontransgenic and Alzheimer's transgenic mice using error-score data. However, the use of response-latency data, either exclusively or in combination with error-scores, resulted in a marked decline of both accuracy and specificity for the transgenic animals.

Nontransgenic (NT) vs. Nontransgenic-Treatment (NT+GMCSF) Groups

Error-Scores Only. The support vector machine-based classifier correctly identified 63% of individuals (Kappa = 0.25), with 63% sensitivity and 63% specificity, as shown in Table 8.3.

Response-Latencies Only. The accuracy of the classifier increased to 69% (Kappa = 0.38) when only response-latency data from the interference paradigm were used to distinguish between GMCSF-treated and untreated nontransgenic mice. The observed sensitivity was 63% and the specificity was 75%. Three-fourths of the untreated nontransgenics were correctly identified.

Errors and Latencies. Training with both error-scores and response-latencies resulted in a classifier which exhibited only 56% accuracy (Kappa = 0.13), as reported in Table 8.3. This classifier showed 63% sensitivity and 50% specificity, erroneously identifying half of the untreated animals as having received GMCSF treatment.

Summary: Overall, behavioral response-latency measures provided the best criteria for distinguishing between GMCSF-treated and untreated nontransgenic mice using support vector machine-based classifiers, although the discrimination provided was modest at best. Classifier performance degradation occurred, however, when error-scores were used exclusively or in combination with response-latencies.

Transgenic (Tg) vs. Transgenic-Treatment (Tg+GMCSF) Groups

Error-Scores Only. As reported in Table 8.3, support vector machine-based classifiers trained exclusively with behavioral error-score data showed 75% accuracy (Kappa = 0.50), with 100% sensitivity and 50% specificity (nonsignificant). Hence, all GMCSF-treated transgenics were correctly identified, while one-half of untreated animals were erroneously recognized as having received GMCSF treatment.

Response-Latencies Only. Classifiers trained only with behavioral response-latency data exhibited 75% overall accuracy (Kappa = 0.50), with 75% sensitivity and specificity. Hence, individuals from both groups were equally likely to be misclassified.

Errors and Latencies. When both errors and latencies were used to train the classifier, the accuracy increased to 82% (Kappa = 0.63), as indicated in Table 8.3, with 100% sensitivity and 63% specificity. Similar to the results obtained from using error-scores exclusively, all GMCSF-treated Alzheimer's transgenic mice were correctly identified.

Summary: Support vector machine-based classifiers trained with either error-score or response-latency data were comparable, with respect to overall discriminability between GMCSF-treated and untreated transgenic mice, the combination of both behavioral measures significantly increased overall performance while preserving the observed sensitivity of the Error-Scores Only classifier to treatment effect in transgenic animals.

All Four Groups

Error-Scores Only. Support vector machine-based classifiers failed to distinguish among all four groups of mice using only error-score data from the interference paradigm. Indeed, only one-half of the untreated Alzheimer's transgenic animals, and none of the untreated nontransgenics, were correctly identified.

Response-Latencies Only. When response-latency data were used to train the classifier, the overall accuracy was a poor, albeit significant, 28% (Kappa = 0.04), as shown in Table 8.3. However, none of the GMCSF-treated Alzheimer's transgenic mice or untreated nontransgenics were correctly identified. Only 63% of the untreated transgenic animals were correctly classified.

Errors and Latencies. The classifier did not successfully distinguish among the four groups of animals using both error and latency measures from the interfer-

ence paradigm. Similar to the Response-Latencies Only result, neither the GMCSF-treated transgenics nor the untreated nontransgenic mice were recognized, and only 63% of untreated transgenics were identified correctly. In addition, fewer than half of the GMCSF-treated nontransgenics were correctly assigned to their treatment group.

Summary: Overall, support vector machine-based classifiers were generally unsuccessful in distinguishing among the four groups of mice. Significant discriminability was observed when behavioral response-latency data were used exclusively as predictor variables, but the resulting classifier showed exclusive bias favoring the untreated Alzheimer's transgenic mice, and was otherwise ineffective.

8.5 Discussion

Standard statistical analyses (using ANOVA) of final-block behavioral measures indicate significant groupwise differences in both the post-test Radial Arm water maze working memory trial T5 (error-score), as well as in three of the four tasks of the interference paradigm (as shown in Figures 8.1 and 8.2). These differences were observed between Tg animals and the other three groups. Additionally, nontransgenic mice (both NT and NT+GMCSF) and GMCSF-treated Alzheimer's transgenic mice did not differ significantly in cognitive performance in the RAWM trial T5 measure. These results are consistent with prior findings (e.g., Leighty et al., 2004; Arendash et al, 2006) showing significant ($p < .05$) working memory impairment (RAWM T4 and T5 measures) in Alzheimer's transgenic-control (Tg) mice, relative to nontransgenic-controls (NT). Significant differences in final-block performance in both the three-trial recall and delayed-recall tasks of the interference paradigm suggest cognitive impairment in Alzheimer's transgenic-control mice, relative to nontransgenic-controls. Moreover, GMCSF-treated Alzheimer's transgenic mice performed significantly better in the delayed-recall task during the final block

of the interference paradigm, relative to transgenic-control animals, indicative of a therapeutic effect of GMCSF treatment. By contrast, no significant differences were found between GMCSF-treated and untreated nontransgenic mice in any of the component measures of the interference paradigm, suggesting that GMCSF may not provide benefits in individuals not predisposed for Alzheimer-like pathology. Finally, although only a single non-interference paradigm measure was compared among the groups – final-block RAWM trial T5 error-scores – the component tasks of the interference paradigm provide superior groupwise discriminability using standard statistical methods of analysis.

Extensive significant pairwise correlation was observed among both error-score and response-latency measures of all four behavioral tasks. Corresponding error and latency measures for each task were also strongly correlated, underscoring the comparability of these behavioral indices for evaluating cognitive performance. Both behavioral measures of the proactive interference task were significantly correlated with the corresponding measure (error or latency) of each of the other tasks, except for three-trial recall. These correlation results were similar to those obtained in the GRK5-related investigation (refer to Chapter 7), with the exception of the Proactive Interference task correlations with other tasks.

Exploratory factor analysis provided evidence for segregation by learning condition (measures related to Pool A vs. measures associated with Pool B) in the preceding ("The Interference Task: A Novel Assessment Paradigm") study, although segregation-by-condition was not observed in this study. Indeed, learning associated with both Pool A and Pool B was reflected in the second factor. Not surprisingly, in light of the extensive pairwise intercorrelations between/among measures of the Three-Trial Recall and Delayed-Recall tasks, these two components of the interference paradigm together comprised the primary factor (Table 8.2). Consequently, the

primary factor may be interpreted as representing performance in the initial learning task (Three-Trial Recall) and following a substantial delay (Delayed-Recall), thus reflecting successful consolidation of early learning (Pool A) in the interference paradigm reminiscent of the primacy effect observed in human serial recall. Similarly, significant intercorrelation between/among measures of the Proactive Interference and Retroactive Interference tasks is reflected in the coexistence of these two tasks in Factor II, in contrast to the factor structure of the interference paradigm recovered in the preceding study wherein only Proactive Interference measures comprised the second factor. The second factor may be related to context-switching or the capacity to adapt to new learning conditions (Pool A vs. Pool B). It is important to recognize, however, that both correlation and factor structures may be different if either first-block or overall data were used for these analyses (instead of final-block measures). This may occur, in part, as a consequence of the nonlinear time-dependence of learning and memory (e.g., Gallistel et al., 2004), whereby discrepancies in groupwise analytic results may arise from behavioral sampling at different stages of learning/memory acquisition. Moreover, distinctive (often, subtle) statistical features of the subject pool (e.g., significantly unequal group sizes, differential treatment effect, discrepancies in group variances, etc.) may introduce bias or other undesirable side-effects in the analysis.

The discriminant analysis results were generally consistent with prior studies, wherein stepwise-forward analyses often return superior discriminability relative to the standard direct-entry (complete) approach (e.g., Arendash and King, 2002; Leighty et al., 2004). Discriminant analysis-based classifiers successfully distinguished both transgenicity (Tg vs. NT) and GMCSF-treatment effect in Alzheimer's transgenic mice (Tg vs. Tg+GMCSF), and showed moderate discriminability (53%), albeit superior to all other classifiers, in the four-group comparison. Indeed, despite over-

all poor performance (69% accuracy), the variance-optimizing stepwise-forward approach successfully distinguished between the NT and NT+GMCSF groups, although the consensus among classifiers was that these two groups do not differ with respect to cognitive measures in the interference paradigm.

The decision tree-based classifiers successfully distinguished between nontransgenic and Alzheimer's transgenic mice at a level comparable to direct-entry discriminant analysis, using interference task error-scores (with or without response-latency data). In addition, decision trees distinguished between GMCSF-treated and untreated Alzheimer's transgenic mice at a level comparable to direct-entry discriminant analysis. In contrast to discriminant analysis, the inclusion of response-latencies to the error data did not improve the overall accuracy of the decision tree. Overall discriminability among the four groups of animals demonstrated by decision trees, using interference task error-scores alone, was between that of the two discriminant analysis-based techniques.

Neural networks were particularly effective for distinguishing between GMCSF-treated and untreated Alzheimer's transgenic mice using error-scores alone. However, neural networks performed comparably to stepwise-forward discriminant analysis when using latency data for detecting the main effect of Alzheimer's transgenicity and, indeed, somewhat better than decision trees for distinguishing between NT and NT+GMCSF individuals.

Support vector machines performed comparably to direct-entry discriminant analysis for distinguishing transgenicity using either error-score or response-latency, but not both, behavioral measures from the interference paradigm. When comparing between the Tg and Tg+GMCSF groups, however, the support vector machine-based classifier was poorer than discriminant analysis and markedly inferior to neural network implementations. Moreover, support vector machines showed the poorest over-

all discriminability among the four groups, achieving an acceptable performance level (25% or better) only for response-latency data.

In addition to comparisons by classifier-type, to determine the relative strengths and weaknesses of each methodology, it is informative to compare performance by subject group, examining both pairwise contrasts and the composite (four-group) results. This study examined two main effects, Alzheimer's transgenicity and GMCSF administration, in a classic 2 x 2 design. In addition, two cognitive metrics (error-score and response-latency) were measured for each behavioral task component of the interference paradigm. The main effect of Alzheimer's transgenicity was best detected by stepwise-forward discriminant analysis, resulting in 94% overall accuracy, although the other classifiers also performed remarkably well. Additionally, classification using error-scores alone was generally superior to efforts including response-latencies (which provided no significant advantage). As discussed previously, the APP genotype produces a distinct behavioral phenotype, with respect to progressive cognitive impairment, and represents a useful model for investigating Alzheimer-like neuropathology through its overt behavioral manifestations. Indeed, standard statistical analyses (ANOVA) of both final-block RAWM trial T5 error-scores and two tasks of the interference paradigm (Three-Trial Recall and Delayed-Recall, both errors and latencies) clearly differentiated between Tg mice and the other groups. The consensus among classifiers is that the GMCSF-treatment effect in nontransgenic animals is minimal, i.e., these mice do not benefit significantly from receiving GM-CSF treatment. Indeed, only modest discriminability was consistently observed in both neural networks and support vector machines for these two groups. ANOVA-based analyses of the final-block RAWM trial T5 failed to distinguish between the NT and NT+GMCSF groups, and modest (but nonsignificant) differences were observed in both error-scores and response-latencies of the Proactive Interference and

Retroactive Interference tasks of the interference paradigm. The GMCSF-treatment effect in Alzheimer's transgenic mice, by contrast, is significant. Indeed, all classifiers distinguished this treatment effect using error-scores and/or response-latencies obtained in the interference paradigm. In fact, only decision trees were unable to distinguish between the groups using only latencies. Standard statistical analysis of final-block RAWM trial T5 showed a significant impairment in the Tg group, relative to Tg+GMCSF. Similarly, ANOVA-based analyses of both Three-Trial Recall and Delayed-Recall measures (errors and latencies) underscore the cognitive deficits of Alzheimer's transgenic-control mice relative to transgenics receiving GMCSF treatment. Moreover, the Tg and Tg+GMCSF groups were often more easily distinguishable than were the Tg and NT groups. Error-scores were generally more effective for achieving composite (four-group) discriminability, although the stepwise-forward discriminant analysis provided the best overall performance using a combination of errors and latencies (59%).

Several important conclusions and cautions may be drawn from the results. First, the Alzheimer's transgenic genotype (APP) reliably produces a distinctive cognitive-behavioral phenotype against which diverse genetic manipulations and therapeutic interventions may be explored, analyzed, and interpreted in the context of human Alzheimer's disease diagnosis, treatment, and management. Indeed, the cognitive repertoire of this behavioral phenotype is sufficiently rich, with respect to discernible and quantifiable domains (e.g., working memory), that comprehensive, multitask-multimetric assessment returns an extraordinarily detailed portrait of abilities, as previously reported (e.g., Leighty et al., 2004). Second, the differential sensitivity of classifiers to cognitive-behavioral features suggests not only the salience of particular attributes (individual measures, such as three-trial recall error-score), with respect to groupwise discriminability, but, moreover, the likelihood that nonlinear interactions

between/among attributes contribute to the observed discriminability. Analysis of variance, for example, although widely accepted in neurobehavioral research, displays significant limitations with respect to groupwise discriminability (e.g., Figure 8.1, discussed earlier) in part because of its dependence upon specific, and often unrealistic, measurement conditions. Indeed, many statistical methodologies are dependent upon – and optimized for – linear models of predictor variables (e.g., ANOVA, multiple regression, linear discriminant analysis). However, nonlinear interactions are quite common in natural physiological systems (e.g., Glass, 2001; Westbury et al., 2003), and require special mathematical treatment for proper analysis. As discussed earlier, neural networks are particularly suitable for applications in which the relationship within and between sets of variables (measures) is likely to be nonlinear or, indeed, unknown. Fourth, this study further supports adoption of the mouse-based interference paradigm as a standard component of neurobehavioral assessment for research in Alzheimer's transgenic mice. The component tasks of the paradigm were shown to contribute unique information reflecting cognitive status in individuals, as well as discriminative criteria for distinguishing among groups by transgenic and/or treatment effect. Finally, although this study did not address the therapeutic window (e.g., first-block vs. final-block vs. overall efficacy), duration of response (permanent vs. transient), dose-dependence, or age-dependence of GMCSF efficacy, in light of the observed therapeutic effect of GMCSF treatment in Alzheimer's transgenic animals (and, possibly, to a lesser extent in nontransgenics), additional investigation of GMCSF's potential role in AD therapy is indicated.

CHAPTER 9

CONCLUSIONS

Alzheimer's disease is a multifactorial disorder, exhibiting a remarkably complex portrait of neuropathological phenomena and cognitive/behavioral manifestations. Contemporary research methodologies must necessarily address these features in order to develop and refine valid experimental models (e.g., transgenic animals), comprehensive neurobehavioral assessment paradigms, and rigorous data analysis protocols. Indeed, the sophistication of modeling, evaluation, and analysis closely parallels advances in both computational- and bio-technology. Consequently, today's medical scientists and clinicians possess a powerful array of tools for investigating both pathological (from molecular-genetic to organ-system scale) and behavioral (sensorimotor, cognitive, daily living) aspects of Alzheimer's disease.

The research presented in this dissertation addressed neurobehavioral assessment and analysis in both Alzheimer's transgenic mice and human AD patients. The utility of comprehensive sensorimotor and cognitive behavioral evaluation for both transgenic and treatment effects was demonstrated, underscoring the importance of multiple tasks and/or measures for identifying and quantifying group-differences in animal models of human AD. Data mining techniques (decision trees, neural networks, support vector machines), underrepresented in behavioral research, were compared with conventional statistical analyses (correlation, discriminant analysis, factor analysis) to explore the diagnostic potential of data mining methodologies in both human- and mouse-based studies. A semantic interference learning task originally developed

for human AD assessment was adapted for mice, by replacing verbal-semantic elements with visual-spatial components. The resulting parallel evaluation paradigm effectively distinguished both transgenic and treatment effects in animal models.

Comprehensive assessment, using multiple tasks and measures of sensorimotor and cognitive ability, provides the best opportunity to detect anomalous behaviors suggestive of underlying neuropathology. By sampling across diverse behavioral domains (e.g., balance, agility, spatial learning, reference memory) and recording multiple intratask response metrics (e.g., acquisition and retention measures in the Morris water maze task), a broader cross-section of the behavioral repertoire is revealed, thus providing the investigator with a more complete portrait of the subject's overall neurobehavioral status. Subsequent analysis of behavioral measures reveals evidence of the complex interplay among, as well as within, sensorimotor and cognitive domains which, collectively, are expressed as a distinct syndrome. The significant correlations between/among behavioral measures observed within individual cognitive-loaded tasks, such as the radial arm water maze, reflects both the extent of integration among diverse cognitive functions necessary for successful overall performance in these tasks, as well as the sweeping impact of incipient neuropathology across multiple cognitive domains. Strong consensus among cognitive measures obtained from the water maze-type tasks (RAWM, Morris water maze, platform recognition task) underscores the ability of these protocols to evaluate similar features of learning and memory processes. Indeed, factor analyses of the comprehensive task battery consistently returned a primary factor comprised largely of measures from the RAWM and platform recognition tasks, representing overall cognitive functioning, which may be analogous to the putative general mental ability "g" construct in humans.

Statistical approaches for identifying and quantifying group differences form the analytic foundation of both experimental and clinical neurobehavioral research. Data mining techniques, by contrast, have been confined largely to industrial and commercial applications, and are only recently gaining recognition in biomedical research. Classifier models produced by stepwise-forward discriminant analysis were generally superior to those returned by direct-entry (complete) discriminant analysis for distinguishing between/among groups based on treatment and/or transgenic effects, using behavioral measures. The variance-optimization capacity of the stepwise-forward approach, in conjunction with the availability of a relatively large sampling of the mouse behavioral repertoire (i.e., many measures from which to select predictor variables), increased the likelihood for successful discriminability. The resulting classifiers often selected a collection of sensorimotor, anxiety, and cognitive measures, stressing the importance of multimetric assessment inventories. Moreover, these findings underscore the utility of the assessment battery for mouse behavioral phenotyping (characterizing groups of animals by treatment and/or transgenic effects using discernable behavioral features).

Decision trees, which select predictor variables based on information-bias capacity, did not exhibit consistent performance but, rather, idiosyncratic (subject to differential performance with different datasets) behavior in the relatively-smaller datasets of mouse-based studies. For this reason, unless the investigator has substantial prior experience with the experimental treatment conditions, to make informed decisions concerning predictor variable selection, reliance on decision tree-based classifiers is not recommended, particularly when small sample sizes are present. Neural networks and support vector machines also exhibited modestly idiosyncratic, albeit more consistent, behavior for analyzing groupwise behavioral measures. Indeed, either/both of these two methods were found to be superior to discriminant

analyses, with respect to groupwise discriminability. These findings suggest that memory/learning processes, mirroring the underlying physiological substrates, may exhibit nonlinear behavioral manifestations to which neural network and support vector machine architectures may be more sensitive (compared to multivariate statistical analyses). Moreover, these results indicate that statistical- and data mining-based classifiers may be complementary techniques, rather than substitutes and, in addition, that sample size may be an important consideration when selecting among alternative analytic methodologies.

Two studies examining the cognitive effects of caffeine admininstration in mice were used to develop and refine advanced statistical and data mining analytic techniques for neurobehavioral research in both nontransgenic and Alzheimer's transgenic animals. The first study (Chapter 4), involving aged nontransgenic mice which received oral caffeine (1.5 mg/day) for approximately ten months suggests that, although caffeine may exert subtle influence on both sensorimotor and cognitive behavior, it is unlikely that chronic caffeine intake provides cognitive benefits in normal (nontransgenic) animals, as reflected in poor groupwise discriminability by both advanced statistical and data mining-based classifiers. The effects of long-term caffeine administration in Alzheimer's transgenic mice were investigated in the second study (Chapter 5). After consuming oral caffeine (1.5 mg/day) for four to five weeks, caffeine-treated transgenic animals performed comparably to age-matched nontransgenic mice and, indeed, superior to untreated transgenic animals, with respect to cognitive measures of working memory, as indicated by superior discriminability between/among treatment groups by both sets of classifiers.

The semantic interference task (Loewenstein et al., 2004) was recently developed as a clinical diagnostic/screening instrument for probable Alzheimer's disease. The original (human) version consists of verbally-reported immediate- and delayed-recall

memory components, which assess the subject's ability to maintain two distinct collections of familiar items in memory. When the dataset from the Loewenstein et al. (2004) study was analyzed using advanced statistical and data mining-based techniques (Chapter 6), these methodologies were found to be comparable with respect to overall accuracy, sensitivity, and specificity for distinguishing between/among probable Alzheimer's disease, mildly cognitively-impaired, and normal aged individuals (humans). Collectively, these results underscore the diagnostic importance of the semantic interference paradigm for Alzheimer's screening, as well as the use of data mining techniques for clinical applications typically confined to statistical methods. A mouse-based adaptation of the task was created by substituting spatial learning elements (similar apparatus to the RAWM) for the verbal components of the original task. The resulting paradigm generated performance measures (both error-scores and response-latencies) which were analogous to those of the original human-based design, reflecting: general recall and learning capacity, proactive interference, retroactive interference, and delayed-recall ability.

The mouse-based interference paradigm was compared to conventional neurobehavioral assessment protocols (e.g., RAWM, comprehensive behavioral task battery) in two studies, involving: (1) the cognitive effects of an additional genetic manipulation (elimination of GRK5 expression) superimposed on the Alzheimer's transgenic genotype, as well as nontransgenics, and, (2) therapeutic efficacy of GM-CSF administration in Alzheimer's transgenic mice and age-matched nontransgenic animals. The GRK5-knockout genotype is being investigated (e.g., Suo et al., 2007) as a model for Alzheimer's-like neuropathology. In Chapter 7, one-year-old mice (both nontransgenic and Alzheimer's transgenic) with an additional genetic modification (with/without GRK5 expression) completed both the comprehensive behavioral task battery and the mouse-based interference paradigm, and the results were

analyzed using standard statistics (ANOVA), advanced statistics (e.g., discriminant analysis), and data mining methods (decision trees, neural networks, support vector machines). Nontransgenic and Alzheimer's transgenic mice, both expressing GRK5, were found to differ in both sensorimotor and cogitive measures of the comprehensive battery, using both standard and advanced statistics, as well as data mining methods. These two groups also exhibit significant differences in component measures of the interference paradigm, based on standard and advanced statistics, as well as data mining analyses. By contrast, only modest discriminability with respect to GRK5 expression was observed in both nontransgenic and Alzheimer's transgenic animals. However, comparisons among nontransgenic mice expressing GRK5 and both Alzheimer's transgenic groups across measures in the interference paradigm suggest that eliminating expression of GRK5 may ameliorate cognitive impairment associated with the Alzheimer's transgene, bringing the Tg-KO animals closer to the baseline cognitive performance of GRK5-expressing nontransgenics. Indeed, GRK5-expressing nontransgenics only differ significantly from GRK5-knockout Alzheimer's transgenics with respect to a single behavioral measure of the interference paradigm, retroactive interference response-latency. Finally, the therapeutic efficacy of GM-CSF admininstration, with respect to cognitive function, was examined in both nontransgenic and Alzheimer's transgenic animals (Chapter 8). The inflammatogenic cytokine GM-CSF has been proposed as a treatment for Alzheimer's disease, to attenuate beta-amyloid plaque-associated structural and functional impairment, thereby improving or restoring cognitive function. One-year-old Alzheimer's transgenic mice received daily injections of GM-CSF (5 mcg/day) for two weeks, then completed both RAWM post-testing and the interference paradigm. GMCSF-treated transgenic animals performed signficantly superior to untreated transgenics and comparably to age-matched nontransgenic mice (receiving either GMCSF or plain vehicle) in the

RAWM trial T5 working memory error-score measure, as well as in the three-trial recall and delayed-recall measures of the interference paradigm. GMCSF, therefore, may ameliorate Alzheimer's transgene-associated cognitive impairment and, thus, represent a cognitive-protective treatment for individuals predisposed for Alzheimer's disease.

In theory, early detection and therapeutic intervention represents the best strategy for treating progressive disorders, such as Alzheimer's disease. Although specific genetic markers have been identified as risk factors for early-onset, familial Alzheimer's disease, the etiology of most AD cases (over ninety percent) remains unclear. Many situational (e.g., lifestyle, diet) and dispositional (e.g., other genetic factors, epigenetics) candidate risk factors are currently under investigation. Meanwhile, diagnostic, therapeutic, and management protocols currently exist for patients. Subtle manifestations of nascent cognitive impairment, for example, may be revealed through early psychometric indices (e.g., academic or mental ability test subscores). Certain individual differences in learning patterns (e.g., memory-task learning curve) which are commonly dismissed as psychological variability may, in fact, represent very early indications of cognitive dysfunction or, possibly, predisposition to dementia-like syndromes. Advanced analytic methodologies, as well, may lead to new insights through the integration of neurobehavioral and neurological databases. Customized hybrid and ensemble classifier systems, for instance, could be used to detect specific distinguishing features, patterns, or trends which reliably identify pathologic states, therapeutic efficacy, or differential treatment effects in both experimental (e.g., transgenic animals) and clinical settings. Finally, in recognition of the spatial and temporal aspects of Alzheimer's disease, investigative paradigms which exploit the spatiotemporal dynamics of AD should be pursued. For example, certain computational architectures, such as self-organizing maps, are particularly

suitable for multidimensional modeling of nonlinear dynamical processes, such as regiospecific patterns of brain activation.

Contemporary Alzheimer's research has become a multidisciplinary endeavor, bringing together the diverse insights and talents of both academic and professional specialties. Alzheimer's disease represents a significant public health concern which calls for a commensurate share of social, political, and scientific attention. Indeed, there is growing public awareness of Alzheimer's disease and related research activities, as reported by the news media and portrayed in popular entertainment. Each advance renews our hopes, and each setback strengthens our resolve: We recognize this familiar ebb and flow as the natural pace of progress, and find solace in its continuity.

REFERENCES

Abdi, H. (1994). A neural network primer. *Journal of Biological Systems*, **2**, 247-283.

Ackley, D. H., Hinton, G. E., and Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, **9**, 147-169.

Adlard, P. A., and Cummings, B. J. (2004). Alzheimer's disease – A sum greater than its parts? *Neurobiology of Aging*, **25**, 725-733.

Adrian, E. D., and Yamagiwa, K. (1935). The origin of the Berger rhythm. *Brain*, **58**, 323-351.

Aftanas, L. I., and Golocheikine, S. A. (2002). Non-linear dynamic complexity of the human EEG during meditation. *Neuroscience Letters*, **330**, 143-146.

Aggarwal, B. B., and Shishodia, S. (2004). Suppression of the nuclear factor κB activation pathway by spice-derived phytochemicals: Reasoning for seasoning. *Annals of the New York Academy of Science*, **1030**, 434-441.

Alkon, D. L. (1974). Associative training of *Hermissenda*. *Journal of General Physiology*, **64**, 70-84.

Almeida, J. S. (2002). Predictive non-linear modeling of complex data by artificial neural networks. *Current Opinion in Biotechnology*, **13**, 72-76.

Alonso, A. del C., Zaidi, T., Novak, M., Grundke-Iqbal, I., and Iqbal, K. (2001). Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments / straight filaments. *Proceedings of the National Academy of Sciences USA*, **98**, 6923-6928.

Altman, D. G., and Bland, J. M. (1994). Diagnostic tests. 1. Sensitivity and specificity. *British Medical Journal*, **308**, 1552.

Alzheimer, A. (1907). Über eine eigenartige Erkrankung der Hirnrinde. *Allgemeine Zeitschrift fur Psychiatrie und psychiatrisch-gerichtliche Medizin* (Berlin), **64**, 146-148.

Alzheimer's Association. (2008). 2008 Alzheimer's Disease Facts and Figures. *Alzheimer's and Dementia*, **4(2)**.

Alzheimer's Disease Collaborative Group (1995). The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families. *Nature Genetics*, **11**, 219-222.

Amari, S., Beltrame, F., Bjaalie, J. G., Dalkara, T., De Schutter, E., Egan, G. F., Goddard, N. H., Gonzalez, C., Grillner, S., Herz, A., Hoffmann, K. P., Jaaskelainen, I., Koslow, S. H., Lee, S. Y., Matthiessen, L., Miller, P. L., Da Silva, F. M., Novak, M., Ravindranath, V., Ritz, R., Ruotsalainen, U., Sebestra, V., Subramaniam, S., Tang, Y., Toga, A. W., Usui, S., Van Pelt, J., Verschure, P., Willshaw, D., Wrobel, A., and the OECD Neuroinformatics Working Group. (2002). Neuroinformatics: The integration of shared databases and tools toward integrative neuroscience. *Journal of Integrative Neuroscience*, **1**, 117-128.

Ambree, O., Leimer, U., Herring, A., Gortz, N., Sachser, N., Heneka, M. T., Paulus, W., and Keyvani, K. (2006). Reduction of amyloid angiopathy and Abeta plaque burden after enriched housing in TgCRND8 mice. *American Journal of Pathology*, **169**, 544-552.

Amieva, H., Lafont, S., Rouch-Leroyer, I., Rainville, C., Dartiques, J. F., Orgogozo, J. M., and Fabriqoule, C. (2004). Evidencing inhibitory deficits in Alzheimer's disease through interference effects and shifting disabilities in the Stroop test. *Archives of Clinical Neuropsychology*, **19**, 791-803.

Amouyel, P., Vidal, O., Launay, J. M., and Laplanche, J. L. (1994). The apolipoprotein E alleles as major susceptibility factors for Creutzfeldt-Jakob disease. *Lancet*, **344**, 1315-1318.

Angelucci, M. E., Vital, M. A., Cesario, C., Zadusky, C. R., Rosalen, P. L., and Du Cunha, C. (1999). The effect of caffeine in animal models of learning and memory. *European Journal of Pharmacology*, **373**, 135-140.

Arendash, G. W., Garcia, M. F., Costa, D. A., Cracchiolo, J. R., Wefes, I. M., and Potter, H. (2004a). Environmental enrichment improves cognition in aged Alzheimer's transgenic mice despite stable β-amyloid deposition. *NeuroReport*, **15**, 1751-1754.

Arendash, G. W., Jensen, M. T., Salem, N., Jr., Hussein, N., Cracchiolo, J., Dickson, A., Leighty, R., and Potter, H. (2007). A diet high in omega-3 fatty acids does not improve or protect cognitive performance in Alzheimer's transgenic mice. *Neuroscience*, **149**, 286-302.

Arendash, G. W., and King, D. L. (2002). Intra- and intertask relationships in a behavioral test battery given to Tg2576 transgenic mice and controls. *Physiology and Behavior*, **75**, 643-652.

Arendash, G., King, D., Gordon, M., Morgan, D., Hatcher, J., Hope, C., and Diamond, D. (2001). Progressive, age-related behavioral impairments in transgenic mice carrying both mutant amyloid precursor protein and presenilin-1 transgenes. *Brain Research*, **891**, 42-53.

Arendash, G. W., Lewis, J., Leighty, R. E., McGowan, E., Cracchiolo, J. R., Hutton, M., and Garcia, M. F. (2004b). Multi-metric behavioral comparison of APPsw and P301L models for Alzheimer's Disease: Linkage of poorer cognitive performance to tau pathology in forebrain. *Brain Research*, **1012**, 29-41.

Arendash, G. W., Mori, T., Cao, C., Mamcarz, M., Runfeldt, M., Dickson, A., Rezai-Zadeh, K., Tan, J., Citron, B. A., Lin, X., Echeverria, V., and Potter, H. (2009). Caffeine reverses cognitive impairment and decreases brain Aβ levels in aged Alzheimer's mice. *Journal of Alzheimer's Disease*. **In press.**

Arendash, G. W., Schleif, W., Rezai-Zadeh, K., Jackson, E. K., Zacharia, L. C., Cracchiolo, J. R., Shippy, D., and Tan, J. (2006). Caffeine protects Alzheimer's mice against cognitive impairment and reduces brain β-amyloid production. *Neuroscience*, **142**, 941-952.

Astur, R. S., Tropp, J., Sava, S., Constable, R. T., and Markus, E. J. (2004). Sex differences and correlations in a virtual Morris water task, a virtual radial arm maze, and mental rotation. *Behavioral Brain Research*, **151**, 103-115.

Atkinson, G. F. (1966). Designs for sequences of treatments with carryover effects. *Biometrics*, **22**, 292-309.

Backman, L., Jones, S., Berger, A.-K., Laukka, E. J., and Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology*, **19**, 520-531.

Backman, L., Small, B. J., and Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*, **124**, 96-102.

Baddeley, A. D. (1992). Working memory. *Science*, **255**, 556-559.

Baddeley, A. D., and Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.) *Recent Advances in Learning and Motivation*, **8**, 47-90. New York: Academic Press.

Bai, D. L., Tang, X. C., and He, X. C. (2000). Huperzine A: A potential therapeutic agent for treatment of Alzheimer's disease. *Current Medicinal Chemistry*, **7**, 355-374.

Bales, K., Verina, T., Cummins, D., Du, Y., Dodel, R., Saura, J., Fishman, C., DeLong, C., Piccardo, P., Petegnief, V., Ghetti, B. and Paul, S. (1999). Apolipoprotein E is essential for amyloid deposition in the APP (V171F) transgenic mouse model of Alzheimer's disease. *Proceedings of the National Academy of Sciences USA*, **96**, 15233-15238.

Balestreri, R., Fontana, L., and Astengo, F. (1987). A double-blind placebo controlled evaluation of the safety and efficacy of vinpocetine in the treatment of patients with chronic vascular senile cerebral dysfunction. *Journal of the American Geriatric Society*, **35**, 425-430.

Baños, J. H., and Franklin, L. M. (2002). Factor structure of the Mini-Mental State Examination in adult psychiatric inpatients. *Psychological Assessment*, **14**, 397-400.

Barnes, C. A. (1979). Memory deficits associated with senescence: A neurophysiological and behavioral study in the rat. *Journal of Comparative and Physiological Psychology*, **93**, 74-104.

Barnes, J., Scahill, R. I., Boyes, R. G., Frost, C., Lewis, E. B., Rossor, C. L., Rossor, M. N., and Fox, N. C. (2004). Differentiating Alzheimer's disease from aging using semiautomated measurement of hippocampal atrophy rates. *Neuroimage*, **23**, 574-581.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society, Series A*, **160**, 268-282.

Bartlett, M. S. (1947). The use of transformations. *Biometrics*, **3**, 39-52.

Bassett, S. S., Avramopoulos, D., and Fallin, M. D. (2002). Evidence for parent of origin effect in late-onset Alzheimer disease. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, **114**, 679-686.

Bassett, S. S., Avramopoulos, D., Perry, R. T., Wiener, H., Watson, B., Go, R. C. P., and Fallin, M. D. (2006). Further evidence of a maternal parent-of-origin effect on Chromosome 10 in late-onset Alzheimer's disease. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, **141B**, 537-540.

Bauer, E., and Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, **36**, 105-142.

Baxt, W. G. (1995). Application of artificial neural networks to clinical medicine. *Lancet*, **346**, 1135-1138.

Benzing, W., Wujek, J., Ward, E., Shaffer, D., Ashe, K., Younkin, S., and Brunden, K. (1999). Evidence for glial-mediated inflammation in aged APPsw transgenic mice. *Neurobiology of Aging*, **20**, 581-589.

Berman, K., and Brodaty, H. (2004). Tocopherol (vitamin E) in Alzheimer's disease and other neurodegenerative disorders. *CNS Drugs*, **18**, 807-825.

Biegler, R., McGregor, A., Krebs, J. R., and Healy, S. D. (2001). A larger hippocampus is associated with longer-lasting spatial memory *Proceedings of the National Academy of Sciences USA*, **98**, 6941-6944.

Billings, L. M., Oddo, S., Green, K. N., McGaugh, J. L., and LaFerla, F. M. (2005). Intraneuronal Abeta causes the onset of early Alzheimer's disease-related cognitive deficits in transgenic mice. *Neuron*, **45**, 675-688.

Bjornsson, H. T., Sigurdsson, M. I., Fallin, M. D., Irizarry, R. A., Aspelund, T., Cui, H., Yu, W., Rongione, M. A., Ekstrom, T. J., Harris, T. B., Launer, L. J., Eiriksdottir, G., Leppert, M. F., Sapienza, C., Gudnason, V., and Feinberg, A. P. (2008). Intra-individual change over time in DNA methylation with familial clustering. *Journal of the American Medical Association*, **299**, 2877-2883.

Blacker, D., Albert, M. S., Bassett, S. S., Go, R. C., Harrell, L. E., and Folstein, M. F. (1994). Reliability and validity of NINCDS-ADRDA criteria for Alzheimer's disease: The National Institute of Mental Health genetics initiative. *Archives of Neurology*, **51**, 1198-1204.

Blacker, D., Bertram, L., Saunders, A. J., Moscarillo, T. J., Albert, M. S., Wiener, H., Perry, R. T., Collins, J. S., Harrell, L. E., Go, R. C. P., Mahoney, A., Beaty, T., Fallin, M. D., Avramopoulos, D., Chase, G. A., Folstein, M. F., McInnis, M. G., Bassett, S. S., Doheny, K. J., Pugh, E. W., and Tanzi, R. E. (2003). Results of a high-resolution genome scan of 437 Alzheimer's disease families. *Human Molecular Genetics*, **12**, 23-32.

Bland, J. M., and Altman, D. G. (1996a). Statistics notes: Transforming data. *British Medical Journal*, **312**, 770.

Bland, J. M., and Altman, D. G. (1996b). Statistics notes: Transformations, means, and confidence intervals. *British Medical Journal*, **312**, 1079.

Bliss, T., and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, **232**, 331-356.

Bloom, F. E., Reilly, J. F., Redwine, J. M., Wu, C.-C., Young, W. G., and Morrison, J. H. (2005). Mouse models of human neurodegenerative disorders: Requirements for medication development. *Archives of Neurology*, **62**, 185-187.

Bock, C., and Lengauer, T. (2008). Computational epigenetics. *Bioinformatics*, **24**, 1-10.

Bohbot, V. D., Jech, R., Ruzicka, E., Nadel, L., Kalina, M., Stepankova, K., and Bures, J. (2002). Rat spatial memory tasks adapted for humans: Characterization in subjects with intact brain and subjects with selective medial temporal lobe thermal lesions. *Physiological Research*, **51 Suppl. 1**, S49-S65.

Bondi, M. W., Serody, A. B., Chan, A. S., Eberson-Shumate, S. C., Delis, D. C., Hansen, L. A., and Salmon, D. P. (2002). Cognitive and neuropathologic correlates of Stroop Color-Word Test performance in Alzheimer's disease. *Neuropsychology*, **16**, 335-343.

Boneau, A. (1960). The effects of violating assumptions underlying the t test. *Psychological Bulletin*, **57**, 49-64.

Borchelt, D., Ratovitski, T., van Lare, J., Lee, M., Gonzales, V., Jenkins, N., Copeland, N., Price, D., and Sisodia, S. (1997). Accelerated amyloid deposition in the brains of transgenic mice coexpressing mutant presenilin 1 amyloid precursor proteins. *Neuron*, **19**, 939-945.

Borras, D., Ferrer, I., and Pumarola, M. (1999). Age-related changes in the brain of the dog. *Veterinary Pathology*, **36**, 202-211.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In: Haussler, D. (Ed.) *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pp. 144-152. ACM Press.

Bossy-Wetzel, E., Schwarzenbacher, R., and Lipton, S. A. (2004). Molecular pathways to neurodegeneration. *Nature Medicine*, **10**, S2-S9.

Bourin, M., Chenu, F., Ripoll, N., and David, D. J. P. (2005). A proposal of decision tree to screen putative antidepressants using forced swim and tail suspension tests. *Behavioral Brain Research*, **164**, 266-269.

Bourre, J. M. (2004). Roles of unsaturated fatty acids (especially omega-3 fatty acids) in the brain at various ages and during aging. *Journal of Nutrition, Health and Aging*, **8**, 163-174.

Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, **26**, 211-252.

Bracco, L., Gallato, R., Grigoletto, F., Lippi, A., Lepore, V., Bino, G., Lazzaro, M. P., Carella, F., Piccolo, T., and Pozzilli, C. (1994). Factors affecting course and survival in Alzheimer's disease: A 9-year longitudinal study. *Archives of Neurology*, **51**, 1213-1219.

Brasnjevic, I., Steinbusch, H. W. M., and Schmitz, C. (2006). Altered gene expression and neuropathology in Alzheimer's disease. *Neurobiology of Aging*, **27**, 1081-1083.

Brenner, R. P., Ulrich, R. F., Spiker, D. G., Sclabassi, R. J., Reynolds, C. F., Marin, R. S., and Boller, F. (1986). Computerized EEG spectral analysis in elderly normal, demented, and depressed subjects. *Electroencephalography and Clinical Neurophysiology*, **64**, 483-492.

Brenner, R. P., Reynolds, C. F., and Ulrich, R. F. (1988). Diagnostic efficacy of computerized spectral versus visual EEG analysis in elderly normal, demented, and depressed subjects. *Electroencephalography and Clinical Neurophysiology*, **69**, 110-117.

Breslow, N. (1970). A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika*, **57**, 579-594.

Briley, M., Chopin, P., and Veigner, M. (1986). The "plus maze test of anxiety," validation in different rat strains and effect of a wide variety of antidepressants. *British Journal of Pharmacology*, **87**, 217P.

Britschgi, M., and Wyss-Coray, T. (2007). Systemic and acquired immune responses in Alzheimer's disease. *International Review of Neurobiology*, **82**, 205-233.

Broadhurst, P. L. (1961). Analysis of maternal effects in the inheritance of behavior. *Animal Behavior*, **9**, 129-141.

Brookmeyer, R., Gray, S., and Kawas, C. (1998). Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *American Journal of Public Health*, **88**, 1337-1342.

Brouwers, N., Sleegers, K., Engelborghs, S., Bogaerts, V., Serneels, S., Kamali, K., Corsmit, E., De Leenheir, E., Martin, J.-J., De Deyn, P. P., Van Broeckhoven, C., and Theuns, J. (2006). Genetic risk and transcriptional variability of amyloid precursor protein in Alzheimer's disease. *Brain*, **129**, 2984-2991.

Brown, J. A. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, **10**, 12-21.

Buchman, A. S., Schneider, J. A., Wilson, R. S., Bienias, J. L., and Bennett, D. A. (2006). Body mass index in older persons is associated with Alzheimer disease pathology. *Neurology*, **67**, 1949-1954.

Bures, J., and Buresova, D. (1963). Cortical spreading depression as a memory disturbing factor. *Journal of Comparative Physiology and Psychology*, **56**, 268-272.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, **2**, 121-167.

Burn, D., Emre, M., McKeith, I., De Deyn, P. P., Aarsland, D., Hsu, C., and Lane, R. (2006). Effects of rivastigmine in patients with and without visual hallucinations in dementia associated with Parkinson's disease. *Movement Disorders*, **21**, 1899-1907.

Butler, J. C. (1982). The closure problem as reflected in discriminant function analysis. *Chemical Geology*, **37**, 367-375.

Cacabelos R, Takeda M, and Winblad B. (1999). The glutamatergic system and neurodegeneration in dementia: Preventive strategies in Alzheimer's disease. *International Journal of Geriatric Psychiatry*, **14**, 3-47.

Caeyenberghs, K., Balschun, D., Roces, D. P., Schwake, M., Saftig, P., and D'Hooge, R. (2006). Multivariate neurocognitive and emotional profile of a mannosidosis murine model for therapy assessment. *Neurobiology of Disease*, **23**, 422-432.

Cai, H., Wang, Y., McCarthy, D., Wen, H., Borchelt, D. R., Price, D. L., and Wong, P. C. (2001). BACE1 is the major β-secretase for generation of Aβ peptides by neurons. *Nature Neuroscience*, **4**, 233-234.

Calhoun, M., Wiederhold, K., Abramowski, D., Phinney, A., Probst, A., Sturchler-Pierrat, C., Staufenbiel, M., Sommer, B., and Jucker, M. (1998). Neuron loss in APP transgenic mice. *Nature*, **395**, 755-756.

Campion, D., Brice, A., Hannequin, D., Tardieu, S., Dubois, B., Calenda, A., Brun, E., Penet, C., Tayot, J., Martinez, M., Bellis, M., Mallet, J., Agid, Y., and Clerget-Darpoux, F. (1995). A large pedigree with early-onset Alzheimer's disease: clinical, neuropathologic, and genetic characterization. *Neurology*, **45**, 80-85.

Canevari, L., Abramov, A. Y., and Duchen, M. R. (2004). Toxicity of amyloid beta peptide: Tales of calcium, mitochondria, and oxidative stress. *Neurochemical Research*, **29**, 637-650.

Carlson, G. A., Borchelt, D. R., Dake, A., Turner, S., Danielson, V., Coffin, J. D., Eckman, C., Meiners, J., Nilsen, S. P., Younkin, S. G., and Hsiao, K. K. (1997). Genetic modification of the phenotypes produced by amyloid precursor protein overexpression in transgenic mice. *Human Molecular Genetics*, **6**, 1951-1959.

Carter, R. J., Lione, L. A., Humby, T., Mangiarini, L., Mahal, A., Bates, G. P., Morton, A. J., and Dunnett, S. B. (1999). Characterization of progressive motor deficits in mice transgenic for the human Huntington's disease mutation. *Journal of Neuroscience*, **19**, 3248-3257.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, **1**, 245-276.

Cerhan, J. H., Ivnik, R. J., Smith, G. E., Tangalos, E. C., Petersen, R. C., and Boeve, B. F. (2002). Diagnostic utility of letter fluency, category fluency, and fluency difference scores in Alzheimer's disease. *Clinical Neuropsychologist*, **16**, 35-42.

Chae, J. H., Jeong, J., Peterson, B. S., Kim, D. J., Bahk, W. M., Jun, T. Y., Kim, S. Y., and Kim, K. S. (2004). Dimensional complexity of the EEG in patients with posttraumatic stress disorder. *Psychiatry Research*, **131**, 79-89.

Chai, C. K. (2007). The genetics of Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, **22**, 37-41.

Chandrasekaran, B., and Jain, A. K. (1979). On balancing decision functions. *Journal of Cybernetics and Information Science*, **2**, 12-15.

Chapman, P. F., White, G. L., Jones, M. W., Cooper-Blacketer, D., Marshall, V. J., Irizarry, M., Younkin, L., Good, M. A., Bliss, T. V., Mayman, B. T., Younkin, S. G., and Hsiao, K. K. (1999). Impaired synaptic plasticity and learning in aged amyloid precursor protein transgenic mice. *Nature Neuroscience*, **2**, 271-276.

Chen, G., Chen, K., Knox, J., Inglis, J., Bernard, A., Martin, S., Justice, A., McConlogue, L., Games, D., Freedman, S., and Morris, R. (2000). A learning deficit related to age and beta-amyloid plaques in a mouse model of Alzheimer's disease. *Nature*, **408**, 975-979.

Cheng, D. H., and Tang, X. C. (1998). Comparative studies of huperzine A, E2020, and tacrine on behavior and cholinesterase activities. *Pharmacology, Biochemistry, and Behavior*, **60**, 377-386.

Cho, Y. H., Delcasso, S., Israel, A., and Jeantet, Y. (2007). A long list visuo-spatial sequential learning in mice. *Behavioral Brain Research*, **179**, 152-158.

Chou, T. (1992). Wake up and smell the coffee: Caffeine, coffee, and the medical consequences. *Western Journal of Medicine*, **157**, 544-553.

Christen, Y. (2000). Oxidative stress and Alzheimer's disease. *American Journal of Clinical Medicine*, **71**, 621S-629S.

Christie, M. A., and Dalrymple-Alford, J. C. (2004). A new rat model of the human serial reaction time task: Contrasting effects of caudate and hippocampal lesions. *Journal of Neuroscience*, **24**, 1034-1039.

Christie, M. A., and Hersch, S. M. (2004). Demonstration of nondeclarative sequence learning in mice: Development of an animal analog of the human serial reaction time task. *Learning and Memory*, **11**, 720-723.

Citron, M., Diehl, T. S., Gordon, G., Biere, A. L., Seubert, P., and Selkoe, D. J. (1996). Evidence that the 42- and 40-amino acid forms of amyloid beta protein are generated from the beta-amyloid precursor protein by different protease activities. *Proceedings of the National Academy of Sciences USA*, **93**, 13170-13175.

Cohen, F. E. (1999). Protein misfolding and prion diseases. *Journal of Molecular Biology*, **293**, 313-320.

Cohen, F. E., and Kelly, J. W. (2003). Therapeutic approaches to protein-misfolding diseases. *Nature*, **426**, 905-909.

Cohen, I. L., Sudhalter, V., Landon-Jimenez, D., and Keogh, M. (1993). A neural network approach to the classification of autism. *Journal of Autism and Developmental Disorders*, **23**, 443-466.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37-46.

Cohen, J. D., Dunbar, K., and McClelland, J. M. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, **97**, 332-361.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., and Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, **386**, 604-608.

Cohen, J. S., Sturdy, C., and Hicks, M. (1996). Intratrial proactive interference in rats' serial alternation performance in the radial maze. *Animal Learning and Behavior*, **24**, 300-309.

Cohen, N. J., Eichenbaum, H., Deacedo, B. S., and Corkin, S. (1985). Different memory systems underlying acquisition of procedural and declarative knowledge. *Annals of the New York Academy of Sciences*, **444**, 54-71.

Cohen-Mansfield, J. (2001). Nonpharmacologic interventions for inappropriate behaviors in dementia: A review, summary, and critique. *American Journal of Geriatric Psychiatry*, **9**, 361-381.

Cohen-Salmon, C., Venault, P., Martin, B., Raffalli-Sebille, M. J., Barkats, M., Clostre, F., Pardon, M. C., Christen, Y., and Chapouthier, G. (1997). Effects of Ginkgo biloba extract (EGb 761) on learning and possible actions on aging. *Journal of Physiology*, **91**, 291-300.

Cole, M. R., and Chappell-Stephenson, R. (2003). Exploring the limits of spatial memory in rats, using very large mazes. *Learning and Behavior*, **31**, 349-368.

Come, J. H., Fraser, P. E., and Lansbury, P. T., Jr. (1993). A kinetic model for amyloid formation in the prion diseases: Importance of seeding. *Proceedings of the National Academy of Sciences USA*, **90**, 5959-5963.

Conover, W. J., and Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *American Statistician*, **35**, 124-129.

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L., and Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, **261**, 921-923.

Cortes, C., and Vapnik, V. (1995). Support vector networks. *Machine Learning*, **20**, 273-297.

Costa, D. A., Cracchiolo, J. R., Bachstetter, A. D., Hughes, T. F., Bales, K. R., Paul, S. M., Mervis, R. F., Arendash, G. W., and Potter, H. (2007). Enrichment improves cognition in AD mice by amyloid-related and unrelated mechanisms. *Neurobiology of Aging*, **28**, 831-844.

Cotman, C. W., Head, E., Muggenburg, B. A., Zicker, S., and Milgram, N. W. (2002). Brain aging in the canine: A diet enriched in antioxidants reduces cognitive dysfunction. *Neurobiology of Aging*, **23**, 809-818.

Cox, P. A., and Balick, M. J. (1994). The ethnobotanical approach to drug discovery. *Scientific American*, **270(6)**, 82-87.

Crabbe, J. C., and Morris, R. G. M. (2004). *Festina lente*: Late-night thoughts on high-throughput screening of mouse behavior. *Nature Neuroscience*, **7**, 1175-1179.

Cracchiolo, J. R., Mori, T., Nazian, S. J., Tan, J., Potter, H., and Arendash, G. W. (2007). Enhanced cognitive activity – over and above social or physical activity – is required to protect Alzheimer's mice against cognitive impairment, reduce Aβ deposition, and increase synaptic immunoreactivity. *Neurobiology of Learning and Memory*, **88**, 277-294.

Crawley, J. N. (1999). Behavioral phenotyping of transgenic and knockout mice: Experimental design and evaluation of general health, sensory functions, motor abilities, and specific behavioral tests. *Brain Research*, **835**, 18-26.

Crawley, J. N., Belknap, J. K., Collins, A., Crabbe, J. C., Frankel, W., Henderson, N., Hitzemann, R. J., Maxson, S. C., Miner, L. L., Silva, A. J., Wehner, J. M., Wynshaw-Boris, A., and Paylor, R. (1997). Behavioral phenotypes of inbred mouse strains: Implications and recommendations for molecular studies. *Psychopharmacology*, **132**, 107-124.

Crisler, S., Morrissey, M. J., Anch, A. M., and Barnett, D. W. (2008). Sleep-stage scoring in the rat using a support vector machine. *Journal of Neuroscience Methods*, **168**, 524-534.

Cristianini, N., and Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines*. Cambridge University Press.

Crum, R. M., Anthony, J. C., Bassett, S. S., and Folstein, M. F. (1993). Population-based norms for the Mini Mental State Examination by age and educational level. *Journal of the American Medical Association*, **269**, 2386-2391.

Crusio, W. E., and Schwegler, H. (2005). Learning spatial orientation tasks in the radial-maze and structural variation in the hippocampus in inbred mice. *Behavioral and Brain Functions*, **1**, 3.

Crusio, W. E., Schwegler, H., and Brust, I. (1995). Covariations between hippocampal mossy fibers and working and reference memory in spatial and non-spatial radial maze tasks in mice. *European Journal of Neuroscience*, **5**, 1413-1420.

Cummings, B. J., Head, E., Afagh, A. J., Milgram, N. W., and Cotman, C. W. (1996a). Beta-amyloid accumulation correlates with cognitive dysfunction in the aged canine. *Neurobiology of Learning and Memory*, **66**, 11-23.

Cummings, B. J., Head, E., Ruehl, W., Milgram, N. W., and Cotman, C. W. (1996b). The canine as an animal model of human aging and dementia. *Neurobiology of Aging*, **17**, 259-268.

Cummings, B. J., Satou, T., Head, E., Milgram, N. W., Cole, G. M., Savage, M. J., Podlisny, M. B., Selkoe, D. J., Siman, R., Greenberg, B. D., and Cotman, C. W. (1996c). Diffuse plaques contain C-terminal Aβ_{42} and not Aβ_{40}: Evidence from cats and dogs. *Neurobiology of Aging*, **17**, 653-659.

Cummings, J. L. (2000). Cognitive and behavioral heterogeneity in Alzheimer's disease: Seeking the neurobiological basis. *Neurobiology of Aging*, **21**, 845-861.

Cummings, J. L., and Zhong, K. (2006). Treatments for behavioral disorders in neurodegenerative diseases: Drug development strategies. *Nature Reviews*, **5**, 64-74.

Cunningham, P., Carney, J., and Jacob, S. (2000). Stability problems with artificial neural networks and the ensemble solution. *Artificial Intelligence in Medicine*, **20**, 217-225.

Cupp, C. J., Jean-Philippe, C., Kerr, W. W., Patil, A. R., and Perez-Camargo, G. (2006). Effect of nutritional interventions on longevity of senior cats. *International Journal of Applied Research in Veterinary Medicine*, **4**, 34-50.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, **2**, 303-314.

Dall'Igna, O. P., Fett, P., Gomes, M. W., Souza, D. O., Cunha, R. A., and Lara, D. R. (2007). Caffeine and adenosine A_{2a} receptor antagonists prevent β-amyloid (25-35)-induced cognitive deficits in mice. *Experimental Neurology*, **203**, 241-245.

DaSilva, K. A., Brown, M. E., Westaway, D., and McLaurin, J. (2006). Immunization with amyloid-β using GM-CSF and IL-4 reduces amyloid burden and alters plaque morphology. *Neurobiology of Disease*, **23**, 433-444.

DeCoteau, W. E., and Kesner, R. P. (2000). A double dissociation between the rat hippocampus and medial caudo-putamen in processing two forms of knowledge. *Behavioral Neuroscience*, **114**, 1096-1108.

DeFigueiredo, R. J. P., Shankle, W. R., Maccato, A., Dick, M. B., Mundkur, P., Mena, I., and Cotman, C. W. (1995). Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain. *Proceedings of the National Academy of Sciences USA*, **92**, 5530-5534.

DeKosky, S., Fitzpatrick, A., Ives, D., Saxton, J., Williamson, J., Lopez, O., Burke, G., Fried, L., Kuller, L., Robbins, J., Tracy, R., Woolard, N., Dunn, L., Kronmal, R., Nahin, R., and Furberg, C. (2006). The Ginkgo Evaluation of Memory (GEM) study: Design and baseline data of a randomized trial of Ginkgo biloba extract in prevention of dementia. *Contemporary Clinical Trials*, **27**, 238-253.

Del-Favero, J., Goossens, D., Van den Bossche, D., and Van Broeckhoven, C. (1999). YAC fragmentation with repetitive and single-copy sequences: detailed physical mapping of the presenilin 1 gene on chromosome 14. *Gene*, **229**, 193-201.

DeMager, P. P., Penke, B., Walter, R., Harkany, T., and Hartignny, W. (2002). Pathological peptide folding in Alzheimer's disease and other conformational disorders. *Current Medicinal Chemistry*, **9**, 1763-1780.

Den Dunnen, W. F. A., Brouwer, W. H., Bijlard, E., Kamphuis, J., van Linschoten, K., Eggens-Meijer, E., and Holstege, G. (2008). No disease in the brain of a 115-year-old woman. *Neurobiology of Aging*, **29**, 1127-1132.

De Quervain, D. J., Henke, K., Aerni, A., Coluccia, D., Wollmer, M. A., Hock, C., Nitsch, R. M., and Papassotiropoulos, A. (2003). A functional genetic variation of the 5-HT2a receptor affects human memory. *Nature Neuroscience*, **6**, 1141-1142.

De Quervain, D. J., and Papassotiropoulos, A. (2006). Identification of a genetic cluster influencing memory performance and hippocampal activity in humans. *Proceedings of the National Academy of Sciences USA*, **103**, 4270-4274.

Dere, E., Huston, J. P., De Souza Silva, M. A. (2005). Episodic-like memory in mice: Simultaneous assessment of object, place and temporal order memory. *Brain Research Protocols*, **16**, 10-19.

Devanand, D. P., Jacobs, D. M., Tang, M. X., Del Castillo-Castenada, C., Sano, M., Marder, K., Bell, K., Bylsma, F. W., Brandt, J., Albert, M., and Stern, Y. (1997). The course of psychopathologic features in mild to moderate Alzheimer's disease. *Archives of General Psychiatry*, **54**, 257-263.

D'Hooge, R., and De Deyn, P. P. (2001). Applications of the Morris water maze in the study of learning and memory. *Brain Research Reviews*, **36**, 60-90.

Dierckx, E., Engelborghs, S., DeRaedt, R., DeDeyn, P. P., and Ponjaert-Kristoffersen, I. (2007). Differentiation between mild cognitive impairment, Alzheimer's disease and depression by means of cued recall. *Psychological Medicine*, **37**, 747-755.

Dietterich, T. (1997). Machine learning research: Four current directions. *AI Magazine*, **18**, 97-136.

Dodart, J.-C., Mathis, C., Bales, K. R., and Paul, S. M. (2002). Does my mouse have Alzheimer's disease? *Genes, Brain and Behavior*, **1**, 142-155.

Dodart, J., Mathis, C., Saura, J., Bales, K., Paul, S., and Ungerer, A. (2000). Neuroanatomical abnormalities in behaviorally characterized APP[V717F] transgenic mice. *Neurobiology of Disease*, **7**, 71-85.

Dodart, J., Meziane, H., Mathis, C, Ungerer, A., Bales, K., and Paul, S. (1999). Behavioral disturbances in transgenic mice overexpressing the V717F beta-amyloid precursor protein. *Behavioral Neuroscience*, **113**, 982-990.

Dolinoy, D. C., Weidman, J. R., and Jirtle, R. L. (2007). Epigenetic gene regulation: Linking early developmental environment to adult disease. *Reproductive Toxicology*, **23**, 297-307.

Donovan, M. H., Yazdani, U., Norris, R. D., Games, D., German, D. C., and Eisch, A. J. (2006). Decreased adult hippocampal neurogenesis in the PDAPP mouse model of Alzheimer's disease. *Journal of Comparative Neurology*, **495**, 70-83.

Dougall, N. J., Bruggink, S., and Ebmeier, K. P. (2004). Systematic review of the diagnostic accuracy of 99mTc-HMPAO-SPECT in dementia. *American Journal of Geriatric Psychiatry*, **12**, 554-570.

Doyon, B. (1992). On the existence and role of chaotic processes in the nervous system. *Acta Biotheoretica*, **40**, 113-119.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P. J., and Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS-ADRDA criteria. *Lancet Neurology*, **6**, 734-746.

Duin, R. P. W. (1996). A note on comparing classifiers. *Pattern Recognition Letters*, **17**, 529-536.

Dunkin, J. J., Leuchter, A. F., Newton, T. F., and Cook, I. A. (1994). Reduced EEG coherence in dementia: State or trait marker? *Biological Psychiatry*, **35**, 870-879.

Ebert, U., Grossmann, M., Oertel, R., Gramatte, T., and Kirch, W. (2001). Pharmacokinetic-pharmacodynamic modeling of the electroencephalogram effects of scopolamine in healthy volunteers. *Journal of Clinical Pharmacology*, **41**, 51-60.

Ebert, U., and Kirch, W. (1998). Scopolamine model of dementia: Electroencephalogram findings and cognitive performance. *European Journal of Clinical Investigation*, **28**, 944-949.

Eckelinen, M., Ngandu, T., Tuomilehto, J., Soininen, H., and Kivipelto, M. (2008). Midlife coffee and tea drinking and the risk of late-life dementia: A population-based CAIDE study. *Journal of Alzheimer's Disease*, **16**, xxx-xxx.

Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association*, **78**, 316-331.

Egan, M. F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., Zaitsev, E., Gold, B., Goldman, D., Dean, M., Lu, B., and Weinberger, D. R. (2003). The BDNF Val66Met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*, **112**, 257-269.

Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. *Behavioral Brain Research*, **103**, 123-133.

Eichenbaum, H., Stewart, C., and Morris, R. G. M. (1990). Hippocampal representation in place learning. *Journal of Neuroscience*, **10**, 3531-3542.

El-Naqa, I., Yongyi, Y., Wernick, M. N., Galatsanos, N. P., and Nishikawa, R. M. (2002). A support vector machine approach for detection of microcalcifications. *IEEE Transactions on Medical Imaging*, **21**, 1552-1563.

Emilien , G., Maloteaux, J.-M., Beyreuther, K., and Masters, C. L. (2000). Alzheimer disease: Mouse models pave the way for therapeutic opportunities. *Archives of Neurology*, **57**, 176-181.

Emre, M., Aarsland, D., Albanese, A., Byrne, E. J., Deuschl, G., De Deyn, P. P., Durif, F., Kulisevsky, J., van Laar, T., Lees, A., Poewe, W., Robillard, A., Rosa, M. M., Wolters, E., Ouarg, P., Tekin, S., and Lane, R. (2004). Rivastigmine for dementia associated with Parkinson's disease. *New England Journal of Medicine*, **351**, 2509-2518.

Eriksen, J. L., Sagi, S. A., Smith, T. E., Weggen, S., Das, P., McLendon, D. C., Ozols, V. V., Jessing, K. W., Zavitz, K. H., Koo, E. H., and Golde, T. E. (2003). NSAIDs and enantiomers of flurbiprofen target gamma-secretase and lower Abeta42 in vivo. *Journal of Clinical Investigation*, **112**, 440-449.

Ethell, D. W., Shippy, D., Cao, C., Cracchiolo, J. R., Runfeldt, M., Blake, B., and Arendash, G. W. (2006). Aβ-specific T-cells reverse cognitive decline and synaptic loss in Alzheimer's mice. *Neurobiology of Disease*, **23**, 351-361.

Evans, D. A., Funkenstein, H. H., Albert, M. S., Scherr, P. A., Cook, N. R., Chown, M. J., Hebert, L. E., Hennekens, C. H., and Taylor, J. O. (1989). Prevalence of Alzheimer's disease in a community population of older persons. Higher than previously reported. *Journal of the American Medical Association*, **262**, 2551-2556.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, **3**, 272-299.

Farias, S. T., Harrell, E., Neumann, C., and Houtz, A. (2003). The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: Ecological validity of neuropsychological tests. *Archives of Clinical Neuropsychology*, **18**, 655-672.

Farina, E., Fioravanti, R., Chiavari, L., Imbornone, E., Alberoni, M., Pomanti, S., Pinardi, G., Pignatti, R., and Mariani, C. (2002). Comparing two programs of cognitive training in Alzheimer's disease: A pilot study. *Acta Neurologica Scandinavica*, **105**, 365-371.

Farlow, M., Gracon, S. I., Hershey, L. A., Lewis, K. W., Sadowsky, C. H., Dolan-Ureno, J. (1992). A controlled trial of tacrine in Alzheimer's disease. *Journal of the American Medical Association*, **268**, 2523-2529.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, **27**, 861-874.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996a). From data mining to knowledge discovery in databases. *AI Magazine*, **Fall**, 37-54.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996b). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, **39(11)**, 27-34.

Feldmann, H., Gauthier, S., Hecker, J., Vellas, B., Subbiah, P. and Whalen, E. (2001). A 24-week, randomized, double-blind study of donepezil in moderate to severe Alzheimer's disease. *Neurology*, **57**, 613-620.

Fell, J., Fernandez, G., and Elger, C. E. (2003). More than synchrony: EEG chaoticity may be necessary for conscious brain functioning. *Medical Hypotheses*, **61**, 158-160.

Fernandes, C., Gonzalez, M. I., Wilson, C. A., and File, S. E. (1999). Factor analysis shows that female rat behavior is characterized primarily by activity, male rats are driven by sex and anxiety. *Pharmacology, Biochemistry, and Behavior*, **64**, 731-738.

Fillenbaum, G. G., Heyman, A., Wilkinson, W. E., and Haynes, C. S. (1987). Comparison of two screening tests in Alzheimer's disease. *Archives of Neurology*, **44**, 924-927.

Fisher, N. J., Rourke, B. P., and Bieliauskas, L. A. (1999). Neuropsychological subgroups of patients with Alzheimer's disease: An examination of the first 10 years of CERAD data. *Journal of Clinical and Experimental Neuropsychology*, **21**, 488-518.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, **7**, 179-188.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, **76**, 378-382.

Fleminger, S., Oliver, D. L., Lovestone, S., Rabe-Hesketh, S., and Giora, A. (2003). Head injury as a risk factor for Alzheimer's disease: The evidence 10 years on. A partial replication. *Journal of Neurology, Neurosurgery and Psychiatry*, **74**, 857-862.

Flint, J. (1999). The genetic basis of cognition. *Brain*, **122**, 2015-2031.

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

Forbes, K. E., Shanks, M. F., and Venneri, A. (2004). The evolution of dysgraphia in Alzheimer's disease. *Brain Research Bulletin*, **63**, 19-24.

Forstl, H., and Kurtz, A. (1999). Clinical features of Alzheimer's disease. *European Archives of Psychiatry and Clinical Neurosciences*, **249**, 288-290.

Foster, D. J., Morris, R. G. M., and Dayan, P. (2000). A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus*, **10**, 1-16.

Fox, K. R., Stathi, A., McKenna, J., and Davis, M. G. (2007). Physical activity and mental well-being in older people participating in the Better Ageing Project. *European Journal of Applied Physiology*, **100**, 591-602.

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suner, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y.-Z., Plass, C., and Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences USA*, **102**, 10604-10609.

Fratiglioni, L., Paillard-Borg, S., and Winblad, B. (2004). An active and socially integrated lifestyle in late life might protect against dementia. *Lancet Neurology*, **3**, 343-353.

Fredholm, B., Battig, K., Holmen, J., Nehlig, A., and Zvartau, E. (1999). Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews*, **51**, 83-133.

French, B. M., Dawson, M. R. W., and Dobbs, A. R. (1997). Classification and staging of dementia of the Alzheimer type: A comparison between neural networks and linear discriminant analysis. *Archives of Neurology*, **54**, 1001-1009.

Frohlich, H., Hoenselaar, A., Eichner, J., Rosenbrock, H., Birk, G., and Zell, A. (2008). Automated classification of the behavior of rats in the forced swimming test with support vector machines. *Neural Networks*, **21**, 92-101.

Fuld, P. A. (1981). *Fuld Object-Memory Evaluation.* WoodDale, IL: Stoelting Co.

Gaasterland, T., and Bekiranov, S. (2000). Making the most of microarray data. *Nature Genetics*, **24**, 204-206.

Gaito, J. (1958). The single Latin square design in psychological research. *Psychometrika*, **23**, 369-378.

Gaito, J. (1961). Repeated measurement designs and counterbalancing. *Psychological Bulletin*, **58**, 46-54.

Gallagher, M., and Rapp, P. R. (1997). The use of animal models to study the effects of aging on cognition. *Annuals Reviews in Psychology*, **48**, 339-370.

Gallez, D., and Babloyantz, A. (1991). Predictability of human EEG: A dynamical approach. *Biological Cybernetics*, **64**, 381-391.

Gallistel, C. R., Fairhurst, S., and Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences USA*, **101**, 13124-13131.

Galsworthy, M., Paya-Cano, J., Liu, L., Monleon, S., Gregoryan, G., Fernandes, C., Schalkwyk, L., and Plomin, R. (2005). Assessing reliability, heritability and general cognitive ability in a battery of cognitive tasks for laboratory mice. *Behavior Genetics*, **35**, 675-692.

Games, D., Adams, D., Alessandrini, R., Barbour, R., Berthelette, P., Blackwell, C., Carr, T., Clemens, J., Donaldson, T., Gillespie, F., Guido, T., Hagopian, S., Johnson-Wood, K., Khan, K., Lee, M., Leibowitz, P., Lieberburg, I., Little, S., Masliah, E., McConlogue, L., Montoya-Zavala, M., Mucke, L., Paganini, L., Penniman, E., Power, M., Schenk, D., Seubert, P., Snyder, B., Soriano, F., Tan, H., Vitale, J., Wadsworth, S., Wolozin, B., and Zhao, J. (1995). Alzheimer-type neuropathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *Nature*, **373**, 523-527.

Gamst, G., Meyers, L. S., and Guardino, A. J. (2008). *Analysis of Variance Designs: A Conceptual and Computational Approach with SPSS and SAS*. New York: Cambridge University Press. Pp. 420-450.

Ganguli, M., Chandra, V., Kamboh, M. I., Johnston, J. M., Dodge, H. H., Thelma, B. K., Juyal, R. C., Pandav, R., Belle, S. H., and DeKosky, S. T. (2000). Apolipoprotein E polymorphism and Alzheimer disease: The Indo-US Cross-National Dementia Study. *Archives of Neurology*, **57**, 824-830.

Garcia-Perez, E., Violante, A., and Cervantes-Perez, F. (1998). Using neural networks for differential diagnosis of Alzheimer disease and vascular dementia. *Expert Systems with Applications*, **14**, 219-225.

Garrard, P., Maloney, L. M., Hodges, J. R., and Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain*, **128**, 250-260.

Gauthier, S. (2001). Alzheimer's disease: Current and future therapeutic perspectives. *Progress in Neuropsychopharmacology and Biological Psychiatry*, **25**, 73-89.

Gauthier, S., Vellas, B., Farlow, M., and Burn, D. (2006). An aggressive course of disease in dementia. *Alzheimer's and Dementia*, **2**, 210217.

Gearing, M., Rebeck, G. W., Hyman, B. T., Tigges, J., and Mirra, S. S. (1994). Neuropathology and apolipoprotein E profile of aged chimpanzees: Implications for Alzheimer disease. *Proceedings of the National Academy of Sciences USA*, **91**, 9382-9386.

Gearing, M., Tigges, J., Mori, H., and Mirra, S. S. (1997). β-amyloid (Aβ) deposition in the brains of aged orangutans. *Neurobiology of Aging*, **18**, 139-146.

Gee, J., Ding, L., Xie, Z., Lin, M., DeVita, C. and Grossman, M. (2003). Alzheimer's disease and frontotemporal dementia exhibit distinct atrophy-behavior correlates: a computer-assisted imaging study. *Academic Radiology*, **10**, 1392-1409.

Gehan, E. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*, **52**, 203-223.

Gerlai, R. (1996). Gene-targeting studies of mammalian behavior: Is it the mutation or the background genotype? *Trends in Neuroscience*, **19**, 177-186.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., and Rush, G. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, **40**, 79-91.

Gevins, A., Smith, M. E., McEvoy, L., and Yu, D. (1997). High resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, **7**, 374-385.

Giaquinto, S., and Nolfe, G. (1986). The EEG in the normal elderly: A contribution to the interpretation of aging and dementia. *Electroencephalography and Clinical Neurophysiology*, **63**, 540-546.

Giarratano, J. C., and Riley, G. D. (2004). *Expert Systems: Principles and Programming*, 4th Ed. Florence, KY: Course Technology.

Gimenez-Llort, L., Blazquez, G., Canete, T., Johansson, B., Oddo, S., Tobena, A., LaFerla, F. M., and Fernandez-Teruel, A. (2007). Modeling behavioral and neuronal symptoms of Alzheimer's disease in mice: A role for intraneuronal amyloid. *Neuroscience and Biobehavioral Reviews*, **31**, 125-147.

Glass, L. (2001). Synchronization and rhythmic processes in physiology. *Nature*, **410**, 277-284.

Glenner, G. G., and Wong, C. W. (1984a). Alzheimer's disease: Initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochemical and Biophysical Research Communications*, **120**, 885-890.

Glenner, G. G., and Wong, C. W. (1984b). Alzheimer's disease and Down's syndrome: Sharing of a a unique cerebrovascular amyloid fibril protein. *Biochemical and Biophysical Research Communications*, **122**, 1131-1135.

Godkar, P. B., Gordon, R. K., Ravindran, A., and Doctor, B. P. (2004). *Celastrus paniculatus* seed water soluble extracts protect against glutamate toxicity in neuronal cultures from rat forebrain. *Journal of Ethnopharmacology*, **93**, 213-219.

Goldgaber, D., Lerman, M. I., McBride, O. W., Saffiotti, U., and Gajdusek, D. C. (1987). Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. *Science*, **235**, 877-880.

Golub, T., Slonim, D., Tamayo, P., Huard, C., Gaansenbeek, M., Mesirov, J. H. H. C., Loh, M., Downing, J., Caligiuri, M., Bloomfield, C., and Lander, E. (1999). Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531-537.

Gomez, R. G., and White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology*, **21**, 771-775.

Gonzalez, F. J. (2002). The peroxisome proliferator-activated receptor α (PPARα): Role in hepatocarcinogenesis. *Molecular and Cellular Endocrinology*, **193**, 71-79.

Good, M. A., Hale, G., and Staal, V. (2007). Impaired "episodic-like" object memory in adult APPswe transgenic mice. *Behavioral Neuroscience*, **121**, 443-448.

Gordon, M. N., Holcomb, L. A., Jantzen, P. T., DiCarlo, G., Wilcock, D., Boyett, K. W., Connor, K., Melachrino, J., OCallaghan, J. P., and Morgan, D. (2002). Time course of the development of Alzheimer-like pathology in the doubly transgenic PS1+APP mouse. *Experimental Neurology*, **173**, 183-195.

Gordon, M, King, D., Diamond, D., Jantzen, P., Boyett, K., Hope, C., Hatcher, J., DiCarlo, G., Gottschalk, P., Morgan, D. and Arendash, G. (2001). Correlation between cognitive deficits and Ab deposits in transgenic APP+PS1 mice. *Neurobiology of Aging*, **22**, 377-386.

Goutte, C. (1997). Note on free lunches and cross-validation. *Neural Computation*, **9**, 1211-1215.

Graeber, M. B. (1999). No man alone: The rediscovery of Alois Alzheimer's original cases. *Brain Pathology*, **9**, 237-240.

Graeber, M. B., Kosel, S., Grasbon-Frodi, F., Moller, H. J., and Mahraein, P. (1998). Histopathology and ApoE genotype of the first Alzheimer disease patient, Auguste D. *Neurogenetics*, **1**, 223-228.

Graf, P., and Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 501-518.

Graf, P., and Schacter, D. L. (1987). Selective effects of interference on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 45-53.

Grassberger, P., and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D*, **9**, 183-208.

Graziano, A., Petrosini, L., and Bartoletti, A. (2003). Automatic recognition of exploratory strategies in the Morris water maze. *Journal of Neuroscience Methods*, **130**, 33-44.

Green, R. C., Schneider, L. S., Zavitz, K. H., Amato, D. A., Beelen, A. P., Swabb, E. A., and the Tarenflurbil Phase 3 Study Group. (2008). Safety and efficacy of tarenflurbil in subjects with mild Alzheimer's disease: Results from an 18-month multi-center phase 3 trial. (Presentation). *Alzheimer's Association International Conference on Alzheimer's Disease*, July 26-31, Chicago, IL.

Grice, J. W. (2001). A comparison of factor scores under conditions of factor obliquity. *Psychological Bulletin*, **6**, 67-83.

Griffith, J. S. (1967). Self-replication and scrapie. *Nature*, **215**, 1043-1044.

Gron, G., and Riepe, M. W. (2004). Neural basis for the cognitive continuum in episodic memory from health to Alzheimer's disease. *American Journal of Geriatric Psychiatry*, **12**, 648-652.

Grossi, E., Buscema, M. P., Snowdon, D., and Antuono, P. (2007). Neuropathological findings processed by artificial neural networks (ANNs) can perfectly distinguish Alzheimer's patients from controls in the Nun Study. *BioMed Central Neurology*, **7**, 15.

Grundke-Iqbal, I., Iqbal, K., Tung, Y. C., Quinlan, M., Wisniewski, H. M., and Binder, L. I. (1986). Abnormal phophorylation of the microtubule-associated protein τ (tau) in Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences USA*, **83**, 4913-4917.

Guadagnoli, E., and Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, **103**, 265-275.

Gunn-Moore, D., McVee, J., Bradshaw, J. M., Pearson, G. R., Head, E., and Gunn-Moore, F. J. (2006). Ageing changes in cat brains demonstrated by beta-amyloid and AT8-immunoreactive phosphorylated tau deposits. *Journal of Feline Medicine and Surgery*, **8**, 234-242.

Gureje O., Ogunniyi A., Baiyewu O., Price B., Unverzagt F. W., Evans R. M., Smith-Gamble V., Lane K. A., Gao S, Hall K. S., Hendrie H. C., and Murrell J. R. (2006). APOE-ϵ4 is not associated with Alzheimer's disease in elderly Nigerians. *Annals of Neurology*, **59**, 182-185.

Gustafson, L., Brun, A., Englund, E., Hagnell, O., Nilsson, K., Stensmyr, M., Ohlin, A. K., and Abrahamson, M. (1998). A 50-year perspective of a family with chromosome-14-linked Alzheimer's disease. *Human Genetics*, **102**, 253-257.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika*, **19**, 149-161.

Haas, C., and DeStrooper, B. (1999). The presenilins in Alzheimer's disease: Proteoloysis holds the key. *Science*, **286**, 916-919.

Habeck, C., Foster, N. L., Perneczky, R., Kurz, A., Alexopoulos, P., Koeppe, R. A., Drzezga, A., and Stern, Y. (2008). Multivariate and univariate neuroimaging biomarkers of Alzheimer's disease. *NeuroImage*, **40**, 1503-1515.

Hall, C. S. (1934). Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *Journal of Comparative and Physiological Psychology*, **18**, 385-403.

Hall, C. S. (1936). Emotional behavior in the rat. III. The relationship between emotionality and ambulatory activity. *Journal of Comparative Psychology*, **22**, 345-452.

Hampton, R. R., and Schwartz, B. L. (2004). Episodic memory in nonhumans: What, and where, is when? *Current Opinion in Neurobiology*, **14**, 192-197.

Han, M.-E., Park, K.-H., Baek, S.-Y., Kim, B.-S., Kim, J.-B., Kim, H.-J., and Oh, S.-O. (2007). Inhibitory effects of caffeine on hippocampal neurogenesis and function. *Biochemical and Biophysical Research Communications*, **356**, 976-980.

Hanley, J. A., and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29-36.

Hardy, J. (1997). The Alzheimer family of diseases: Many etiologies, one pathogenesis? *Proceedings of the National Academy of Sciences USA*, **94**, 2095-2097.

Hardy, J. (2006). Has the amyloid cascade hypothesis for Alzheimer's disease been proved? *Current Alzheimers Research*, **3**, 71-73.

Hart, D. J., Craig, D., Compton, S. A., Critchlow, S., Kerrigan, B. M., McIlroy, S. P., and Passmore, A. P. (2003). A retrospective study of the behavioral and psychological symptoms of mid and late phase Alzheimer's disease. *International Journal of Geriatric Psychiatry*, **18**, 1037-1042.

Hartman, R. E., Shah, A., Fagan, A. M., Schwetye, K. E., Parsadanian, M., Schulman, R. N., Finn, M. B., and Holtzman, D. M. (2006). Pomegranate juice decreases amyloid load and improves behavior in a mouse model of Alzheimer's disease. *Neurobiology of Disease*, **24**, 506-515.

Hashimoto, M., Rockenstein, E., Crews, L., and Masliah, E. (2003). Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases. *Neuromolecular Medicine*, **4**, 21-36.

Hasselmo, M. E., and Wyble, B. P. (1997). Free recall and recognition in a network model of the hippocampus. *Behavioral Brain Research*, **89**, 1-34.

Hayashi, H., Kimura, N., Yamaguchi, H., Hasegawa, K., Yokoseki, T., Shibata, M., Yamamoto, N., Michikawa, M., Yoshikawa, Y., Terao, K., Matsuzaki, K., Lemere, C. A., Selkoe, D. J., Naiki, H., and Yanagisawa, K. (2004). A seed for Alzheimer amyloid in the brain. *Journal of Neuroscience*, **24**, 4894-4902.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation, 2nd Ed.* Upper Saddle River, NJ: Prentice Hall.

Head, E., Moffat, K., Das, P., Sarsoza, F., Poon, W. W., Landsberg, G., Cotman, C. W., and Murphy, M. P. (2005). Beta-amyloid deposition and tau phosphorylation in clinically characterized aged cats. *Neurobiology of Aging*, **26**, 749-763.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory.* New York: Wiley.

Hebb, D. O., and Williams, K. A. (1946). A method for rating animal intelligence. *Journal of General Psychology*, **34**, 59-65.

Hebert, L. E., Scherr, P. A., Bienias, J. L., Bennett, D. A., and Evans, D. A. (2003). Alzheimer's disease in the U. S. Population: Prevalence estimates using the 2000 census. *Archives of Neurology*, **60**, 1119-1122.

Heeren, D. J., and Cools, A. R. (2000). Classifying postures of freely moving rodents with the help of Fourier descriptors and a neural network. *Behavioral Research Methods, Instrumentation, and Computers*, **32**, 56-62.

Hepler, D. J., Wenk, G. L., Cribbs, B. L., Olton, D. S., and Coyle, J. T. (1985). Memory impairments following basal forebrain lesions. *Brain Research*, **346**, 8-14.

Holcomb, L. A., Dhanasekaran, M., Hitt, A. R., Young, K. A., Riggs, M., and Manyam, B. V. (2006). *Bacopa monniera* extract reduces amyloid levels in PSAPP mice. *Journal of Alzheimer's Disease*, **9**, 243-251.

Holcomb, L. A., Gordon, M. N., Jantzen, P., Hsiao, K., Duff, K., and Morgan, D. (1999). Behavioral changes in transgenic mice expressing both amyloid precursor protein and presenilin-1 mutations: Lack of association with amyloid deposits. *Behavior Genetics*, **29**, 177-185.

Holcomb, L., Gordon, M., McGowan, E., Yu, X., Benkovic, S., Jantzen, P., Wright, K., Saad, I., Mueller, R., Morgan, D., Sanders, S., Zehr, C., OCampo, K., Hardy, J., Prada, C., Eckman, C., Younkin, S., Hsaio, K., and Duff, K. (1998). Accelerated Alzheimer-type phenotype in transgenic mice carrying both mutant amyloid precursor protein and presenilin 1 transgenes. *Nature Medicine*, **4**, 97-100.

Holmes, C., Wilkinson, D., Dean, C., Vethanayagam, S., Olivieri, S., Langley, A., Pandita-Gunawardena, H. D., Hogg, F., Clare, C., and Damms, J. (2004). The efficacy of donezepil in the treatment of neuropsychiatric symptoms in Alzheimer's disease. *Neurology*, **63**, 214-219.

Holmstrom, L., Koistinen, P., Laaksonen, J., and Oja, E. (1997). Neural and statistical classifiers: Taxonomy and two case studies. *IEEE Transactions on Neural Networks*, **8**, 5-17.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, **79**, 2554-2558.

Hsiao, K., Chapman, P., Nilsen, S., Eckman, C., Harigaya, Y.., Younkin, S., Yang, F., and Cole, G. (1996). Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. *Science*, **274**, 99-102.

Hughes, J. R., Shanmugham, S., Wetzel, L. C., Bellur, S., and Hughes, C. A. (1989). The relationship between EEG changes and cognitive functions in dementia: A study in a VA population. *Clinical Electroencephalography*, **20**, 77-85.

Hyde, L. A., Hoplight, B. J., and Denenberg, V. H. (1998). Water version of the radial-arm maze: Learning in three inbred strains of mice. *Brain Research*, **785**, 236-244.

Ikeuchi, T., and Sisodia, S. S. (2002). Cell-free generation of the notch1 intracellular domain (NICD) and APP-CTfgamma: evidence for distinct intramembranous "gamma-secretase" activities. *Neuromolecular Medicine*, **1**, 43-54.

Ikonomovic, M. D., Klunk, W. E., Abrahamson, E. E., Mathis, C. A., Price, J. C., Tsopelas, N. D., Lopresti, B. J., Ziolko, S., Bi, W., Paljug, W. R., Debnath, M. L., Hope, C. E., Isanski, B. A., Hamilton, R. L., and DeKosky, S. T. (2008). Post-mortem correlates of *in vivo* PiB-PET amyloid imaging in a typical case of Alzheimer's disease. *Brain*, **131**, 1630-1645.

Irizarry, M., McNamara, M., Fedorchak, K., Hsiao, K., and Hyman, B. (1997). APPsw transgenic mice develop age-related Aβ deposits and neuropil abnormalities, but no neuronal loss in CA1. *Journal of Neuropathology and Experimental Neurology*, **56**, 965-973.

Irwin, S. (1968). Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiological state of the mouse. *Psychopharmacologia*, **13**, 222-257.

Izquierdo, J. A., Costas, S. M., Justel, E. A., and Rabiller, G. (1979). Effect of caffeine on the memory of the mouse. *Psychopharmacology*, **61**, 29-30.

Jack, C. R., Marjanska, M., Wengenack, T. M., Reyes, D. A., Curran, G. L., Lin, J., Preboske, G. M., Poduslo, J. F., and Garwood, M. (2007). Magnetic resonance imaging of Alzheimer's pathology in the brains of living transgenic mice: A new tool in Alzheimer's disease research. *The Neuroscientist*, **13**, 38-48.

Jack, C. R., Petersen, R. C., Xu, Y. C., Waring, S. C., OBrien, P. C., Tangalos, E. G., Smith, G. E., Ivnik, R. J., and Kokmen, E. (1997). Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology*, **49**, 786-794.

Jacobs, D., Sano, M., Marder, K., Bell, K., Bylsma, F., Lafleche, G., Albert, M., Brandt, J., and Stern, Y. (1994). Age at onset of Alzheimer's disease: Relation to pattern of cognitive dysfunction and rate of decline. *Neurology*, **44**, 1215-1220.

Jaenisch, R. (1988). Transgenic animals. *Science*, **240**, 1468-1472.

Jankowsky, J. L., Fadale, D. J., Anderson, J., Xu, G. M., Gonzales, V., Jenkins, N. A., Copeland, N. G., Lee, M. K., Younkin, L. H., Wagner, S. L., Younkin, S. G., and Borchelt, D. R. (2004). Mutant presenilins specifically elevate the levels of the 42 residue beta-amyloid peptide *in vivo*: Evidence for augmentation of a 42-specific gamma secretase. *Human Molecular Genetics*, **13**, 159-170.

Jankowsky, J. L., Younkin, L. H., Gonzales, V., Fadale, D. J., Slunt, H. H., Lester, H. A., Younkin, S. G., and Borchelt, D. R. (2007). Rodent Aβ modulates the solubility and distribution of amyloid deposits in transgenic mice. *Journal of Biological Chemistry*, **282**, 22707-22720.

Janus, C., and Westaway, D. (2001). Transgenic mouse models of Alzheimer's disease. *Physiology and Behavior*, **73**, 873-886.

Jarrett, J. T., and Lansbury, P. T. (1993). Seeding "one dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell*, **73**, 1055-1058.

Jausovec, N., and Jausovec, K. (2000). Differences in event-related and induced brain oscillations in the theta and alpha frequency bands related to human intelligence. *Neuroscience Letters*, **293**, 191-194.

Jee, S. W., Cho, J. S., Kim, C. K., Hwang, D. Y., Shim, S. B., Lee, S. H., Sin, J. S., Kim, Y. S., Park, J. H., Lee, S. H., Choi, S. Y., and Kim, Y. K. (2007). Analysis of differentially expressed genes in early- and late-stage APPsw-transgenic and normal mice using cDNA microarray. *International Journal of Molecular Medicine*, **19**, 461-468.

Jee, S. W., Cho, J. S., Oh, J. H., Shim, S. B., Hwang, D. Y., Lee, S. H., Song, Y. S., Lee, S. H., and Kim, Y. (2005). cDNA microarray-based analysis of differentially expressed genes in transgenic brains expressing NSE-controlled APPsw. *International Journal of Molecular Medicine*, **16**, 547-552.

Jelles, B., van Birgelen, J. H., Slaets, J. P., Hekster, R. E., Jonkman, E. J., and Stam, C. J. (1999). Decrease of non-linear structure in the EEG of Alzheimer patients compared to healthy controls. *Clinical Neurophysiology*, **110**, 1159-1167.

Jensen, M. T., Mottin, M. D., Cracchiolo, J. R., Leighty, R. E., and Arendash, G. W. (2005). Lifelong immunization with human β-amyloid (1-42) protects Alzheimer's transgenic mice against cognitive impairment throughout aging. *Neuroscience*, **130**, 667-684.

Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clinical Neurophysiology*, **115**, 1490-1505.

Jeong, J., Gore, J. C., and Peterson, B. S. (2001a). Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clinical Neurophysiology*, **112**, 827-835.

Jeong, J., Kim, S. Y., and Han, S.-H. (1998). Non-linear dynamical analysis of the EEG in Alzheimer's disease with optimal embedding dimension. *Electroencephalography and Clinical Neurophysiology*, **106**, 220-228.

Jeong, J., Kim, D. J., Kim, S. Y., Chae, J. H., Go, H. J., and Kim, K. S. (2001b). Effect of total sleep deprivation on the dimensional complexity of the waking EEG. *Sleep*, **24**, 197-202.

Jitsumori, M., Siemann, M., Lehr, M., and Delius, J. D. (2002). A new approach to the formation of equivalence classes in pigeons. *Journal of the Experimental Analysis of Behavior*, **78**, 397-408.

Johnson-Kozlow, M., Kritz-Silverstein, D., Barrett-Connor, E., and Morton, D. (2002). Coffee consumption and cognitive function among older adults. *American Journal of Epidemiology*, **156**, 842-850.

Johnson-Wood, K., Lee, M., Motter, R., Hu, K., Gordon, G., Barbour, R., Khan, K., Gordon, M., Tan, H., Games, D., Lieberburg, I., Schenk, D., Seubert, P., and McConlogue, L. (1997). Amyloid precursor protein processing and Aβ_{42} deposition in a transgenic mouse model of Alzheimer's disease. *Proceedings of the National Academy of Sciences USA*, **94**, 1550-1555.

Jones, R. N., and Gallo, J. J. (2000). Dimensions of the Mini-Mental State Examination among community dwelling older adults. *Psychological Medicine*, **30**, 605-618.

Jones, S. N., and Ayers, C. R. (2006). Psychometric properties and factor structure of an expanded CERAD neuropsychological battery in an elderly VA sample. *Archives of Clinical Neuropsychology*, **21**, 359-365.

Jorgensen, P., Bus, C., Pallisgaard, N., Bryder, M., and Jorgensen, A. L. (1996). Familial Alzheimer's disease co-segregates with a Met-146-Ile substitution in presenilin-1. *Clinical Genetics*, **50**, 281-286.

Joseph, J. A., Shukitt-Hale, B., and Casadesus, G. (2005). Reversing the deleterious effects of aging on neuronal communication and behavior: Beneficial properties of fruit polyphenolic compounds. *American Journal of Clinical Nutrition*, **81**, S313-S316.

Joseph, J. A., Shukitt-Hale, B., Denisova, N. A., Prior, R. L., Cao, G., Martin, A., Taglialatela, G., and Bickford, P. C. (1998). Long-term dietary strawberry, spinach, or vitamin E supplementation retards the onset of age-related neuronal signal-transduction and cognitive behavioral deficits. *Journal of Neuroscience*, **18**, 8047-8055.

Jost, M. C., and Grossberg, G. T. (1996). The evolution of psychiatric symptoms in Alzheimer's disease: A natural history study. *Journal of the American Geriatric Society*, **44**, 1078-1081.

Kafkafi, N. (2003). Extending SEE for large-scale phenotyping of mouse open-field behavior. *Behavior Research Methods, Instruments, and Computers*, **35**, 294-301.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, **20**, 141-151.

Kalampokis, A., Kotsavasiloglou, C., Argyrakis, P., and Baloyannis, S. (2003). Robustness in biological neural networks. *Physica A*, **317**, 581-590.

Kandel, E. R. (1979). Small systems of neurons. *Scientific American*, **241(3)**, 66-76.

Kandel, E. R. (2001). The molecular biology of memory storage: A dialogue between genes and synapses. *Science*, **294**, 1030-1038.

Kandel, E. R., and Pittenger, C. (1999). The past, the future and the biology of memory storage. *Philosophical Transactions of the Royal Society of London, Series B*, **354**, 2027-2052.

Kang, J., Lemaire, H. G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K. H., Multhaup, G., Beyreuther, K., and Müller-Hill, B. (1987). The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature*, **325**, 733-736.

Kanne, S. M., Balota, D. A., Storandt, M., McKeel, D. W., Jr., and Morris, J. C. (1998). Relating anatomy to function in Alzheimer's disease: Neuropsychological profiles predict regional neuropathology 5 years later. *Neurology*, **50**, 979-985.

Kanowski, S., Hermann, W. M., Stephan, K., Wierich, W., and Horr, R. (1996). Proof of efficacy of *Ginkgo biloba* special extract EGb 761 in outpatients suffering from mild to moderate primary degenerative dementia of the Alzheimer type or multi-infarct dementia. *Pharmacopsychiatry*, **29**, 47-56.

Kaplan, G. B., Greenblatt, D. J., Ehrenberg, B. L., Goddard, J. E., Cotreau, M. M., Harmatz, B. A., and Shader, R. I. (1997). Dose-dependent pharmacokinetics and psychomotor effects of caffeine in humans. *Journal of Clinical Pharmacology*, **37**, 693-703.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., and Schimmel, H. (1983). Validation of a short orientation-memory-concentration test of cognitive impairment. *American Journal of Psychiatry*, **140**, 734-739.

Kaufer, D., Cummings, J. and Christine, D. (1996). Effect of tacrine on behavioral symptoms in Alzheimer's disease: An open-label study. *Journal of Geriatric Psychiatry and Neurology*, **9**, 1-6.

Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, **300**, 486-489.

Kelly, P., Bondolfi, L., Hunziker, D., Schlecht, H., Carver, K., Maquire, E., Abramowski, D., Wiederhold, K., Sturchler-Pierrat, C., Jucker, M., Bergmann, R., Staufenbiel, M., and Sommer, B. (2003). Progressive age-related impairment of cognitive behavior in APP23 transgenic mice. *Neurobiology of Aging*, **24**, 365-378.

Kelso, S. R., Ganong, A. H., and Brown, T. H. (1986). Hebbian synapses in hippocampus. *Proceedings of the National Academy of Sciences USA*, **83**, 5326-5330.

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

Kiang, Y. M. (2003). A comparative assessment of classification methods. *Decision Support Systems*, **35**, 441-454.

King, D., and Arendash, G. (2002a). Behavioral characterization of the Tg2576 transgenic model of Alzheimer's disease through 19 months. *Physiology & Behavior*, **75**, 627-642.

King, D., and Arendash, G. (2002b). Maintained synaptophysin immunoreactivity in Tg2576 transgenic mice during aging: Correlations with cognitive impairment. *Brain Research*, **926**, 58-68.

King, D. L., Arendash, G. W., Crawford, F., Sterk, T., Menendez, J., and Mullan, M. J. (1999). Progressive and gender-dependent cognitive impairment in the APPsw transgenic mouse model for Alzheimer's disease. *Behavioral Brain Research*, **103**, 145-162.

Kirsch, P., Besthorn, C., Klein, S., Rindfleisch, J., and Olbrich, R. (2000). The dimensional complexity of the EEG during cognitive tasks reflects the impaired information processing in schizophrenic patients. *International Journal of Psychophysiology*, **36**, 237-246.

Kitzawa, M., Yamasaki, T. R., and LaFerla, F. M. (2004). Microglia as a potential bridge between the amyloid beta-peptide and tau. *Annals of the New York Academy of Sciences*, **1035**, 85-103.

Klafki, H. W., Staufenbiel, M., Kornhuber, J., and Wiltfang, J. (2006). Therapeutic approaches to Alzheimer's disease. *Brain* **129**, 2840-2855.

Klunk, W. E., Engler, H., Nordberg, A., Bacskai, B. J., Wang, Y., Price, J. C., Bergstrom, M., Hyman, B. T., Langstrom, B., and Mathis, C. A. (2003). Imaging the pathology of Alzheimer's disease: Amyloid-imaging with positron emission tomography. *Neuroimaging Clinics of North America*, **13**, 781-789.

Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., Bergstrom, M., Savitcheva, I., Huang, G. F., Estrada, S., Ausen, B., Debnath, M. L., Barletta, J., Price, J. C., Sandell, J., Lopresti, B. J., Wall, A., Koivisto, P., Antoni, G., Mathis, C. A., and Langstrom, B. (2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Annals of Neurology*, **55**, 306-319.

Knapp, M. J., Knopman, D. S., Solomon, P. R., Pendlebury, W. W., Davis, C. S., and Gracon, S. I. (1994). A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. *Journal of the American Medical Association*, **271**, 985-991.

Knopman, D., and Nissen, M. J. (1991). Procedural learning is impaired in Huntington's disease: Evidence from the serial reaction time task. *Neuropsychologia*, **29**, 245-254.

Knott, V., Engeland, C., Mohr, E., Mahoney, C., and Ilivitsky, V. (2000). Acute nicotine administration in Alzheimer's disease: An exploratory EEG study. *Neuropsychobiology*, **41**, 210-220.

Koenig, T., Prichep, L., Dierks, T., Hubl, D., Wahlund, L. O., John, E. R., and Jelic, V. (2005). Decreased EEG synchronization in Alzheimer's disease and mild cognitive impairment. *Neurobiology of Aging*, **26**, 165-171.

Kogan, E. A., Korczyn, A. D., Virchovsky, R. G., Klimovizky, S. S., Treves, T. A., and Neufeld, M. Y. (2001). EEG changes during long-term treatment with donepezil in Alzheimer's disease patients. *Journal of Neural Transmission*, **108**, 1167-1173.

Koh, H. C., and Tan, G. (2005). Data mining applications in healthcare. *Journal of Healthcare Information Management*, **19**, 64-72.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 59-69.

Koizumi, K., Nakajima, M., Yuasa, S., Saga, Y., Sakai, T., Kuriyama, T., Shirasawa, T., and Koseki, H. (2001). The role of presenilin 1 during somite segmentation. *Development*, **128**, 1391-1402.

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., and Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiology of Learning and Memory*, **84**, 241-246.

Kolmogorov, A. N. (1957). On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition. *Doklady Akademeii Nauk SSSR*, **114**, 953-956.

Kovacs, D. M., Fausett, H. J., Page, K. J., Kim, T.-W., Moir, R. D., Merriam, D. E., Hollister, R. D., Hallmark, O. G., Mancini, R., Felsenstein, K. M., Hyman, B. T., Tanzi, R. E., Wasco, W. (1996). Alzheimer-associated presenilins 1 and 2: neuronal expression in brain and localization to intracellular membranes in mammalian cells. *Nature Medicine*, **2**, 224-229.

Kowalik-Jankowska, T., Ruta-Dolejsz, M., Wisniewska, K., Lanklewicz, L., and Kozlowski, H. (2002). Possible involvement of copper (II) in Alzheimer disease. *Environmental Health Perspectives*, **110 (Suppl. 5)**, 869-870.

Kowalski, J. W., Gawel, M., Pfeffer, A., and Barcikowska, M. (2001). The diagnostic value of EEG in Alzheimer's disease: Correlation with the severity of mental impairment. *Journal of Clinical Neurophysiology*, **18**, 570-575.

Kramer, A. F., and Erickson, K. I. (2007). Capitalizing on cortical plasticity: Influence of physical activity on cognition and brain function. *Trends in Cognitive Sciences*, **11**, 342-348.

Kristensen, M., and Hansen, T. (2004). Statistical analyses of repeated measures in physiological research: A tutorial. *Advances in Physiology Education*, **28**, 2-14.

Kuncel, N. R., Hezlett, S. A., and Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, **86**, 148-161.

LaFerla, F. M. (2006). An array of genes implicated in Alzheimer's disease. *Neurobiology of Aging*, **27**, 1078-1080.

LaFerla, F. M., and Oddo, S. (2005). Alzheimer's disease: Abeta, tau and synaptic dysfunction. *Trends in Molecular Medicine*, **11**, 170-176.

Lalonde, R., Dumont, M., Staufenbiel, M., Sturchler-Pierrat, C., and Strazielle, C. (2002). Spatial learning, exploration, anxiety, and motor coordination in female APP23 transgenic mice with the Swedish mutation. *Brain Research*, **956**, 36-44.

Lamour, Y., Bassant, M. H., Jobert, A., and Joly, M. (1989). Septo-hippocampal neurons in the aged rat: Relation between their electrophysiological and pharmacological properties and behavioral performances. *Neurobiology of Aging*, **10**, 181-186.

Lance, C. E., Butts, M. M., and Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, **9**, 202-220.

Lance, R. F., Kennedy, M. L., and Leberg, P. L. (2000). Classification bias in discriminant function analyses used to evaluate putatively different taxa. *Journal of Mammalogy*, **81**, 245-249.

Landauer, T. K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science*, **10**, 477-493.

Larder, B., Wang, D., Revell, A., Montaner, J., Harrigan, R., De Wolf, F., Lange, J., Wegner, S., Ruiz, L., Perez-Elias, M. J., Emery, S., Gatell, J., Monforte, A. D., Torti, C., Zazzi, M., and Lane, C. (2007). The development of artificial neural networks to predict virological response to combination HIV therapy. *Antiviral Therapy*, **12**, 15-24.

Larson, E., Wang, L., Bowen, J., McCormick, W., Teri, L., Crane, P., and Kukull, W. (2006). Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Annals of Internal Medicine*, **144**, 73-81.

Larson, J., Lynch, G., Games, D., and Seubert, P. (1999). Alterations in synaptic transmission and long-term potentiation in hippocampal slices from young and aged PDAPP mice. *Brain Research*, **840**, 23-35.

Lasko, T. A., Bhagwat, J. G., Zou, K. H., and Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics*, **38**, 404-415.

Lazarov, O., Robinson, J., Tang, Y.-P., Hairston, I. S., Korade-Mirnics, Z., Lee, V. M.-Y., Hersh, L. B., Sapolsky, R. M., Mirnics, K., and Sisodia, S. S. (2005). Environmental enrichment reduces Aβ levels and amyloid deposition in transgenic mice. *Cell*, **120**, 701-713.

Le Bars, P. L., Katz, M. M., Berman, N., Itil, T. M., Freedman, A. M., and Schatzberg, A. F. (1997). A placebo-controlled double-blind, randomized trial of an extract of *Ginkgo biloba* for dementia. North American EGb study group. *Journal of the American Medical Association*, **278**, 1327-1332.

LeBlanc, A. C. (2005). The role of apoptotic pathways in Alzheimer's disease neurodegeneration and cell death. *Current Alzheimer's Research*, **2**, 389-402.

Ledbetter, D. H., Riccardi, V. M., Airhart, S. D., Strobel, R. J., Keenan, B. S., and Crawford, J. D. (1981). Deletions of chromosome 15 as a cause of the Prader-Willi syndrome. *New England Journal of Medicine*, **304**, 325-329.

Lee, C.-K., Weindruch, R., and Prolla, T. A. (2000). Gene-expression profile of the aging brain in mice. *Nature Genetics*, **25**, 294-297.

Lee, M. K., Borchelt, D. R., Wong, P. C., Sisodia, S. S., and Price, D. L. (1996). Transgenic models of neurodegenerative diseases. *Current Opinion in Neurobiology*, **6**, 651-660.

Leighty, R., Nilsson, L., Potter, H., Costa, D., Low, M., Bales, K. Paul, S., and Arendash, G. (2004). Use of multimetric statistical analysis to characterize and discriminate between the performance of four Alzheimer's transgenic mouse lines differing in Abeta deposition. *Behavioural Brain Research*, **153**, 107-121.

Leighty, R. E. (2003). *Correlation, discriminant function, and factor analyses of a behavioral assessment battery for transgenicity and strain-differences in mice.* Master's thesis, University of South Florida.

Leighty, R. E., Runfeldt, M. J., Berndt, D. J., Schleif, W. S., Cracchiolo, J. R., Potter, H., and Arendash, G. W. (2008). Use of artificial neural networks to determine cognitive impairment and therapeutic effectiveness in Alzheimer's transgenic mice. *Journal of Neuroscience Methods*, **167**, 358-366..

Lesne, S., Koh, M. T., Kotilinek, L., Kayed, R., Glabe, C. G., Yang, A., Gallagher, M., and Ashe, K. H. (2006). A specific amyloid-beta protein assembly in the brain impairs memory. *Nature*, **440**, 352-357.

Letemendia, F., and Pampiglione, G. (1958). Clinical and electroencephalographic observations in Alzheimer's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, **21**, 167-172.

Letenneur, L., Larrieu, S., and Barberger-Gateau, P. (2004). Alcohol and tobacco consumption as risk factors of dementia: A review of epidemiological studies. *Biomedicine and Pharmacotherapy*, **58**, 95-99.

Leung, H.-C., Skudlarski, P., Gatenby, J. C., Peterson, B. S., and Gore, J. C. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cerebral Cortex*, **10**, 552-560.

Levine, M. S., Lloyd, R. L., Fisher, R. S., Hull, C. D., and Buchwald, N. A. (1987). Sensory, motor and cognitive alterations in aged cats. *Neurobiology of Aging*, **8**, 253-263.

Levy-Lahad, E., Poorkaj, P., Wang, K., Fu, Y. H., Oshima, J., Mulligan, J., and Schellenberg, G. D. (1996). Genomic structure and expression of STM2, the chromosome 1 familial Alzheimer disease gene. *Genomics*, **34**, 198-204.

Lewin, B. (2008). *Genes IX*. Sudbury, MA: Jones and Bartlett Publishers, Inc.

Lilliefors, H. W. (1968). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, **62**, 399-402.

Lim, G. P., Chu, T., Yang, F., Beech, W., Frautschy, S. A., and Cole, G. M. (2001). The curry spice curcumin reduces oxidative damage and amyloid pathology in an Alzheimer transgenic mouse. *Journal of Neuroscience*, **21**, 8370-8377.

Lin, F. H., Lin, R., Wisniewski, H. M., Hwang, Y.-W., Grundke-Iqbal, I., Healy-Louie, G., and Iqbal, K. (1992). Detection of point mutations in codon 331 of mitochondrial NADH dehydrogenase subunit 2 in Alzheimer's brains. *Biochemical & Biophysical Research Communications*, **182**, 238-246.

Lindsay, J., Laurin, D., Verreault, R., Hebert, R., Helliwell, B., Hill, G. B., and McDowell, I. (2002). Risk factors for Alzheimer's disease: A prospective analysis from the Canadian Study of Health and Aging. *American Journal of Epidemiology*, **156**, 445-453.

Lipton, S. A. (2005). The molecular basis of memantine action in Alzheimer's disease and other neurological disorders: Low-affffinity, uncompetitive antagonism. *Current Alzheimer Research*, **2**, 155-165.

Lisboa, P. J., and Taktak, A. F. G. (2006). The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks*, **19**, 408-415.

Lister, R. G. (1987). The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology*, **92**, 180-185.

Liu, B., Jiang, T., Ma, S., Zhao, H., Li, J., Jiang, X., and Zhang, J. (2006). Exploring candidate genes for human brain diseases from a brain-specific gene network. *Biochemical and Biophysical Research Communications*, **349**, 1308-1314.

Lleo, A., Blesa, R., Gendre, J., Castellvi, M., Pastor, P., Queralt, R., and Oliva, R. (2001). A novel presenilin2 gene mutation (D439A) in a patient with early-onset Alzheimer's disease. *Neurology*, **57**, 1926-1928.

Locascio, J. J., Growdon, J. H., and Corkin, S. (1995). Cognitive test performance in detecting, staging, and tracking Alzheimer's disease. *Archives of Neurology*, **52**, 1087-1099.

Loewenstein, D. A., Acevedo, A., Luis, C., Crum, T., Barker, W. W., and Duara, R. (2004). Semantic interference deficits and the detection of mild Alzheimer's disease and mild cognitive impairment without dementia. *Journal of the International Neuropsychological Society*, **10**, 91-100.

Loewenstein, D. A., Acevedo, A., Schram, L., Ownby, R., White, G., Mogosky, B., Barker, W. W., and Duara, R. (2003). Semantic interference in mild Alzheimer's disease: Preliminary findings. *American Journal of Geriatric Psychiatry*, **11**, 252-255.

Loewenstein, D. A., Amigo, A., Duara, R., Gutterman, A., Hurwitz, D., Berkowitz, N., Wilkie, F., Weinberg, G., Black, B., Gittelman, B., and Eisdorfer, C. (1989). A new scale for the assessment of functional status in Alzheimer's disease and related disorders. *Journal of Gerontology*, **44**, 114-121.

Loewenstein, D. A., Arguelles, T., Barker, W. W., Schram, L., Ownby, R., Acevedo, A., Mogosky, B., White, G., and Duara, R. (2001). The utility of a modified object memory test in distinguishing between three different age groups of Alzheimer's disease patients and normal controls. *Journal of Mental Health and Aging*, **7**, 317-324.

Loewenstein, D. A., Rubert, M. P., Arguelles, T., and Duara, R. (1995). Neuropsychological test performance and prediction of functional capacities among Spanish-speaking and English-speaking patients with dementia. *Archives of Clinical Neuropsychology*, **10**, 75-88.

Logue, S. F., Paylor, R., and Wehner, J. M. (1997). Hippocampal lesions cause learning deficits in inbred mice in the Morris water maze and conditioned-fear task. *Behavioral Neuroscience*, **111**, 104-113.

Lopez-Rangel, E., and Lewis, M. (2006). Loud and clear evidence for gene silencing by epigenetic mechanisms in autism spectrum and related neurodevelopmental disorders. *Clinical Genetics*, **69**, 21-25.

Lutzenberger, W., Birbaumer, N., Flor, H., Rockstroh, B., and Elbert, T. (1992). Dimensional analysis of the human EEG and intelligence. *Neuroscience Letters*, **143**, 10-14.

Lynch, M. A. (2004). Long-term potentiation and memory. *Physiological Reviews*, **84**, 87-136.

MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, **4**, 84-99.

Maccioni, R., Munoz, J. and Barbeito, L. (2001). The molecular bases of Alzheimer's disease and other neurodegenerative disorders. *Archives of Medical Research*, **32**, 367-381.

Magenis, R. E., Brown, M. G., Lacy, D. A., Budden, S., and LaFranchi, S. (1987). Is Angelman syndrome an alternate result of del(15)(q11q13)? *American Journal of Medical Genetics*, **28**, 829-838.

Maia, L., and DeMendonca, A. (2002). Does caffeine intake protect from Alzheimer's disease? *European Journal of Neurology*, **9**, 377-382.

Markesbery, W. R. (1997). Neuropathological criteria for the diagnosis of Alzheimer's disease. *Neurobiology of Aging*, **18(S4)**, S13-S19.

Markowska, A. L., Stone, W. S., Ingram, D. K., Reynolds, J., Gold, P. E., Conti, L. H., Pontecorvo, M. J., Wenk, G. L., and Olton, D. S. (1989). Individual differences in aging: Behavioral and neurobiological correlates. *Neurobiology of Aging*, **10**, 31-43.

Marnett, L. J., and Kalgutkar, A. S. (1999). Cyclooxygenase 2 inhibitors: discovery, selectivity and the future. *Trends Pharmacological Sciences*, **20**, 465469.

Martone, M., Butters, N., Payne, M., Becker, J., and Sax, D. (1984). Dissociations between skill learning and verbal recognition in amnesia and dementia. *Archives of Neurology*, **41**, 965-970.

Masaki, K. H., Losonczy, K. G., and Izmirlian, G. (2000). Association of vitamin E and C supplement use with cognitive function and dementia in elderly men. *Neurology*, **54**, 1265-1272.

Masliah, E., Sisk, A., Mallory, M., and Games, D. (2001). Neurofibrillary pathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *Journal of Neuropathology and Experimental Neurology*, **60**, 357-368.

Masliah, E., Sisk, A., Mallory, M., Mucke, L., Schenk, D., and Games, D. (1996). Comparison of neurodegenerative pathology in transgenic mice overexpressing V717F beta-amyloid precursor protein and Alzheimer's disease. *Journal of Neuroscience*, **16**, 5795-5811.

Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. (1985). Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proceedings of the National Academy of Sciences USA*, **82**, 4245-4249.

Mastropaolo, J., Nadi, N. S., Ostrowski, N. L., and Crawley, J. N. (1988). Galanin antagonizes acetylcholine on a memory task in basal forebrain-lesioned rats. *Proceedings of the National Academy of Sciences USA*, **85**, 9841-9845.

Mathis, C. A., Wang, Y., and Klunk, W. E. (2004). Imaging beta-amyloid plaques and neurofibrillary tangles in the aging human brain. *Current Pharmaceutical Design*, **13**, 1469-1492.

Mattson, M. P. (2004). Pathways towards and away from Alzheimer's disease. *Nature*, **430**, 631-639.

Mattson, M. P., Chan, S. L., and Duan, W. (2002). Modification of brain aging and neurodegenerative disorders by genes, diet, and behavior. *Physiological Reviews*, **82**, 637-672.

Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., Specht, S. M., and Gandhi, C. C. (2003). Individual differences in the expression of a "general" learning ability in mice. *Journal of Neuroscience*, **23**, 6423-6433.

Maurer, K., Volk, S., and Gerbaldo, H. (1997). Auguste D. and Alzheimer's disease. *Lancet*, **349**, 1546-1549.

Maviel, T., Durkin, T. P., Menzaghi, F., and Bontempi, B. (2004). Sites of neocortical reorganization critical for remote spatial memory. *Science*, **305**, 96-99.

McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., Ahern, F., Petrill, S. A., and Plomin, R. (1997). Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science*, **276**, 1560-1563.

McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, **5**, 115-133.

McDonald, M. P., and Overmier, J. B. (1998). Present imperfect: A critical review of animal models of the mnemonic impairments in Alzheimer's disease. *Neuroscience and Biobehavioral Reviews*, **22**, 99-120.

McGeer, P. L., and McGeer, E. G. (2007). NSAIDs and Alzheimer's disease: Epidemiological, animal model and clinical studies. *Neurobiology of Aging*, **28**, 639-647.

McGeer, P. L., Schulzer, M., and McGeer, E. G. (1996). Arthritis and anti-inflammatory agents as possible protective factors for Alzheimer's disease: a review of 17 epidemiologic studies. *Neurology*, **47**, 425432.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, **34**, 939-944.

Meyer, D. E., and Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, **90**, 227-234.

Meyer-Luehmann, M., Coomaraswamy, J., Bolmont, T., Kaeser, S., Schaefer, C., Kilger, E., Neuenschwander, A., Abramowski, D., Frey, P., Jaton, A. L., Vigouret, J.-M., Paganetti, P., Walsh, D. M., Mathews, P. M., Ghiso, J., Staufenbiel, M., Walker, L. C., and Jucker, M. (2006). Exogenous induction of cerebral beta-amyloidogenesis is governed by agent and host. *Science*, **313**, 1781-1784.

Middei, S., Daniele, S., Caprioli, A., Ghirardi, O., and Ammassari-Teule, M. (2006). Progressive cognitive decline in a transgenic mouse model of Alzheimer's disease overexpressing mutant hAPPswe. *Genes, Brain and Behavior*, **5**, 249-256.

Milgram, N. W., Head, E., Zicker, S. C., Ikeda-Douglas, C. J., Murphey, H., Muggenburg, B., Siwak, C., Tapp, D., and Cotman, C. W. (2005). Learning ability in aged beagle dogs is preserved by behavioral enrichment and dietary fortification: A two-year longitudinal study. *Neurobiology of Aging*, **26**, 77-90.

Miller, A. S., Blott, B. H., and Hames, T. K. (1992). Review of neural network applications in medical imaging and signal processing. *Medical and Biological Engineering and Computing*, **30**, 449-464.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.

Miller, R. M., and Federoff, H. J. (2008). Isoform-specific effects of ApoE on HSV immediate early gene expression and establishment of latency. *Neurobiology of Aging*, **29**, 71-77.

Minkeviciene, R., Banerjee, P., and Tanila, H. (2004). Memantine improves spatial learning in a transgenic mouse model of Alzheimer's disease. *Journal of Pharmacology and Experimental Therapeutics*, **311**, 677-682.

Minoshima, S. (2003). Imaging Alzheimer's disease: Clinical applications. *Neuroimaging Clinics of North America*, **13**, 769-780.

Minsky, M., and Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.

Mirakur, A., Craig, D., Hart, D. J., McIlroy, S. P., and Passmore, A. P. (2004). Behavioral and psychological syndromes in Alzheimer's disease. *International Journal of Geriatric Psychiatry*, **19**, 1035-1039.

Mirnics, K., and Pevsner, J. (2004). Progress in the use of microarray technology to study the neurobiology of disease. *Nature Neuroscience*, **7**, 434-439.

Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 31-49.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., and Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, **130**, 621-640.

Miyashita, Y. (2004). Cognitive memory: Cellular and network machineries and their top-down control. *Science*, **306**, 435-440.

Moechars, D., Lorent, K., De Strooper, B., Dewachter, I., and Van Leuven, F. (1996). Expression in brain of amyloid precursor protein mutated in the alpha-secretase site causes disturbed behavior, neuronal degeneration and premature death in transgenic mice. *EMBO Journal*, **15**, 1265-1274.

Mohajeri, M., Madani, R., Saini, K., Lipp, H., Nitsch, R., and Wolfer, D. (2004). The impact of genetic background on neurodegeneration and behavior in seizure mice. *Genes, Brain and Behavior*, **3**, 228-239.

Molle, M., Marshall, L., Wolf, B., Fehm, H. L., and Born, J. (1999). EEG complexity and performance measures of creative thinking. *Psychophysiology*, **36**, 95-104.

Montgomery, K. C. (1953). Concerning the use of analysis of variance on latency data. *American Journal of Psychology*, **66**, 131-135.

Montgomery, K. C. (1955). The relation between fear induced by novel stimulation and exploratory behavior. *Journal of Comparative and Physiological Psychology*, **48**, 254-260.

Morelli, L., Prat, M. I., Levy, E., Mangone, C. A., Castano, E. M. (1998), Presenilin 1 met146leu variant due to an A-T transversion in an early-onset familial Alzheimer's disease pedigree from Argentina. *Genetics*, **53**, 469-473.

Morihara, T., Chu, T., Ubeda, O., Beech, W., and Cole, G. M. (2002). Selective inhibition of Abeta42 production by NSAID R-enantiomers. *Journal of Neurochemistry*, **83**, 1009-1012.

Morris, J. C., Edland, S., Clark, C., Galasko, D., Koss, E., Mohs, R., van Belle, G., Fillenbaum, G., and Heyman, A. (1993). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part IV. Ratings of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology*, **43**, 2457-2465.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., and Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part 1. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, **39**, 1159-1165.

Morris, M., Evans, D., Tangney, C., Bienias, J., and Wilson, R. (2005). Fish consumption and cognitive decline with age in a large community study. *Archives of Neurology*, **62**, 1849-1853.

Morris, M. C., Beckett, L. A., Scherr, P. A., Hebert, L. E., Bennett, D. A., Field, T. S., and Evans, D. A. (1998). Vitamin E and vitamin C supplement use and risk of incident Alzheimer disease. *Alzheimer Disease and Associated Disorders*, **12**, 121-126.

Morris, R., Anderson, E., Lynch, G., and Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature*, **319**, 774-776.

Morris, R. G. M. (1981). Spatial localization does not depend on the presence of local cues. *Learning and Motivation*, **12**, 239-260.

Morris, R. G. M. (1984). Developments of water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods*, **11**, 47-60.

Morris, R. G. M. (2001). Episodic-like memory in animals: Psychological criteria, neural mechanisms and the value of episodic-like tasks to investigate animal models of neurodegenerative disease. *Philosophical Transactions of the Royal Society of London, Series B*, **356**, 1453-1465.

Morris, R. G. M., Garrud, P., Rawlins, J. N. P., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, **297**, 681-683.

Morrison, J. H., and Hof, P. R. (1997). Life and death of neurons in the aging brain. *Science*, **278**, 412-419.

Moser, V. C., Cheek, B. M., and MacPhail, R. C. (1995). A multidisciplinary approach to toxicological screening. III. Neurobehavioral toxicology. *Journal of Toxicology and Environmental Health*, **45**, 173-210.

Moulin, C. J. A., Laine, M., Rinne, J. O., Kaasinen, V., Sipila, H., Hiltunen, J., and Kangasmaki, A. (2007). Brain function during multi-trial learning in mild cognitive impairment: A PET activation study. *Brain Research*, **1136**, 132-141.

Mudher, A. and Lovestone, S. (2002). Alzheimer's disease: Do tauists and baptists finally shake hands? *TRENDS in Neurosciences*, **25**, 22-26.

Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B., and Lannfelt, L. (1992). A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nature Genetics*, **1**, 345-347.

Mundt, J. C., Freed, D. M., and Greist, J. H. (2000). Lay person-based screening for early detection of Alzheimer's disease: Development and validation of an instrument. *Journal of Gerontology: Psychological Sciences*, **55B**, P163-P170.

Murrell, J., Farlow, M., Ghetti, B., and Benson, M. (1991). A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science*, **254**, 97-99.

Nair, J., Nair, S. S., Kashani, J. H., Reid, J. C., Mistry, S. I., and Vargas, V. G. (1999). Analysis of the symptoms of depression: A neural network approach. *Psychiatry Research*, **87**, 193-201.

Nakayama, H., Uchida, K., and Doi, K. (2004). A comparative study of age-related brain pathology: Are neurodegenerative diseases present in nonhuman animals? *Medical Hypotheses*, **63**, 198-202.

Nalini, K., Karanth, K. S., Rao, A., and Aroor, A. R. (1995). Effects of *Celastrus paniculatus* on passive avoidance performance and biogenic amine turnover in albino rats. *Journal of Ethnopharmacology*, **47**, 101-108.

Neely, R. (1977). Discriminant analysis for prediction of college graduation. *Educational and Psychological Measurement*, **37**, 965-970.

Nehlig, A., Daval, J.-L., and Debry, G. (1992). Caffeine and the central nervous system: Mechanisms of action, biochemical, metabolic and psychostimulant effects. *Brain Research Reviews*, **17**, 139-170.

Nelson, P. T., Greenberg, S. G., and Saper, C. B. (1994). Neurofibrillary tangles in the cerebral cortex of sheep. *Neuroscience Letters*, **170**, 187-190.

Neyman, J., and Pearson, E. S. (1933). The testing of statistical hypotheses in relation to probabilities *a priori*. Reprinted in: Neyman, J., and Pearson, E. S. (1967). *Joint Statistical Papers*. Cambridge: Cambridge University Press. pp. 186-202.

Nguyen, P. V., Abel, T., Kandel, E. R., and Bourtchouladze, R. (2000). Strain-dependent differences in LTP and hippocampus-dependent memory in inbred mice. *Learning and Memory*, **7**, 170-179.

Nichol, K. E., Parachikova, A. I., and Cotman, C. W. (2007). Three weeks of running wheel exposure improves cognitive performance in the aged Tg2576 mouse. *Behavioral Brain Research*, **184**, 124-132.

Nilsson, J., Ohlsson, M., Thulin, L., Höglund, P., Nashef, S. A. M., and Brandt, J. (2006). Risk factor identification and mortality prediction in cardiac surgery using artificial neural networks. *Journal of Thoracic and Cardiovascular Surgery*, **132**, 12-19.

Nilsson, L. N., Arendash, G. W., Leighty, R. E., Costa, D. A., Low, M. A., Garcia, M. F., Cracchiolo, J. R., Rojiani, A., Wu, X., Bales, K. R., Paul, S. M., and Potter, H. (2004). Cognitive impairment in PDAPP mice depends on ApoE and ACT-catalyzed amyloid formation. *Neurobiology of Aging*, **25**, 1153-1167.

Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, **19**, 1-32.

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, **24**, 1565-1567.

Nunn, J. A., Graydon, F. J., Polkey, C. E., and Morris, R. G. (1999). Differential spatial memory impairment after right temporal lobectomy demonstrated using temporal titration. *Brain*, **122**, 47-59.

Nyberg, L., Habib, R., and Herlitz, A. (2000). Brain activation during episodic memory retrieval: Sex differences. *Acta Psychologica*, **105**, 181-194.

Obuchowski, N. A. (2003). Receiver operating characteristic curves and their use in radiology. *Radiology*, **229**, 3-8.

Oddo, S., Caccamo, A., Kitazawa, M., Tseng, B. P., and LaFerla, F. M. (2003a). Amyloid deposition precedes tangle formation in a triple transgenic model of Alzheimer's disease. *Neurobiology of Aging*, **24**, 1063-1070.

Oddo, S., Caccamo, A., Shepherd, J. D., Murphy, M. P., Golde, T. E., Kayed, R., Metherate, R., Mattson, M. P., Akbari, Y., and LaFerla, F. M. (2003b). Triple-transgenic model of Alzheimer's disease with plaques and tangles: Intracellular Abeta and synaptic dysfunction. *Neuron*, **39**, 409-421.

Oddo, S., Caccamo, A., Tran, L., Lambert, M. P., Glabe, C. G., Klein, W. L., and LaFerla, F. M. (2006). Temporal profile of amyloid-beta (Abeta) oligomerization in an in vivo model of Alzheimer's disease. A link between Abeta and tau pathology. *Journal of Biological Chemistry*, **281**, 1599-1604.

O'Donnell, B. F., Drachman, D. A., Lew, R. A., and Swearer, J. M. (1988). Measuring dementia: Assessment of multiple deficit domains. *Journal of Clinical Psychology*, **44**, 916-923.

Olazaran, J., Muniz, R., Reisberg, B., Pena-Casanova, J., del Ser, T., Cruz-Jentoft, A., Serrano, P., Navarro, E., Garcia de la Rocha, M., Frank, A., Galiano, M., Fernandez-Bullido, Y., Serra, J., Gonzalez-Salvador, M., and Sevilla, C. (2004). Benefits of cognitive-motor intervention in MCI and mild to moderate Alzheimer disease. *Neurology*, **63**, 2348-2353.

Olson, J. M., Goddard, K. A. B., and Dudek, D. M. (2002). A second locus for very-late-onset Alzheimer disease: a genome scan reveals linkage to 20p and epistasis between 20p and the amyloid precursor protein region. *American Journal of Human Genetics*, **71**, 154-161.

Olton, D. S. (1977). Spatial memory. *Scientific American*, **236(6)**, 82-98.

Olton, D. S., and Samuelson, R. J. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, **2**, 97-116.

Ono, K., Hasegawa, K., Naiki, H., and Yamada, M. (2004). Curcumin has potent anti-amyloidogenic effects for Alzheimer's beta-amyloid fibrils in vitro. *Journal of Neuroscience Research*, **75**, 742-750.

Ossenkopp, K. P., Sorenson, L., and Mazmanian, D. S. (1994). Factor analysis of open field behavior in the rat (*Rattus norvegicus*): Application of the PARAFAC model to a longitudinal data set. *Behavioral Processes*, **31**, 129-144.

Ownby, R. L., Loewenstein, D. A., Schram, L., and Acevedo, A. (2004). Assessing the cognitive abilities that differentiate patients with Alzheimer's disease from normals: Single and multiple factor models. *International Journal of Geriatric Psychiatry*, **19**, 232-242.

Papassotiropoulos, A., Wollmer, M. A., Aguzzi, A., Hock, C., Nitsch, R. M., and de Quervain, D. J. (2005). The prion gene is associated with human long-term memory. *Human Molecular Genetics*, **14**, 2241-2246.

Papassotiropoulos, A., Stephan ,D. A., Huentelman, M. J., Hoerndli, F. J., Craig, D. W., Pearson, J. V., Huynh, K.-D., Brunner, F., Corneveaux, J., Osborne, D., Wollmer, M. A., Aerni, A., Coluccia, D., Hänggi, J., Mondadori, R. A., Buchmann, A., Reiman, E. M., Caselli, R. J., Henke, K., and de Quervain, D. J.-F. (2006). Common *Kibra* alleles are associated with human memory performance. *Science*, **314**, 475-478.

Papik, K., Molnar, B., Schaefer, R., Dombovari, Z., Tulassay, Z., and Feher, J. (1996). Application of neural networks in medicine: A review. *Medical Science Monitor*, **4**, 538-546.

Pappas, B. A., Bayley, P. J., Bui, B. K., Hansen, L. A., and Thal, L. J. (2000). Choline acetyltransferase activity and cognitive domain scores of Alzheimer's patients. *Neurobiology of Aging*, **21**, 11-17.

Pardo, L. M., and van Duijn, C. M. (2005). In search of genes involved in neurodegenerative disorders. *Mutation Research*, **592**, 89-101.

Parsons, S., and Jones, G. (2000). Acoustic identification of twelve species of echolocating bat by discriminant function analysis and artificial neural networks. *Journal of Experimental Biology*, **203**, 2641-2656.

Pascuale-Leone, A., Wassermann, E. M., Grafman, J., and Hallett, M. (1996). The role of the dorsolateral prefrontal cortex in implicit procedural learning. *Experimental Brain Research*, **107**, 479-485.

Patel, N. S., Paris, D., Mathura, V., Quadros, A. N., Crawford, F. C., and Mullan, M. J. (2005). Inflammatory cytokine levels correlate with amyloid load in transgenic mouse models of Alzheimer's disease. *Journal of Neuroinflammation*, **2**, 9. DOI: 10.1186/1742-2094-2-9.

Peek, M. S., Russek-Cohen, E., Wait, D. A., and Forseth, I. N. (2002). Physiological response curve analysis using nonlinear mixed models. *Oecologia*, **132**, 175-180.

Pendharkar, P. C. (2002). A computational study on the performance of artificial neural networks under changing structural design and data distribution. *European Journal of Operational Research*, **138**, 155-177.

Pennanen, C., Kivipelto, M., Tuomainen, S., Hartikainen, P., Hanninen, T., Laakso, M. P., Hallikainen, M., Vanhanen, M., Nissinen, A., Helkala, E. L., Vainio, P., Vanninen, R., Partanen, K., and Soininen, H. (2004). Hippocampus and entorhinal cortex in mild cognitive impairment and early Alzheimer's disease. *Neurobiology of Aging*, **25**, 303-310.

Periani, D., Bressi, S., Cappa, S. F., Vallar, G., Alberoni, M., Grassi, F., Caltagirone, C., Ciplotti, L., Franceschi, M., Leniz, G. L., and Fazio, F. (1993). Evidence of multiple memory systems in the human brain. *Brain*, **106**, 903-909.

Perry, E. K., Pickering, A. T., Wang, W. W., Houghton, P. J., and Perry, N. S. (1999). Medicinal plants and Alzheimer's disease: From ethnobotany to phytotherapy. *Journal of Pharmacy and Pharmacology*, **51**, 527-534.

Peterson, B. S., Skudlarski, P., Gatenby, J. C., Zhang, H., Anderson, A. W., and Gore, J. C. (1999). An fMRI study of Stroop word-color interference: Evidence for cinguate subregions subserving multiple distributed attentional systems. *Biological Psychiatry*, **45**, 1237-1258.

Peterson, L. R., and Peterson, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology*, **58**, 193-198.

Pfeiffer, E. (1975). A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatric Society*, **23**, 433-441.

Philipson, J. D. (2003). 50 years of medicinal plant research – Every progress in methodology is a progress is science. *Planta Medica*, **69**, 491-495.

Phillips-Wren, G., Sharkey, P., and Dy, S. M. (2007). Mining lung cancer patient data to assess healthcare resource utilization. *Expert Systems with Applications*, **35**, 1611-1619.

Picciotto, M. R., and Wickman, K. (1998). Using knockout and transgenic mice to study neurophysiology and behavior. *Physiological Review*, **78**, 1131-1163.

Plomin, R. (2001). The genetics of *g* in human and mouse. *Nature Reviews Neuroscience*, **2**, 136-141.

Pompl, P. N., Mullan, M. J., Bjugstad, K., and Arendash, G. W. (1999). Adaptation of the circular platform spatial memory task for mice: Use in detecting cognitive impairment in the APPsw transgenic mouse model for Alzheimer's disease. *Journal of Neuroscience Methods*, **87**, 87-95.

Posner, M. I., DiGirolamo, G. J., and Fernandez-Duque, D. (1997). Brain mechanisms of cognitive skills. *Consciousness and Cognition*, **6**, 267-290.

Postman, L., and Underwood, B. J. (1973). Critical issues in interference theory. *Memory and Cognition*, **1**, 19-40.

Potter, H., Wefes, I. and Nilsson, L. (2001). The inflammation-induced pathological chaperones ACT and apo-E are necessary catalysts of Alzheimer amyloid formation. *Neurobiology of Aging*, **22**, 932-930.

Pradhan, N., and Dutt, D. N. (1993). A nonlinear perspective in understanding the neurodynamics of EEG. *Computers in Biology and Medicine*, **23**, 425-442.

Pradhan, N., and Sadasivan, P. K. (1996). The nature of dominant Lyapunov exponent and attractor dimension curves of EEG in sleep. *Computers in Biology and Medicine*, **26**, 419-428.

Pratico, D., Uryu, K., Leight, S., Trojanowski, J. K., and Lee, V. M.-Y. (2001). Increased lipid peroxidation precedes amyloid plaque formation in an animal model of Alzheimer amyloidosis. *Journal of Neuroscience*, **21**, 4183-4187.

Price, D. L., and Sisodia, S. S. (1998). Mutant genes in familial Alzheimer's disease and transgenic models. *Annual Review of Neuroscience*, **21**, 479-505.

Price, J., Ko, A., Wade, M., Tsou, S., McKeel, D. and Morris, J. (2001). Neuron number in the entorhinal cortex and CA1 in preclinical Alzheimer disease. *Archives of Neurology*, **58**, 1395-1402.

Price, R. K., Spitznagel, E. L., Downey, T. J., Meyer, D. J., and Risk, N. K. (2000). Applying artificial neural network models to clinical decision making. *Psychological Assessment*, **12**, 40-51.

Pritchard, W. S., and Duke, D. W. (1995). Measuring chaos in the brain: A tutorial review of EEG dimension estimation. *Brain and Cognition*, **27**, 353-397.

Provost, F., and Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning*, **42**, 203-231.

Prusiner, S. B. (1991). Molecular biology of prion diseases. *Science*, **252**, 1515-1522.

Pugh, P. L., Ahmed, S. F., Smith, M. I., Upton, N., and Hunter, A. J. (2004). A behavioral characterization of the FVB/N mouse strain. *Behavioral Brain Research*, **155**, 283-289.

Pugh, P. L., Richardson, J. C., Bate, S. T., Upton, N., and Sunter, D. (2007). Non-cognitive behaviors in an APP/PS1 transgenic model of Alzheimer's disease. *Behavioral Brain Research*, **178**, 18-28.

Qizilbash, N., Whitehead, A., Higgins, J., Wilcock, G., Schneider, L., and Farlow, M. (1998). Cholinesterase inhibition for Alzheimer disease: A meta-analysis of the tacrine trials. *Journal of the American Medical Association*, **280**, 1777-1782.

Quayhagen, M. and Quayhagen, M. (1989). Differential effects of family-based strategies on Alzheimer's disease. *The Gerontologist*, **29**, 150-155.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, **1**, 81-106.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers.

Ramirez, B. G., Blazquez, C., del Pulgar, T. G., Guzman, M., and de Ceballos, M. L. (2005). Prevention of Alzheimer's disease pathology by cannabinoids: Neuroprotection mediated by blockade of microglial activation. *Journal of Neuroscience*, **25**, 1904-1913.

Rammes G, Rupprecht R, Ferrari U, Zieglgansberger W, and Parsons CG. (2001). The N-methyl-D-aspartate receptor channel blockers memantine, MRZ 2/579 and other amino-alkyl-cyclohexanes antagonise 5-HT(3) receptor currents in cultured HEK-293 and N1E-115 cell systems in a non-competitive manner. *Neuroscience Letters*, **306**, 81-4.

Ramos, A., and Mormede, P. (1998). Stress and emotionality: A multidimensional and genetic approach. *Neuroscience and Biobehavioral Reviews*, **22**, 33-57.

Raskind, M., Sadowsky, C., Sigmund, W., Beitler, P. and Auster, S. (1997). Effect of tacrine on language, praxis, and noncognitive behavioral problems in Alzheimer disease. *Archives of Neurology*, **54**, 836-840.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, **97**, 285-308.

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., Bush, G., Breiter, H. C., and Rosen, B. R. (1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Human Brain Mapping*, **5**, 124-132.

Raudys, S. J., and Jain, A. K. (1991). Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, **13**, 252-264.

Ray, S., Britschgi, M., Herbert, C., Takeda-Uchimura, Y., Boxer, A., Blennow, K., Friedman, L. F., Galasko, D. R., Jutel, M., Karydas, A., Kaye, J. A., Leszek, J., Miller, B. L., Minthon, L., Quinn, J. F., Rabinovici, G. D., Robinson, W. H., Sabbagh, M. N., So, Y. T., Sparks, D. L., Tabaton, M., Tinklenberg, J., Yesavage, J. A., Tibshirani, R., and Wyss-Coray, T. (2007). Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nature Medicine*, **13**, 1359-1362.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, **6**, 855-863.

Rechtschaffen, A., and Kales, A. (1968). *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages in Human Subjects*. (NIH Publication No. 204). Washington, DC: Public Health Services, U. S. Government Printing Office.

Reddy, P. H. (2006). Amyloid precursor protein-mediated free radicals and oxidative damage: Implications for the development and progression of Alzheimer's disease. *Journal of Neurochemistry*, **96**, 1-13.

Reed, J., and Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 585-594.

Reilly, J. F., Games, D., Rydel, R. E., Freedman, S., Schenk, D., Young, W. G., Morrison, J. H., and Bloom, F. E. (2003). Amyloid deposition in the hippocampus and entorhinal cortex: Quantitative analysis of a transgenic mouse model. *Proceedings of the National Academy of Sciences USA*, **100**, 4837-4842.

Reiman, E. M., Webster, J. A., Myers, A. J., Hardy, J., Dunckley, T., Zismann, V. L., Joshipura, K. D., Pearson, J. V., Hu-Lince, D., Huentelman, M. J., Craig, D. W., Coon, K. D., Liang, W. S., Herbert, R. H., Beach, T., Rohrer, K. C., Zhao, A. S., Leung, D., Bryden, L., Marlowe, L., Kaleem, M., Mastroeni, D., Grover, A., Heward, C. B., Ravid, R., Rogers, J., Hutton, M. L., Melquist, S., Petersen, R. C., Alexander, G. E., Caselli, R. J., Kukull, W., Papassotiropoulos, A., and Stephan, D. A. (2007). GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron*, **54**, 713-720.

Reisberg, B., Doody, R., Stöffler, A., Schmitt, F., Ferris, S., and Möbius, H. J. (2003). Memantine in moderate-to-severe Alzheimer's disease. *New England Journal of Medicine*, **348**, 1333-1341.

Reisberg, B., Franssen, E. H., Hasan, S. M., Monteiro, I., Boksay, I., Souren, L. E., Kenowsky, S., Auer, S. R., Elahi, S., and Kluger, A. (1999). Retrogenesis: Clinical, physiologic, and pathologic mechanisms in brain aging, Alzheimer's and other dementing processes. *European Archives of Psychiatry and Clinical Neurosciences*, **249** Suppl 3, 28-36.

Riek, R. (2006). Infectious Alzheimer's disease? *Nature*, **444**, 429-431.

Risner, M. E., Saunders, A. M., Altman, J. F. B., Ormandy, G. C., Craft, S., Foley, I. M., Zvartau-Hind, M. E., Hosford, D. A., and Roses, A. D. (2006). Efficacy of rosiglitazone in a genetically defined population with mild-to-moderate Alzheimer's disease. *The Pharmacogenomics Journal*, **6**, 246-254.

Ritchie, K., Carriere, I., de Mendonca, A., Portet, F., Dartigues, J. F., Rouaud, O., Barberger-Gateau, P., and Ancelin, M. L. (2007). The neuroprotective effects of caffeine. A prospective population study (the Three City Study). *Neurology*, **69**, 536-545.

Roberson, E. D., Scearce-Levie, K., Palop, J. J., Yan, F., Cheng, I. H., Wu, T., Gerstein, H., Yu, G. Q., and Mucke, L. (2007). Reducing endogenous tau ameliorates amyloid beta-induced deficits in an Alzheimer's disease mouse model. *Science*, **316**, 750-754.

Robert, C., Guilpin, C., and Limoge, A. (1997). Comparison between conventional and neural network classifiers for rat sleep-wake stage discrimination. *Neuropsychobiology*, **35**, 221-225.

Robert, C., Karasinski, P., Natowicz, R., and Limoge, A. (1996). Adult rat vigilance states discrimination by artificial neural networks using a single EEG channel. *Physiology and Behavior*, **59**, 1051-1060.

Roberts, W. A., and Dale, R. H. I. (1981). Remembrance of places lasts: Proactive interference and patterns of choice in rat spatial memory. *Learning and Motivation*, **12**, 261-281.

Robinson, D. M., and Keating, G. M. (2006). Memantine: A review of its use in Alzheimer's disease. *Drugs*, **66**, 1515-1534.

Rocchi, A., Pellegrini, S., Siciliano, G., and Murri, L. (2003). Causative and susceptibility genes for Alzheimer's disease: A review. *Brain Research Bulletin*, **61**, 1-24.

Rockenstein, E. M., McConlogue, L., Tan, H., Power, M., Masliah, E., and Mucke, L. (1995). Levels and alternative splicing of amyloid beta protein precursor (APP) transcripts in brains of APP transgenic mice and humans with Alzheimer's disease. *Journal of Biological Chemistry*, **270**, 28257-28267.

Rodgers, R. J., and Johnson, N. J. T. (1995). Factor analysis of spatiotemporal and ethological measures in the murine elevated plus-maze test of anxiety. *Pharmacology Biochemistry and Behavior*, **52**, 297-303.

Roertgen, K. E., Parisi, J. E., Clark, H. B., Barnes, D. L., O'Brien, T. D., and Johnson, K. H. (1996). Aβ-associated cerebral angiopathy and senile plaques with neurofibrillary tangles and cerebral hemorrhage in an aged wolverine (*Gulo gulo*). *Neurobiology of Aging*, **17**, 243-247.

Rogaev, E. I., Sherrington, R., Rogaeva, E. A., Levesque, G., Ikeda, M., Liang, Y., Chi, H., Lin, C., Holman, K., Tsuda, T., Mar, L., Sorbi, S., Nacmias, B., Placentini, S., Amaducci, L., Chumakov, I., Cohen, D., Lannfelt, L., Fraser, P. E., Rommens, J. M., and St George-Hyslop, P. H. (1995). Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature*, **376**, 775-778.

Rogaev, E. I., Sherrington, R., Wu, C., Levesque, G., Liang, Y., Rogaeva, E. A., Ikeda, M., Holman, K., Lin, C., Lukiw, W. J., de Jong, P. J., Fraser, P. E., Rommens, J. M., St. George-Hyslop, P. (1997), Analysis of the 5-prime sequence, genomic structure, and alternative splicing of the presenilin-1 gene (PSEN1) associated with early onset Alzheimer disease. *Genomics*, **40**, 415-424.

Rogaeva, E., Meng, Y., Lee, J. H., Gu, Y., Kawarai, T., Zou, F., Katayama, T., Baldwin, C. T., Cheng, R., Hasegawa, H., Chen, F., Shibata, N., Lunetta, K. L., Pardossi-Piquard, R., Bohm, C.,Wakutani, Y., Cupples, L. A., Cuenco, K. T., Green, R. C., Pinessi, L., Rainero, I., Sorbi, S., Bruni, A., Duara, R., Friedland, R. P., Inzelberg, R., Hampe, W., Bujo, H., Song, Y.-Q., Andersen, O. M., Willnow, T. E., Graff-Radford, N., Petersen, R. C., Dickson, D., Der, S. D., Fraser, P. E., Schmitt-Ulms, G., Younkin, S., Mayeux, R., Farrer, L. A. and St. George-Hyslop, P. (2007). The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nature Genetics*, **39**, 168-177.

Rogers, D. C., Fisher, E. M. C., Brown, S. D. M., Peters, J., Hunter, A. J., and Martin, J. E. (1997). Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment. *Mammalian Genome*, **8**, 711-713.

Rogers, D. C., Jones, D. N. C., Nelson, P. R., Jones, C. M., Quilter, C. A., Robinson, T. L., and Hagan, J. J. (1999). Use of SHIRPA and discriminant analysis to characterize marked differences in the behavioral phenotype of six inbred mouse strains. *Behavioral Brain Research*, **105**, 207-217.

Rogers, S. L., Farlow, M. R., Doody, R. S., Mohs, R., and Friedhoff, L. T. (1998). A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. *Neurology*, **50**, 136-145.

Roher, A., Baudry, J., Chaney, M., Kuo, Y., Stine, W. and Emmerling, M. (2000). Oligomerization and fibril assembly of the amyloid-b protein. *Biochimica et Biophysica Acta*, **1502**, 31-43.

Roher, A., Wolfe, D., Palutke, M., and KuKuruga, D. (1986). Purification, ultrastructure, and chemical analysis of Alzheimer disease amyloid plaque core protein. *Proceedings of the National Academy of Sciences USA*, **83**, 2662-2666.

Rolland, Y., Pillard, F., Klapouszczak, A., Reynish, E., Thomas, D., Andrieu, S., Riviere, D., and Vellas, B. (2007). Exercise program for nursing home residents with Alzheimer's disease: A 1-year randomized, controlled trial. *Journal of the American Geriatrics Society*, **55**, 158-165.

Rosen, R. F., Farberg, A. S., Gearing, M., Dooyema, J., Long, P. M., Anderson, D. C., Davis-Turak, J., Coppola, G., Geschwind, D. H., Pare, J.-F., Duong, T. Q., Hopkins, W. D., Preuss, T. M., and Walker, L. C. (2008). Tauopathy with paired helical filaments in an aged chimpanzee. *Journal of Comparative Neurology*, **509**, 259-270.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, **65**, 386-408.

Rössler, M., Zarski, R., Bohl, J. and Ohm, T. (2002). Stage-dependent and sector-specific neuronal loss in hippocampus during Alzheimer's disease. *Acta Neuropathologica*, **103**, 363-369.

Rosso, A., Mossey, J., and Lippa, C. F. (2008). Caffeine: Neuroprotective functions in cognition and Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, **23**, 417-422.

Rousseau, J. B., Van Lochem, P. B., Gispen, W. H., and Spruijt, B. M. (2000). Classification of rat behavior with an image-processing method and a neural network. *Behavioral Research Methods, Instrumentation, and Computers*, **32**, 63-71.

Rovio, S., Kareholt, I., Viitanen, M., Winblad, B., Tuomilehto, J., Soininen, H., Nissinen, A., and Kivipelto, M. (2007). Work-related physical activity and the risk of dementia and Alzheimer's disease. *International Journal of Geriatric Psychiatry*, **22**, 874-882.

Rowan, A. J., and Tolunsky, E. (2003). *Primer of EEG*. Philadelphia, PA: Elsevier.

Ruehl, W. W., Bruyette, D. S., DePaoli, A., Cotman, C. W., Head, E., Milgram, N. W., and Cummings, B. J. (1995). Canine cognitive dysfunction as a model for human age related cognitive decline, dementia and Alzheimer's disease: Clinical presentation, cognitive testing, pathology and response to L-deprenyl therapy. *Progress in Brain Research*, **106**, 217-225.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, **323**, 533-536.

Rumelhart, D. E., and McClelland, J. L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press.

Salmon, D. P., Thomas, R. G., Pay, M. M., Booth, A., Hofstetter, C. R., Thal, L. J., et al. (2002). Alzheimer's disease can be accurately diagnosed in very mildly impaired individuals. *Neurology*, **59**, 1022-1028.

Sanan, D., Weisgraber, K., Russell, S., Mahley, R., Huang, D., Saunders, A., Schmechel, D., Wisniewski, T., Frangione, B., Roses, A. and Strittmatter. (1994). Apolipoprotein E associates with beta amyloid peptide of Alzheimer's disease to form novel monofibrils. Isoform apoE4 associates more efficiently than apoE3. *Journal of Clinical Investigation*, **94**, 860-869.

Sango, K., McDonald, M. P., Crawley, J. N., Mack, M. L., Tifft, C. J., Skop, E., Starr, C. M., Hoffmann, A., Sandhoff, K., Suzuki, K., and Proia, R. L. (1996). Mice lacking both subunits of lysosomal beta-hexosaminidase display gangliosidosis and mucopolysaccharidosis. *Nature Genetics*, **14**, 348-352.

Savonenko, A., Xu, G. M., Melnikova, T., Morton, J. L., Gonzales, V., Wong, M. P. F., Price, D. L., Tang, F., Markowska, A. L., and Borchelt, D. R. (2005). Episodic-like memory deficits in the APPswe/PS1dE9 mouse model of Alzheimer's disease: Relationships to β-amyloid deposition and neurotransmitter abnormalities. *Neurobiology of Disease*, **18**, 602-617.

Scarmeas, N., Albert, S. M., Manly, J. J., and Stern, Y. (2006a). Education and rates of cognitive decline in incident Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, **77**, 308-316.

Scarmeas, N., Stern, Y., Mayeux, R., and Luschsinger, J. (2006b). Mediterranean diet, Alzheimer's disease, and vascular mediation. *Archives of Neurology*, **63**, 1709-1717.

Scarpini, E., Scheltens, P. and Feldman, H. (2003). Treatment of Alzheimer's disease: Current status and new prospectives. *Lancet Neurology*, **2**, 539-47.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501-518.

Scherder, E., Eggermont, L., Sergeant, J., and Boersma, F. (2007). Physical activity and cognition in Alzheimer's disease: Relationship to vascular risk factors, executive functions and gait. *Reviews of Neuroscience*, **18**, 149-158.

Schöllhorn, W. I. (2004). Applications of artificial neural nets in clinical biomechanics. *Clinical Biomechanics*, **19**, 876-898.

Schlosser, R., Hutchinson, M., Joseffer, S., Rusinek, H., Saarimaki, A., Stevenson, J., Dewey, S. L., and Brodie, J. D. (1998). Functional magnetic resonance imaging of human brain activity in a verbal fluency task. *Journal of Neurology, Neurosurgery and Psychiatry*, **64**, 492-498.

Schwab, C., Hosokawa, M., and McGeer, P. L. (2004). Transgenic mice overexpressing amyloid beta protein are an incomplete model of Alzheimer's disease. *Experimental Neurology*, **188**, 52-64.

Scott, A., and Wild, C. (1991). Transformations and r-squared. *American Statistician*, **45**, 127-128.

Scott, L. J., and Goa, K. L. (2000). Galantamine: A review of its use in Alzheimer's disease. *Drugs*, **60**, 1095-1122.

Seaman, J. W., Walls, S. C., Wide, S. E., and Jaeger, R. G. (1994). *Caveat emptor*: Rank transform methods and interactions. *Trends in Ecology and Evolution*, **9**, 261-263.

Selkoe, D. J. (1989). Biochemistry of altered brain proteins in Alzheimer's disease. *Annual Review of Neuroscience*, **12**, 463-490.

Selkoe, D. J. (2001). Alzheimer's disease: Genes, proteins, and therapy. *Physiological Reviews*, **81**, 741-766.

Selkoe, D. J., Bell, D. S., Podlisny, M. B., Price, D. L., and Cork, L. C. (1987). Conservation of brain amyloid proteins in aged mammals and humans with Alzheimer's disease. *Science*, **235**, 873-877.

Selkoe, D. J., and Podlisny, M. B. (2002). Deciphering the genetic basis of Alzheimer's disease. *Annual Review of Genomics and Human Genetics*, **3**, 67-99.

Serpell, L. (2000). Alzheimer's amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta*, **1502**, 16-30.

Serretti, A., Zanardi, R., Mandelli, L., Smeraldi, E., and Colombo, C. (2007). A neural network model for combining clinical predictors of antidepressant response in mood disorders. *Journal of Affective Disorders*, **98**, 239-245.

Sevush, S., Guterman, A., and Villalon, A. V. (1991). Improved verbal learning after outpatient physostigmine therapy in patients with dementia of the Alzheimer's type. *Journal of Clinical Psychiatry*, **52**, 300-303.

Sevush, S., Peruyera, G., Bertran, A., and Cisneros, W. (2003). A three-factor model of cognition in Alzheimer disease. *Cognitive and Behavioral Neurology*, **16**, 110-117.

Shallice, T., and Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neurophysiological study. *Quarterly Journal of Experimental Psychology*, **22**, 261-273.

Shenk, D. (2001). *The Forgetting. Alzheimers: Portrait of an Epidemic*. New York: Anchor.

Sherrington, R., Rogaev, E. I., Liang, Y., Rogaeva, E. A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K., Tsuda, T., Mar, L., Foncin, J. F., Bruni, A. C., Montesi, M. P., Sorbi, S., Rainero, I., Pinessi, L., Nee, L., Chumakov, I., Pollen, D., Brookes, A., Sanseau, P., Polinsky, R. J., Wasco, W., Da Silva, H. A. R., Haines, J. L., Pericak-Vance, M. A., Tanzi, R. E., Roses, A. D., Fraser, P. E., Rommens, J. M., St George-Hyslop, P. H. (1995), Cloning of a gene bearing mis-sense mutations in early-onset familial Alzheimer's disease. *Nature*, **375**, 754-760.

Shif, O., Gillette, K., Damkaoutis, C. M., Carrano, C., Robbins, S. J., and Hoffman, J. R. (2006). Effects of *Ginkgo biloba* administered after spatial learning on water maze and radial arm maze performance in young adult rats. *Pharmacology, Biochemistry and Behavior*, **84**, 17-25.

Shivers, B. D., Hilbich, C., Multhaup, G., Salbaum, M., Beyreuther, K., and Seeburg, P. H. (1988). Alzheimer's disease amyloidogenic glycoprotein: Expression pattern in rat brain suggests a role in cell contact. *EMBO Journal*, **7**, 1365-1370.

Shore, D. I., Stanford, L., MacInnes, W. J., Klein, R. M., and Brown, R. E. (2001). Of mice and men: Virtual Hebb-Williams mazes permit comparison of spatial learning across species. *Cognitive, Affective, and Behavioral Neuroscience*, **1**, 83-89.

Shortliffe, E. H. (1974). *MYCIN: A rule-based computer program for advising physicians regarding antimicrobial therapy selection*. Doctoral dissertation, Stanford University.

Shortliffe, E. H. (1976). *Computer-Based Medical Consultations: MYCIN*. New York: American Elsevier.

Sigut, J., Pineiro, J., Gonzalez, E., and Torres, J. (2007). An expert system for supervised classifier design: Application to Alzheimer's diagnosis. *Expert Systems with Applications*, **32**, 927-938.

Silvestrelli, G., Lanari, A., Parnetti, L., Tomassoni, D., and Amenta, F. (2006). Treatment of Alzheimer's disease: From pharmacology to a better understanding of disease pathophysiology. *Mechanisms of Ageing and Development*, **127**, 148-157.

Simard, A. R., Soulet, D., Gowing, G., Julien, J. P., and Rivest, S. (2006). Bone marrow-derived microglia play a critical role in restricting senile plaque formation in Alzheimer's disease. *Neuron*, **49**, 489-502.

Sitzer, D. I., Twamley, E. W., and Jeste, D. V. (2006). Cognitive training in Alzheimer's disease: A meta-analysis of the literature. *Acta Psychiatrica Scandinavica*, **114**, 75-90.

Sizova, D., Charbaut, E., Delalande, F., Poirier, F., High, A. A., Parker, F., Van Dorsselaer, A., Duchesne, M., and Diu-Hercend, A. (2007). Proteomic analysis of brain tissue from an Alzheimer's disease mouse model by two-dimensional difference gel electrophoresis. *Neurobiology of Aging*, **28**, 357-370.

Skarda, C. A., and Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral Brain Science*, **10**, 161-195.

Smith, A. (2002). Effects of caffeine on human behavior. *Food and Chemical Toxicology*, **40**, 1243-1255.

Sobow, T., Flirski, M., and Liberski, P. P. (2004). Amyloid-beta and tau proteins as biochemical markers of Alzheimer's disease. *Acta Neurobiologiae Experimentalis*, **64**, 53-70.

Song, J., Burrage, K., Yuan, Z., and Huber, T. (2006). Prediction of *cis/trans* isomerization in proteins using PSI-BLAST profiles and secondary structure information. *BMC Bioinformatics*, **7**, 124.

Souren, L., Franssen, E. and Reisberg, B. (1995). Contractures and loss of function in patients with Alzheimer's disease.*Journal of the American Geriatric Society*, **43**, 650-655.

Spalletta, G., Baldinetti, F., Buccione, I., Fadda, L., Perri, R., Scalmana, S., Serra, L., and Caltagirone, C. (2004). Cognition and behavior are independent and heterogeneous dimensions in Alzheimer's disease. *Journal of Neurology*, **251**, 688-695.

Spear, N. E., Miller, J. S., and Jagielo, J. A. (1990). Animal memory and learning. *Annual Reviews in Psychology*, **41**, 169-211.

Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology*, **15**, 201-293.

Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs*, **74**, 1-29.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195-231.

Squire, L. R., and Zola, S. M. (1997). Amnesia, memory and brain systems. *Philosophical Transactions of the Royal Society of London, Series B*, **352**, 1663-1673.

Stackman, R. W., Eckenstein, F., Frei, B., Kulhanek, D., Nowlin, J., and Quinn, J. F. (2003). Prevention of age-related spatial memory deficits in a transgenic mouse model of Alzheimer's disease by chronic *Ginkgo biloba* treatment. *Experimental Neurology*, **184**, 510-520.

Stein, T. D., and Johnson, J. A. (2002). Lack of neurodegeneration in transgenic mice overexpressing mutant amyloid precursor protein is associated with increased levels of transthyretin and the activation of cell survival pathways. *The Journal of Neuroscience*, **22**, 7380-7388.

Steinberger, D., Reynolds, D. S., Ferris, P., Lincoln, R., Datta, S., Stanley, J., Paterson, A., Dawson, G. R., and Flint, J. (2003). Genetic mapping of variation in spatial learning in the mouse. *Journal of Neuroscience*, **23**, 2426-2433.

Stelzmann, R. A., Schnitzlein, H. N., and Murtagh, F. R. (1995). An English translation of Alzheimer's 1907 paper, "Über eine eigenartige Erkrankung der Hirnrinde." *Clinical Anatomy*, **8**, 429-431.

Stern, Y., Gurland, B., Tatemichi, T. K., Tang, M. X., Wilder, D., and Mayeux, R. (1994). Influence of education and occupation on the incidence of Alzheimer's disease. *Journal of the American Medical Association*, **271(13)**, 1004-1010.

Stern, E. A., Bacskai, B. J., Hickey, G. A., Attenello, F. J., Lombardo, J. A., and Hyman, B. T. (2004). Cortical synaptic integration in vivo is disrupted by amyloid-beta plaques. *Journal of Neuroscience*, **24**, 4535-4540.

Stewart, W. F., Kawas, C., Corrada, M., and Metter, E. J. (1997). Risk of Alzheimer's disease and duration of NSAID use. *Neurology*, **48**, 626632.

St George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Haines, J. L., Nee, L., Watkins, P. C., Myers, R. H., Feldman, R. G., Pollen, D., Drachman, D., *et al.* (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science*, **235**, 885-890.

Stone, M. (1977). Asymptotics for and against cross-validation. *Biometrika*, **64**, 29-35.

Stopford, C. L., Snowden, J. S., Thompson, J. C., and Neary, D. (2007). Distinct memory profiles in Alzheimer's disease. *Cortex*, **43**, 846-857.

Strauss, M. E., and Fritsch, T. (2004). Factor structure of the CERAD neuropsychological battery. *Journal of the International Neuropsychological Society*, **10**, 559-565.

Stricker, L. J., and Rock, D. A. (1987). Factor structure of the GRE General Test in young and middle adulthood. *Developmental Psychology*, **23**, 526-536.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-661.

Stroustrup, B. (1985). *The C++ Programming Language*. New York: Addison-Wesley.

Sturchler-Pierrat, C., Abramowski, D., Duke, M., Wiederhold, K., Mistl, C., Rothacher, S., Ledermann, B., Burki, K., Frey, P., Paganetti, P., Waridel, C., Calhoun, M., Jucker, M., Probst, A., Staufenbiel, M., and Sommer, B. (1997). Two amyloid precursor protein transgenic mouse models with Alzheimer disease-like pathology. *Proceedings of the National Academy of Sciences USA*, **94**, 13287-13292.

Suh, Y.-H., and Checler, F. (2002). Amyloid precursor protein, presenilins, and alpha-synuclein: Molecular pathogenesis and pharmacological applications in Alzheimer's disease. *Pharmacological Reviews*, **54**, 469-525.

Suo, Z., Cox, A. A., Bartelli, N., Rasul, I., Festoff, B. W., Premont, R. T., and Arendash, G. W. (2007). GRK5 deficiency leads to early Alzheimer-like pathology and working memory impairment. *Neurobiology of Aging*, **28**, 1873-1888.

Suo, Z., Wu, M., Citron, B. A., Wong, G. T., and Festoff, B. W. (2004). Abnormality of G-protein-coupled receptor kinases at prodromal and early stages of Alzheimer's disease: An association with early beta-amyloid accumulation. *Journal of Neuroscience*, **24**, 3444-3452.

Suzuki, S., Augerinos, G., and Black, A. H. (1980). Stimulus control of spatial behavior on the eight-arm maze in rats. *Learning and Motivation*, **11**, 1-18.

Sweatt, J. (1999). Toward a molecular explanation for long-term potentiation. *Learning and Memory*, **6**, 399-416.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285-1293.

Swets, J. A., Dawes, R. M., and Monahan, J. (2000). Better decisions through science. *Scientific American*, **283(4)**, 82-87.

Tabachnick, B. G., and Fidell, L. S. (2001). *Using Multivariate Statistics*. New York: Harper Collins.

Taddei, K., Fisher, C., Laws, S. M., Martins, G., Paton, A., Clarnette, R. M., Chung, C., Brooks, W. S., Hallmayer, J., Miklossy, J., Relkin, N., St George-Hyslop, P. H., Gandy, S. E., and Martins, R. N. (2002). Association between presenilin-1 Glu318Gly mutation and familial Alzheimer's disease in the Australian population. *Molecular Psychiatry*, **7**, 776-781.

Tafeit, E., and Reibnegger, G. (1999). Artificial neural networks in laboratory medicine and medical outcome prediction. *Clinical Chemistry and Laboratory Medicine*, **37**, 845-853.

Takeuchi, A., Irizarry, M., Duff, K., Saido, T., Hsiao Ashe, K., Hasegawa, M., Mann, D., Hyman, B., and Iwatsubo, T. (2000). Age-related amyloid β deposition in transgenic mice overexpressing both Alzheimer mutant presenilin 1 and amyloid β precursor protein Swedish mutant is not associated with global neuronal loss. *American Journal of Pathology*, **157**, 331-339.

Tandon, R., Adak, S., and Kaye, J. A. (2006). Neural networks for longitudinal studies in Alzheimer's disease. *Artificial Intelligence in Medicine*, **36**, 245-255.

Tanzi, R. E., and Bertram, L. (2001). New frontiers in Alzheimer's disease genetics. *Neuron*, **32**, 181-184.

Tanzi, R. E., and Bertram, L. (2005). Twenty years of the Alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell*, **120**, 545-555.

Tanzi, R. E., Gusella, J. F., Watkins, P. C., Bruns, G. A., St George-Hyslop, P., Van Keuren, M. L., Patterson, D., Pagan, S., Kurnit, D. M., and Neve, R. L. (1987). Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science*, **235**, 880-884.

Tanzi, R. E., McClatchey, A. I., Lampert, E. D., Villa-Komaroff, L., Gusella, J. F., and Neve, R. L. (1988). Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease. *Nature*, **331**, 528530.

Tariot, P., Solomon, P., Morris, J., Kershaw, P., Lilienfeld, S. and Ding, C. (2000). A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology*, **54**, 2269-2276.

Tariot, P. N., Farlow, M. R., Grossberg, G. T., Graham, S. M., McDonald, S., and Gergel, I. (2004). Memantine treatment in patients with moderate to severe Alzheimer's disease already receiving donezepil: A randomized controlled trial. *Journal of the American Medical Association*, **291**, 317-324.

Tecott, L. H., and Nestler, E. J. (2004). Neurobehavioral assessment in the information age. *Nature Neuroscience*, **7**, 462-466.

Tekirian, T. L., Cole, G. M., Russell, M. J., Yang, F., Wekstein, D. R., Patel, E., Snowdon, D. A., Markesbery, W. R., and Geddes, J. W. (1996). Carboxy terminal of β-amyloid deposits in aged human, canine, and polar bear brains. *Neurobiology of Aging*, **17**, 249-257.

Thompson, P., Hayashi, K., de Zubicaray, G., Janke, A., Rose, S., Semple, J., Herman, D., Hong, M., Dittmer, S., Doddrell, D. and Toga, A. (2003). Dynamics of gray matter loss in Alzheimer's disease. *Journal of Neuroscience*, **23**, 994-1005.

Thompson, R. F., and Kim, J. J. (1996). Memory systems in the brain and localization of a memory. *Proceedings of the National Academy of Sciences USA*, **93**, 13438-13444.

Thorndike, R. L. (1935). Organization of behavior in the albino rat. *Genetic Psychology Monographs*, **17**, 1-70.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, **1**.

Tian, H., and Shang, Z. (2006). Artificial neural network as a classification method of mice by their calls. *Ultrasonics*, **44**, e275-e278.

Toffler, A. (1970). *Future Shock*. New York: Random House.

Toga, A. W., and Thompson, P. M. (2005). Genetics of brain structure and intelligence. *Annual Review of Neuroscience*, **28**, 1-23.

Tomidokoro, Y., Harigaya, Y., Matsubara, E., Ikeda, M., Kawarabayashi, T., Shirao, T., Ishiguro, K., Okamoto, K., Younkin, S. G., and Shoji, M. (2001). Brain Abeta amyloidosis in APPsw mice induces accumulation of presenilin-1 and tau. *Journal of Pathology*, **194**, 500-506.

Tomita, T., Maruyama, K., Saido, T. C., Kume, H., Shinozaki, K., Tokuhiro, S., Capell, A., Walter, J., Grunberg, J., Haass, C., Iwatsubo, T., and Obata, K. (1997). The presenilin 2 mutation (N141I) linked to familial Alzheimer disease (Volga German families) increases the secretion of amyloid beta protein ending at the 42nd (or 43rd) residue. *Proceedings of the National Academy of Science*, **94**, 2025-2030.

Touchon, J., Bergman, H., Bullock, R., Rapatz, G., Nagel, J., Lane, R. (2006). Response to rivastigmine or donepezil in patients with Alzheimer's disease and symptoms suggestive of concomitant Lewy body pathology. *Current Medical Research and Opinion*, **22**, 49-59.

Touzani, K., Puthanveettil, S. V., and Kandel, E. R. (2007). Consolidation of learning strategies during spatial working memory task requires protein synthesis in the prefrontal cortex. *Proceedings of the National Academy of Sciences USA*, **104**, 5632-5637.

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

Tulving, E. (1972). Episodic and semantic memory. Reprinted in: Tulving, E., and Donaldson, W. (Eds.) *Organization and Memory*. New York: Academic Press. pp. 381-403.

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, **40**, 385-398.

Übeyli, E. D. (2008). Multiclass support vector machines for diagnosis of erythematosquamous diseases. *Expert Systems with Applications*, **35**, 1733-1740.

Underwood, B. J., Boruch, R. F., and Malmi, R. A. (1978). Composition of episodic memory. *Journal of Experimental Psychology: General*, **107**, 393-419.

Valenti, P., Cazamajou, E., Scarpettini, M., Aizemberg, A., Silva, W., and Kochen, S. (2006). Automatic detection of interictal spikes using data mining models. *Journal of Neuroscience Methods*, **150**, 105-110.

Van Broeckhoven C, Backhovens H, Cruts M, De Winter G, Bruyland M, Cras P, and Martin J. J. (1992). Mapping of a gene predisposing to early-onset Alzheimer's disease to chromosome 14q24.3. *Nature Genetics*, **2**, 335-339.

Van Dam, D., DHooge, R., Staufenbiel, M., Van Ginneken, C., Van Meir, F., and De Deyn, P. (2003). Age-dependent cognitive decline in the APP23 model precedes amyloid deposition. *European Journal of Neuroscience*, **17**, 388-396.

Van Der Staay, F. J., and Steckler, T. (2001). Behavioral phenotyping of mouse mutants. *Behavioral Brain Research*, **125**, 3-12.

Van Gelder, B. M., Buijsse, B., Tijhuis, M., Kalmijn, S., Giampaoli, S., Nissinen, A., and Kromhout, D. (2007). Coffee consumption is inversely associated with cognitive decline in elderly European men: The FINE Study. *European Journal of Clinical Nutrition*, **61**, 226-232.

Vapnik, V. (1998). *Statistical Learning Theory*. New York: Wiley.

Vapnik, V., and Chervonenkis, A. (1964). A note on one class of perceptrons. *Automation and Remote Control*, **25**, 821-837.

Vapnik, V., and Lerner, A. (1963). Pattern recognition using generalized portrait method. *Automation and Remote Control*, **24**, 774-780.

Vassar, R., Bennett, B. D., Babu-Khan, S., Kahn, S., Mendiaz, E. A., Denis, P., Teplow, D. B., Ross, S., Amarente, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M. A., Biere, A. L., Curran, E., Burgess, T., Louis, J. C., Collins, F., Treanor, J., Rogers, G., and Citron, M. (1999). Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science*, **286**, 735-741.

Vaughan, W., Jr. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 36-42.

Vekovischeva, O. Yu., Verbitskaya, E. V., Aitta-aho, T., Sandnabba, K., and Korpi, E. R. (2007). Multimetric statistical analysis of behavior in mice selected for high and low levels of isolation-induced male aggression. *Behavioral Processes*. **75**, 23-32.

Velicer, W. F., and Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, **3**, 231-251.

Velicer, W. F., and Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, **25**, 1-28.

Venneri, A., Forbes-Mckay, K. E., and Shanks, M. F. (2005). Impoverishment of spontaneous language and the prediction of Alzheimer's disease. *Brain*, **128**, E27.

Vercelli, D. (2004). Genetics, epigenetics, and the environment: Switching, buffering, releasing. *Journal of Allergy and Clinical Immunology*, **113**, 286-381.

Verghese, J., Lipton, R., Katz, M., Hall, C., Derby, C., Kuslansky, G., Ambrose, A., Sliwinski, M., and Buschke, H. (2003). Leisure activities and the risk of dementia in the elderly. *New England Journal of Medicine*, **348**, 2508-2516.

Veurink, G., Fuller, S., Atwood, C., and Martins, R. (2003). Genetics, lifestyle, and the roles of amyloid beta and oxidative stress in Alzheimer's disease. *Annals of Human Biology*, **30**, 639-667.

Viera, A. J., and Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, **37**, 360-363.

Vinores, S., Xiao, W., Zimmerman, R., Whitcup, S. and Wawrousek, E. (2003). Upregulation of vascular endothelial growth factor (VEGF) in the retinas of transgenic mice overexpressing interleukin-1 beta (IL-1beta) in the lens and mice undergoing retinal degeneration. *Histology and Histopathology*, **18**, 797-810.

Waddington, C. H. (1953). Epigenetics and evolution. *Symposium of the Society for Experimental Biology*, **7**, 186-199.

Wahlsten, D., Cooper, S. F., and Crabbe, J. C. (2005). Different rankings of inbred mouse strains on the Morris maze and a refined 4-arm water escape task. *Behavioral Brain Research*, **165**, 36-51.

Walczak, S., and Cerpa, N. (1999). Heuristic principles for the design of artificial neural networks. *Information and Software Technology*, **41**, 107-117.

Walker, P. R., Smith, B., Liu, Q. Y., Famili, A. F., Valdes, J. J., Liu, Z., and Lach, B. (2004). Data mining of gene expression changes in Alzheimer brain. *Artificial Intelligence in Medicine*, **31**, 137-154.

Wall, P. M., and Messier, C. (2000). Ethological confirmatory factor analysis of anxiety-like behavior in the murine elevated plus-maze. *Behavioral Brain Research*, **114**, 199-212.

Walsh, D. M., and Selkoe, D. J. (2004). Oligomers on the brain: The emerging role of soluble protein aggregates in neurodegeneration. *Protein and Peptide Letters*, **11**, 213-228.

Wang, J., Ikonen, S., Gurevicius, K., van Groen, T., and Tanila, H. (2002). Alteration of cortical EEG in mice carrying mutated human APP transgene. *Brain Research*, **943**, 181-190.

Watanabe, S., Sakamoto, J., and Wakita, M. (1995). Pigeons' discrimination of paintings by Monet and Picasso. *Journal of the Experimental Analysis of Behavior*, **63**, 165-174.

Webster, S., Lue, L. F., Brachova, L., Tenner, A. J., McGeer, P. L., Terai, K., Walker, D. G., Bradt, B., Cooper, N. R., and Rogers, J. (1997). Molecular and cellular characterization of the membrane attack complex, C5b-9, in Alzheimer's disease. *Neurobiology of Aging*, **18**, 415-421.

Weggen, S., Eriksen, J. L., Das, P., Sagi, S. A., Wang, R., Pietrzik, C. U., Findlay, K. A., Smith, T. E., Murphy, M. P., Bulter, T., Kang, D. E., Marquez-Sterling, N., Golde, T. E., and Koo, E. H. (2001). A subset of NSAIDs lower amyloidogenic Abeta42 independently of cyclooxygenase activity. *Nature*, **414**, 212216.

Weiner, M. F., Hynan, L. S., Bret, M. E., and White, C. (2005). Early behavioral symptoms and course of Alzheimer's disease. *Acta Psychiatrica Scandinavica*, **111**, 367-371.

Weinhold, B. (2006). Epigenetics: The science of change. *Environmental Health Perspectives*, **114**, A160-A167.

Welsh, K., Butters, N., Hughes, J., Mohs, R., and Heyman, A. (1991). Detection of abnormal memory decline in cases of Alzheimer's disease using CERAD neurological measures. *Archives of Neurology*, **48**, 278-281.

Wesnes, K., Simpson, P. M., and Kidd, A. G. (1988). An investigation of the range of cognitive impairments induced by scopolamine 0.6 mg. *Human Psychopharmacology*, **3**, 27-41.

Westbury, C., Buchanan, L., Sanderson, M., Rhemtulla, M., and Phillips, L. (2003). Using genetic programming to discover nonlinear variable interactions. *Behavior Research Methods, Instruments, and Computers*, **35**, 202-216.

Westerman, M., Cooper-Blacketer, D., Mariash, A., Kotilinek, L., Kawara-bayashi, T., Younkin, L., Carlson, G., Younkin, S., and Ashe, K. (2002). The relationship between Aβ and memory in the Tg2576 mouse model of Alzheimer's disease. *Journal of Neuroscience*, **22**, 1858-1567.

Whelihan, W. M., Thompson, J. A., Piatt, A. L., Caron, M. D., and Chung, T. (1997). The relation of neuropsychological measures to levels of cognitive functioning in elderly individuals: A discriminant analysis approach. *Applied Neuropsychology*, **4**, 160-164.

Whishaw, I. Q., and Tomie, J. A. (1996). Of mice and mazes: Similarities between mice and rats on dry land but not water mazes. *Physiology and Behavior*, **60**, 1191-1197.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika*, **24**, 471-494.

Willingham, D. B., and Koroshetz, W. J. (1993). Evidence for dissociable motor skills in Huntington's disease patients. *Psychobiology*, **21**, 173-182.

Wilson, R. S., Bienias, J. L., Berry-Kravis, E., Evans, D. A., and Bennett, D. A. (2002). The apolipoprotein E ϵ2 allele and decline in episodic memory. *Journal of Neurology, Neurosurgery and Psychiatry*, **73**, 672-677.

Wilson, R. S., Li, Y., Aggarwal, N. T., Barnes, L. L., McCann, J. J., Gilley, D. W., and Evans, D. A. (2004). Education and the course of cognitive decline in Alzheimer's disease. *Neurology*, **63**, 1198-1202.

Wilson, R. S., Scherr, P. A., Schneider, J. A., Tang, Y., and Bennett, D. A. (2007). The relation of cognitive activity to risk of developing Alzheimer disease. *Neurology*, **69**, 1911-1920.

Winblad, B., Engedal, K., Soininen, H., verhey, F., Waldemar, G., Wimo, A., Wetterholm, A., Zhang, R., Haglund, A., Subbiah, P. and the Donepezil Nordic Study Group. (2001). A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology*, **57**, 489-495.

Winocur, G. (1984). The effects of retroactive and proactive interference on learning and memory in old and young rats. *Developmental Psychobiology*, **17**, 537-545.

Wisniewski, T., Castano, E., Golabek, A., Vogel, T. and Frangione, B. (1994). Acceleration of Alzheimer's fibril formation by apolipoprotein E *in vitro*. *American Journal of Pathology*, **145**, 1030-1035.

Witten, I. H., and Frank, E. (2000). *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, 2nd Ed.* New York: Morgan Kaufmann Publishers.

Wolfer, D. P., Stagljar-Bozicevic, M., Errington, M. L., and Lipp, H.-P. (1998). Spatial memory and learning in transgenic mice: Fact or artifact? *News in Physiological Sciences*, **13**, 118-123.

Wolpert, D. H., and Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, **1**, 67-82.

Wong, P. C., Cai, H., Borchelt, D. R., and Price, D. L. (2002). Genetically engineered mouse models of neurodegenerative diseases. *Nature Neuroscience*, **5**, 633-639.

Woodruff-Pak, D. S., Vogel, R. W., and Wenk, G. L. (2001). Galantamine: Effect on nicotinic receptor binding, acetylcholinesterase inhibition, and learning. *Proceedings of the National Academy of Sciences USA*, **98**, 2089-2094.

Wright, J. W., Alt, J. A., Turner, G. D., and Krueger, J. M. (2004). Differences in spatial learning comparing transgenic p75 knockout, New Zealand Black, C57BL/6, and Swiss Webster mice. *Behavioral Brain Research*, **153**, 453-458.

Wu, J., Basha, M. R., and Zawia, N. H. (2008a). The environment, epigenetics and amyloidogenesis. *Journal of Molecular Neuroscience*, **34**, 1-7.

Wu, J., Basha, M. R., Brock, B., Cox, D. P., Cardozo-Pelaez, F., McPherson, C. A., Harry, J., Rice, D. C., Maloney, B., Chen, D., Lahiri, D. K., and Zawia, N. H. (2008b). Alzheimer's disease (AD)-like pathology in aged monkeys after infantile exposure to environmental metal lead (Pb): Evidence for a developmental origin and environmental link for AD. *Journal of Neuroscience*, **28**, 3-9.

Wyss-Coray, T. (2006). Inflammation in Alzheimer disease: Driving force, bystander or beneficial response? *Nature Medicine*, **12**, 1005-1015.

Xiao, Q. X., Wang, R., and Xi, C. T. (2000). Huperzine A and tacrine attenuate beta-amyloid peptide-induced oxidative injury. *Journal of Neuroscience Research*, **61**, 564-569.

Xu, Y., Jack, C. R., OBrien, P. C., Kokmen, E., Smith, G. E., Ivink, R. J., Boeve, B. F., Tangalos, R. G., and Petersen, R. C. (2000). Usefulness of MRI measures of entorhinal cortex versus hippocampus in Alzheimer's disease. *Neurology*, **54**, 1760-1767.

Yamaguchi, H., Hirai, S., Morimatsu, M., Shoji, M., and Harigaya, Y. (1988). Diffuse type of senile plaques in the brains of Alzheimer-type dementia. *Acta Neuropathologica*, **77**, 113-119.

Yao, J., Petanceska, S. S., Montine, T. J., Holtzman, D. M., Schmidt, S. D., Parker, C. A., Callhan, M. J., Lipinski, W. J., Bisgaier, C. L., Turner, B. A., Nixon, R. A., Martins, R. N., Ouimet, C., Smith, J. D., Davies, P., Laska, E., Ehrlich, M. E., Walker, L. C., Mathews, P. M., and Gandy, S. (2004). Aging, gender, and ApoE isotype modulate metabolism of Alzheimer's Aβ peptides and F-isoprostanes in the absence of detectable amyloid deposits. *Journal of Neurochemistry*, **90**, 1011-1018.

Yoshikai, S., Sasaki, H., Doh-ura, K., Furuya, H., and Sakaki, Y. (1990), Genomic organization of the human amyloid beta-protein precursor gene. *Gene*, **87**, 257-263.

Zandi, P. P., Anthony, J. C., Hayden, K. M., Mehta, K., Mayer, L., Breitner, J. C., and Cache County Study Investigators. (2002). Reduced incidence of AD with NSAID but not H2 receptor antagonists: The Cache County Study. *Neurology*, **59**, 880-886.

Zangara, A. (2003). The psychopharmacology of huperzine A: An alkaloid with cognitive enhancing and neuroprotective properties of interest in the treatment of Alzheimer's disease. *Pharmacology Biochemistry and Behavior*, **75**, 675-686.

Zhang, G. P. (2000). Neural networks for classification: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, **30**, 451-462.

Zhuo, J.-M., Prescott, S. L., Murray, M. E., Zhang, H.-Y., Baxter, M. G., and Nicolle, M. M. (2007). Early discrimination reversal learning impairment and preserved spatial learning in a longitudinal study of Tg2576 APPsw mice. *Neurobiology of Aging*, **28**, 1248-1257.

Zillmer, E. A., Fowler, P. C., Gutnick, H. N., and Becker, E. (1990). Comparison of two cognitive bedside screening instruments in nursing home residents: A factor analytic study. *Journal of Gerontology*, **45**, P69-P74.

Zimmer, C. (2008, Oct). The search for intelligence. *Scientific American*, **299(4)**, 68-75.

Zimmermann, M., Colciaghi, F., Cattabeni, F., and Di Luca, M. (2002). *Ginkgo biloba* extract: From molecular mechanisms to the treatment of Alzheimer's disease. *Cellular and Molecular Biology*, **48**, 613-623.

Zubenko, G. S., Hughes, H. B., III, and Stiffler, J. S. (2001). D10S1423 identifies a susceptibility locus for Alzheimer's disease in a prospective, longitudinal, double-blind study of asymptomatic individuals. *Molecular Psychiatry*, **6**, 413-419.

Zweig, M. H., and Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, **39**, 561-577.

Zwick, W. R., and Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, **99**, 432-442.

ABOUT THE AUTHOR

Ralph E. Leighty attended the University of Florida (Gainesville) as an undergraduate, where he completed Bachelor's degrees in Chemistry (1985) and Psychology (1988), followed by postbaccalaureate study in Pharmacy and Neurobiological Sciences. In 1991, he relocated to Tampa, Florida to pursue graduate studies in Experimental Psychology at the University of South Florida. He transferred to the College of Engineering in 1995 to study artificial intelligence and robotics, and received a Master's in Computer Science (1997). In 2001, he resumed his graduate studies, earning a Master's in Biology (2003) with a thesis on behavioral analysis of strain differences and Alzheimer-like transgenic effects in mice. His current research interests include behavioral neurobiology, computational neuroscience, and nonlinear analysis. In addition to his academic activities, he is an actor/comedian, musician (piano, synthesizers), and chess enthusiast. He currently resides in Tampa with his wife, Debby, and their two cats, Jo and Meg.