



Published in final edited form as:

Neurobiol Aging. 2022 January ; 109: 88–99. doi:10.1016/j.neurobiolaging.2021.09.015.

Sex differences in cognitive aging: a 4-year longitudinal study in marmosets

Emily S. Rothwell^{a,*}, Kathryn P. Workman^a, Dongwei Wang^a, Agnès Lacreuse^{a,b}

^aDepartment of Psychological & Brain Sciences, University of Massachusetts, Amherst, MA, 01003, USA

^bNeuroscience & Behavior Graduate Program, University of Massachusetts, Amherst, MA, 01003, USA

Abstract

Longitudinal studies are essential to understand healthy and pathological neurocognitive aging such as Alzheimer's Disease, but longitudinal designs are rare in both humans and non-human primate models of aging because of the difficulty of tracking cognitive change in long-lived primates. Common marmosets (*Callithrix jacchus*) are uniquely suited for aging studies due to their naturally short lifespan (10–12 years), sophisticated cognitive and social abilities and Alzheimer Disease-like neuropathology. We report the first longitudinal study of cognitive aging in marmosets (N=28) as they transitioned from middle- (~ 5 years) to old age (~ 9 years). We characterized aging trajectories using reversal learning with different stimuli each year. Marmosets initially improved on cognitive performance due to practice, but worsened in the final year, suggesting the onset of age-related decline. Cognitive impairment emerged earlier in females than males and was more prominent for discrimination than for reversal learning. Sex differences in cognitive aging could not be explained by differences in motivation or motor abilities, which improved or remained stable across aging. Likewise, males and females did not differ in aging trajectories of overall behavior or reactivity to a social stressor, with the exception of a progressive decline in the initiation of social behavior in females. Patterns of cognitive aging were highly variable across marmosets of both sexes, suggesting the potential for pathological aging for some individuals. Future work will link individual cognitive trajectories to neuropathology in order to better understand the relationships between neuropathologic burden and vulnerability to age-related cognitive decline in each sex.

Keywords

aging; nonhuman primate; marmoset; reversal learning; sex difference; cognition

*Corresponding author: erothwell@umass.edu.

Declarations of Interest

None.

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

1. Introduction

Nearly 13 million people will have Alzheimer's Disease (AD) in the US by 2050 (Alzheimer's Association, 2019). Despite these alarming statistics, factors that drive an individual onto a normal vs. a pathological aging trajectory remain largely unknown. Longitudinal designs, which follow the same individuals over time, are far less common than cross-sectional studies, but are essential for identifying susceptibility to or protection from pathological aging processes (McQuail et al., 2021; Nyberg et al., 2020; Raz & Lindenberger, 2011). Longitudinal studies of cognitive aging have shown that variability in cognitive performance becomes greater at older ages (Nyberg et al., 2020). Interestingly, they also point to differences in cognitive aging processes based on biological sex, but these studies are rare and provide mixed results, with some suggesting that men decline faster than women (McCarrey et al., 2016) while others report the reverse (Levine et al., 2021). The complexity of human cognitive aging is even greater when considering sociocultural factors that are widely acknowledged to influence aging trajectories including gender, race and their interaction (Avila et al., 2019; Rahman et al., 2019). Longitudinal studies of neurocognitive aging in laboratory animals living in controlled environments are necessary to isolate the variables that critically influence cognitive health across the lifespan.

Due to their similarity with humans in brain organization and function, behavior and aging patterns, non-human primates (NHPs) are highly valuable to study neurocognitive aging (Baxter, 2001; Kaas, 2021; Lacreuse & Herndon, 2009; Moss et al., 1999; Voytko & Tinkler, 2004). However, longitudinal studies in NHPs remain sparse due to the difficulty of assessing cognitive change in long-lived species such as the rhesus macaque (up to 42 years old), the traditional model of human cognitive aging. Shorter-lived NHP species have long been proposed as complementary models for human aging (Fischer & Austad, 2011). With an average lifespan of 10–12 years (Nishijima et al., 2012), and signs of aging appearing around age eight (Abbott et al., 2003), marmosets are ideally suited for longitudinal studies. They also show similar age-related changes as humans in a number of biological systems, including the cardiovascular, renal, gastrointestinal, metabolic and muscular systems (Ross, 2019; Ross et al., 2012). The marmoset has a brain organization typical of primates (Kaas, 2021) and its brain also exhibits marked age-related changes, including reduced neurogenesis in the dentate gyrus of the hippocampus and development of neuropathology relevant to AD, including deposition of amyloid- β (A β) proteins and accumulation of hyperphosphorylated tau and dystrophic microglia (Arnsten et al., 2021; Geula et al., 2002; Maclean et al., 2000; Rodriguez-Callejas et al., 2016; Rothwell et al., 2021). Moreover, marmosets have highly developed cognitive abilities (Huber & Voelkl, 2009; Spinelli et al., 2004), a rich behavioral repertoire (Miller, 2017), and a high level of prosocial behavior (Burkart & van Schaik, 2020), thus providing characteristics that are highly relevant for studying human aging. Only a few studies have investigated age-related cognitive decline in this species. Three cross-sectional studies reported age-related differences in working memory (Sadoun et al., 2019) and executive function (Munger et al., 2017; Phillips et al., 2019; Sadoun et al., 2019).

The current study is the first comprehensive analysis of longitudinal cognitive change in male and female marmosets. We studied a cohort of animals as they aged from middle

(~ 5 years old) to old age (~9 years old). We aimed to characterize cognitive aging trajectories using a reversal learning paradigm to evaluate changes in executive function across aging. Executive function, which encompasses several cognitive abilities dependent on the prefrontal cortex (PFC) including updating, task switching, planning, and inhibiting (Miyake et al., 2000) shows robust age-related decline in humans and many other primates and can be considered as a reliable marker of cognitive aging (Lacreuse et al., 2020). We also characterized aging trajectories for comprehensive behavioral assessments including motor competency, behavioral profiles and social stress reactivity.

2. Material and methods

2.1 Subjects

We studied 28 adult common marmosets (*Callithrix jacchus*) aged between 4 and 6 years old at the start of the study (14 females, mean age at start = 4.81; 14 males, mean age at start = 5.14 years) for 4 consecutive years. Three females died of natural causes through the course of the study. One female was removed from the study due to unexpected pregnancy. Six additional animals were humanely euthanized in Year 4 for analysis of brain tissues. Details about subjects and participation in particular tests are provided in Table A.1. Marmosets were housed in opposite sex pairs and males were vasectomized as adults prior to the start of testing to prevent breeding. Thus both males and females maintained their natural endocrine milieu throughout the study. The marmosets were housed at the University of Massachusetts, Amherst under a 12:12 light cycle in steel mesh cages (182 × 54 × 75 cm) outfitted with a nest box, hammock and perches or platforms. They were fed twice daily of a diet made out of fresh fruits, vegetables, nuts and seeds, various breads and ZuPreem marmoset food. Food and water were available *ad libitum* except during two hours prior to cognitive testing. Daily enrichment of foraging toys were provided to all animals. Pairs and housing conditions were consistent across the 4 years of study for most pairs (10 pairs) but repairing was necessary in certain cases (e.g. death of partner). In these instances, testing was paused and resumed following a period of acclimation. The marmosets were cared for in accordance with Guide for the Care and Use of Laboratory Animals (8th edition, 2011) and all testing described was approved by the University of Massachusetts Amherst Institutional Animal Care and Use Committee. Data from this publication are available via Open Access (<https://doi.org/10.5281/zenodo.5497708>).

2.2 Study Timeline

Testing occurred between 2016 and 2020 as outlined in Figure 1A. We collected data to assess cognitive performance (reversal learning of 3 pairs of stimuli per year across 4 years), motor function (once per year across 4 years), stress reactivity (once per year across 3 years) and homecage behavior (weekly across 4 years). The tests of motor function (Hill and Valley tests) were conducted concurrently with the end of cognitive testing within each year. Stress reactivity was assessed through a Social Separation Test that took place one day per year and did not overlap with any other tests for that animal. Some of the methods below have been described in previous publications from these subjects regarding cognitive performance during the first year (Laclair et al., 2019; Lacreuse et al., 2018; Nephew et al., 2020) and first two years (Workman et al., 2018) of data collection.

2.3 Cognitive Assessments

We used reversal learning as the main measure of cognitive performance to be administered over 4 years. Reversal learning is a test of executive function that assesses cognitive flexibility, or the ability to adapt to changing reward-stimulus contingencies. Reversal learning has been well studied in NHPs including marmosets (Kangas et al., 2016; LaClair & Lacreuse, 2016; Lacreuse et al., 2014; Munger et al., 2017; Pryce et al., 2004; Ridley et al., 1981; Sadoun et al., 2019; Strasser & Burkart, 2012; Takemoto et al., 2015) and involves a network of brain regions including the orbitofrontal cortex, caudate and striatum (Clarke et al., 2005; Clarke et al., 2011; Collins et al., 2000; Jackson et al., 2018; Roberts et al., 1990). In year 1, a subset of monkeys were also tested on intradimensional/extradimensional set shifting ($n = 17$) (Laclair et al., 2019), and progressive ratio ($n = 16$) (Carlotto, unpublished results) but these tests were not administered in subsequent years and are not reported here.

Reversal learning is comprised of two testing phases: discrimination and reversal. Cognitive performance was measured by trials to criterion (criterion = 90% on 40 consecutive trials) for discrimination and reversal, and was completed either on a touchscreen interface (Monkey CANTAB) or on a manual version using a modified Wisconsin General Testing Apparatus (WGTA, Figure 1B). Marmosets were presented with three stimulus pairs per year (depicted in Figure 1A). After reaching criterion on discrimination for a stimulus pair, the reward contingencies of the stimulus pair were reversed and the marmoset began the reversal phase for the same pair on the next testing day. After reaching criterion on reversal the marmoset moved on to the second stimulus pair of the year until completion of the 3 pairs. For all tests criterion was set at 90% (36 correct responses) on 40 consecutive trials. Yearly scores for each marmoset were averaged across the three stimulus pairs presented per year. Years one through three had one color mis-matched pair followed by two color-matched pairs of stimuli. All stimuli in year four were color-matched, but we found a similar pattern of results when only the color-matched pairs were analyzed (Figure A.1, Table A.2).

2.3.1 CANTAB—We used a NHP version of the Cambridge Neuropsychological Test Automated Battery (Monkey CANTAB with Liquid Reward, Model 80951A), which consists of a touch screen panel (37.78 cm) housed in stainless-steel frame ($56 \times 38 \times 30$ cm) using Intel based 1.6 GHz CPU operating system. In front of the touch screen panel is a stainless-steel sipper tube that delivers reward (banana milkshake) via a peristaltic pump at rate of 0.2 ml per second. All animals were trained to use the cognitive testing apparatus prior to the beginning of data collection. Training followed procedures detailed in (Roberts et al., 1988) and previously described for this group of marmosets (Laclair et al., 2019; Workman et al., 2018).

Food and water were removed from the marmoset homecages approximately two hours prior to cognitive testing. To begin a cognitive test session, a marmoset entered a transport box ($34.1 \times 20.65 \times 30.8$ cm) affixed to its homecage. The CANTAB apparatus was then placed in front of the transport box so that the marmoset could reach through a metal mesh (2.5×2.5 cm openings) to access the touch screen. During the test session the marmoset had visual, auditory and olfactory access to their partner in the homecage and other marmosets in their colony room. A solid, opaque door on the back of the transport box was closed so

that they could not go back to their homecage until after the test session had ended. If a marmoset appeared distracted by other animals during testing, a cover was added over the transport box to facilitate attention to the cognitive task.

To start a test session, the reversal learning program was loaded to the CANTAB apparatus from a remote computer outside of the marmoset colony room. For the discrimination, marmosets were presented with a pair of stimuli on the CANTAB screen appearing in random positions on the screen at each trial. Pairs of stimuli were selected from a set of images provided by the CANTAB, but we avoided certain color combinations (i.e. red/green/brown/orange), as female marmosets can be di- or tri-chromats but all males are dichromats (Caine et al., 2010; Pessoa et al., 2005). In each pair, one stimulus was associated with a reinforcement (banana milkshake), while the other was incorrect. The rewarded stimulus was counterbalanced across all marmosets such that half of the marmosets were rewarded for one stimulus whereas the other half were rewarded for the opposite stimulus. Marmosets were offered 40 trials per day and five test sessions per week to acquire discrimination. Once a marmoset reached criterion on discrimination, the reward-stimulus contingencies were reversed (reversal). When a marmoset reached criterion in the reversal, the following test day the marmoset moved on to the next pair of stimuli for that testing year.

2.3.2 Wisconsin General Testing Apparatus—As commonly observed in cognitive studies (Schubiger et al., 2019), and computerized tests in particular (Nakamura et al., 2018; Takemoto et al., 2015), a subset of monkeys were unable to complete training on CANTAB. While one monkey did not participate in any task, four marmosets (detailed in Table A.1) completed cognitive testing using a manual version of reversal learning administered with a modified Wisconsin General Testing Apparatus (WGTA) (Harlow, 1949). The WGTA consists of an opaque box ($43.2 \times 42.3 \times 44.5$ cm) with test tray ($40.65 \times 11.15 \times 1.25$ cm) containing two wells (diameter 2.5 cm) that can be baited with food rewards. The wells were covered with stimulus objects ($2 \times 2 \times 0.9$ cm³) that were the same shape and color of the CANTAB stimuli (Figure 1A) but were made of a foam-like material and mounted on a token (diameter 3.5 cm). When in place, the stimulus and token completely covered the food well. An opaque plexiglass door affixed between the marmoset and the WGTA testing box allowed the experimenter to permit or limit access to the stimulus objects during the test session.

To begin a session, the experimenter closed the opaque door and baited one well with a food reward (mini dried marshmallow, 6 mm diameter), covered it with the correct stimulus, while simultaneously covering the empty well with the incorrect stimulus. The experimenter then lifted the opaque door to start the trial and let the marmoset displace one of the stimuli. The experimenter recorded whether the marmoset had chosen the correct (coded as 1) or incorrect (coded as 0) stimulus. The response time was recorded as the time from the door opening to the marmoset's choice. If the marmoset did not choose within two minutes it was coded as omitted and the next trial was administered. Between each trial the experimenter closed the opaque door to set up the next trial.

Marmosets on WGTA were offered 20 trials per day, five days per week. The rewarded stimulus was counterbalanced across the four marmosets. The experimenter followed a test

sheet indicating a randomized location (left vs. right) for the correct stimulus. Once a marmoset reached criterion, the reward contingencies were reversed and the marmoset began the reversal task. When a marmoset reached criterion on reversal, he or she moved on to the next pair for that testing year.

2.3.4 Preliminary Analyses & Cognitive Variables—We assessed differences in trials to criterion based on whether marmosets completed cognitive testing with the CANTAB touchscreen or the WGTA. We conducted general linear mixed models (GLMM) with fixed effects for Test Type (WGTA, CANTAB), Year and a Test Type*Year Interaction and a random effect of individual marmoset. The GLMMs revealed a significant effect of Test Type for discrimination ($F(1, 213) = 5.94, p = 0.015$) and reversal ($F(1, 207) = 4.26, p = 0.04$). Planned comparisons of WGTA vs. CANTAB at each testing year showed that WGTA score were significantly worse than CANTAB only in Year 4 for both discrimination ($t(213) = -2.55, p = 0.01$) and reversal ($t(207) = -2.91, p = 0.004$). Due to sample size attrition, the WGTA dataset in Year 4 included only two marmosets (1 female, 1 male). The effect of Test Type in Year 4 was likely due to chance, as the scores for these marmosets fell within their individual aging trajectories. Therefore, the CANTAB and WGTA data were combined in subsequent analyses. Moreover, the aging trajectories reported in sections 3.1.1 and 3.1.2 remain consistent when only the CANTAB data are analyzed.

We measured cognitive performance using the number of trials to reach criterion and the number of errors made to reach criterion for both discrimination and reversal. We averaged across the three stimulus sets within each year to obtain a yearly score for each marmoset on each cognitive performance variable. We also analyzed two measures that evaluate the motivation to complete the cognitive task: response latency and omissions. Response latency is the time between the presentation of the cognitive stimuli on the CANTAB screen and the marmoset's selection of the correct or incorrect stimulus. We calculated a proportion of omissions as proportion of trials a marmoset declined to respond to out of 40 total trials presented to that marmoset for that specific test. Due to differences in the experimental set-up, motor latencies were much longer and omissions less frequent in the WGTA than CANTAB. Therefore, these analyses excluded the WGTA monkeys.

2.4 Motor Assessments

To assess motor competency, we utilized the Hill and Valley task which has been used to measure motor ability in marmoset models of Parkinson's disease (Bihel et al., 2010; Eslamboli et al., 2003; Marshall & Ridley, 2003; Phillips et al., 2017). Hand preference, which has been shown to be stable over time (Hook & Rogers, 2000), was assessed in Year 1.

2.4.1 Hill & Valley Apparatuses—Two apparatuses were used for motor testing, each made out of clear plexiglass containing two staircases ($9 \times 9 \times 3$ mm) with five steps each (Figure 1C). The Hill task has two lateral openings (7.4×2 mm) on the right and left side with one staircase near each opening. The Valley task has one central opening (7.7×2 cm) with two staircases rising on either side. All animals were trained to use the motor testing apparatuses prior to the beginning of data collection. Training followed procedures

previously described for this group of marmosets (Laclair et al., 2019; Workman et al., 2018).

To begin a motor test session, a marmoset jumped into a transport box affixed to their homecage and transported to a nearby cart in the colony room. The Hill or Valley apparatus, filled with mini-dried marshmallows on each step, was then affixed to the front of the transport box. Then the experimenter lifted an opaque door so that the marmoset could access the apparatus. Marmosets had to use one hand to retrieve the food rewards from each step. In the Hill version, the monkey had to use the right hand to reach the right staircase and the left hand for the left staircase. In the Valley apparatus, the right hand had to be used to reach the left staircase and vice versa. Each marmoset had five minutes maximum to retrieve all five marshmallows from one staircase. Each marmoset participated in all four conditions (Hill Left, Hill Right, Valley Left, Valley Right) in one test session and were given one test session per day across three days. The order of testing was counterbalanced so that on the first testing day half of the marmosets participated in the Hill version and the other half participated on the Valley apparatus. The order of conditions was randomized across test sessions. Each session was video recorded and a trained experimenter decoded the video playbacks to score accuracy and latency to complete the task.

2.4.2 Preliminary Analyses & Motor variables—Maximum accuracy score was 15 points per hand with one point counted for retrieving the marshmallow on the first step, 2 points for the second step and so on until 5 points for the fifth step. If a marshmallow was dropped then one point was deducted. The latency to retrieve all five food rewards was calculated for each test. A GLMM revealed that motor performance of left vs. right hands did not differ for accuracy ($F(1, 128) = 0.66, p = 0.42$) or latency ($F(1, 91) = 0.78, p = 0.38$). Likewise, motor performance of preferred vs. non-preferred hand did not differ for accuracy ($F(1, 128) = 0.84, p = 0.36$) or latency ($F(1, 91) = 0.33, p = 0.57$). Therefore, we took averaged performance scores of accuracy and latency across each hand for each year of testing.

2.5 Homecage Behavior

Behavior observations were recorded twice weekly for each marmoset consisting of one morning and one afternoon observation. Trained observers completed 5-min focal animal observations using the ethogram in Table A.3 and noted a 0/1 for each behavior during 15 second intervals. We summarized behavior quarterly for each year by taking the sum of observed intervals (i.e. 20 15-s intervals per 5-min observation) across 3 months of each year. We then calculated the percentage of observations of the behavior of interest among the total observations. We evaluated calm locomotion, agitated locomotion and two composite behavior scores: Initiate Social Behavior was a composite that included initiating the following behaviors with the cagemate: social contact, sniff/nuzzle, groom, mount, tongue flick, or social play. Agonistic Displays was a composite score that included the following behaviors: scent mark, gouge, piloerection and genital display.

2.6 Social Separation Test

Social separation in NHPs, including marmosets, reliably activates behavioral and physiological distress responses (Cacioppo et al., 2015; Hennessy, 1997). One procedure previously validated for marmosets is the Social Separation Test (French et al., 2007) where an individual marmoset is housed alone away from the colony room for 7 hours during which distress behaviors can be observed along with elevated hypothalamic-pituitary-adrenal (HPA) axis activity, measured via urinary cortisol excretion, both subsiding upon return to the homecage and reunion with the partner. The Social Separation Test occurred on a single test day per year for Years 1, 2 and 3.

The social separation test started at approximately 0800 AM, when the test marmoset was first observed in the homecage prior to separation and we collected a baseline behavioral observation for 5 min as well as a BL urine sample to assess urinary cortisol levels (Saltzman et al., 1998). Then the test marmoset was removed from the homecage (Saltzman et al., 1998) and colony room and housed alone in a similar cage for 7 hours with food and water available *ad libitum*. During these 7 hours, the marmoset was not in visual, auditory and olfactory communication with any other marmosets. Each session was videotaped. Three 5-min behavior observations were collected during separation (first 5 min, 3.5 hours after separation, final 5 min) and compiled to create a total separation score. Experimenters entered the room once each hour to collect any urine that was available on a catch pan underneath the marmoset's cage. After 7 hours had elapsed, the marmoset was returned to the homecage. The following day, a urine sample was collected again at approximately 0800 AM to assess cortisol concentrations.

2.6.1 Urinary Cortisol Assays.—All urine samples were pipetted into 1.5 ml vials, spun in a centrifuge a minimum of 5 min, transferred to a new 1.5 ml vial and stored frozen at -20°C . Urine samples from Year 1 were processed by the Endocrine BioServices Assay Lab at the University of Nebraska Omaha. Urine samples from Year 2 and 3 were processed by the Assay Services at the Wisconsin National Primate Research Center.

2.6.2 Social Separation Test Variables—We assessed behavioral reactivity to the Social Separation Test by taking a difference score between each individual's total score during the Social Separation Test and subtracting it from that individual's pre-separation baseline. This provided scores with negative values representing a lower prevalence of the behaviors during separation and a positive number representing an increased prevalence of the behavior compared to pre-separation. The measures focused on three behaviors commonly observed during the Social Separation Test: agitated locomotion, a measure of stress reactivity, calm locomotion, reflecting normal locomotor activity and alert, representing vigilant behavior (Table A3). For physiological responses, we calculated cortisol reactivity and cortisol regulation according to previous calculations used for marmosets during Social Separation Test (Cavanaugh et al., 2016). Cortisol reactivity was calculated as a percent increase to peak cortisol during the Social Separation Test from the pre-separation baseline sample. Cortisol regulation was calculated as a percent increase from the pre-separation baseline to the urine sample collected the morning following separation.

2.7 Statistical Analyses

All analyses were conducted in SAS (version 9.4), SPSS (IBM version 25), or HLM (version 7.0). Raw data were evaluated by year and data points three standard deviations above or below the mean were considered outliers and removed. Datasets were considered normal if the average skew across years fell between -1 and 1 . If datasets were skewed outside of this range, a square root transformation was used to achieve a normal distribution.

We utilized multilevel mixed effects models to incorporate both individual aging trajectories and predictors of between-individual variation by nesting repeated observations across aging within individual marmosets (Raudenbush & Byrk, 2002). Marmosets with partial or missing data were included in order to stabilize estimates of means and variances and full information maximum likelihood was used to estimate missing data. We first fit a linear or curvilinear (i.e. quadratic) baseline growth model that derives an average aging trajectory for all marmosets based on the individual trajectories of each marmoset. We retained the quadratic growth model when it provided a significantly better fit for the aging trajectory (chi-square test of deviance, $p < 0.05$). We then describe initial performance in the first year of testing (i.e. intercept) and growth trajectories (i.e. slopes) across the subsequent three years. When aging trajectories were linear, we report the slope across all four years. When aging trajectories were quadratic, we reported slopes of change between each year. Sex (Male, Female) was added as a predictor variable in order to assess whether the variability around the average slopes in the baseline growth model were significantly predicted by sex. We also added starting age as a covariate to the baseline growth model to control for variability in age at the start of the study. All results are expressed as mean \pm standard error of the mean.

3. Results

3.1 Cognitive Performance

3.1.1 Discrimination—Results from the baseline growth models are detailed in Table A.4. Cognitive performance across aging followed an inverted U-shaped trajectory and a significant curvature was found for trajectories of discrimination performance across aging ($B = 34.36 \pm 8.99$, $t(25) = 4.13$, $p < 0.001$; Figure A.2a). Curved aging trajectories are found in longitudinal cognitive studies in humans where scores improve initially due to additional practice with a repeated cognitive test, called a practice or retest effect (Salthouse, 2010, 2014). Marmosets took an average of 261.73 ± 12.37 trials to achieve criterion on discrimination in Year 1 and significantly improved in Year 2 (204.75 ± 17.30 trials to criterion; $B = -91.34 \pm 28.25$, $t(25) = -3.23$, $p = 0.003$). Discrimination performance was stable from Year 2 to Year 3 (216.47 ± 20.42 trials to criterion), but significantly declined in Year 4 ($B = 46.09 \pm 12.97$ trials to criterion, $t(25) = 3.55$, $p = 0.002$), with marmosets taking an average of 296.92 ± 28.73 trials to reach the 90% criterion.

The number of errors to criterion on discrimination followed a similar pattern (Figure A.2b). Marmosets made an average of 85.26 ± 8.50 errors in Year 1, then improved in Year 2 (68.33 ± 9.55 errors; $B = -1.39 (0.57)$, $t(26) = -2.42$, $p = 0.023$). The number of errors to

criterion remained stable in Year 3 (58.80 ± 6.71 errors) and then increased in Year 4 (92.60 ± 11.95 errors, $B = 0.57 \pm 0.27$, $t(26) = 2.13$, $p = 0.043$).

We also examined the types of errors made as a function of aging. For this analysis, the total number of errors to criterion was divided into Chance errors and Learning errors (Lai et al., 1995). Errors were categorized as Chance when a marmoset's performance was not significantly better than chance, determined to be a score of 27/40 based on 1-tailed binomial test ($p < 0.02$). Learning errors occurred after a marmoset was performing significantly better than chance (i.e. score 27/40) until reaching criterion (i.e. 36/40). Chance errors were relatively stable across aging, while changes in Learning errors accounted for most of the changes in total errors across the 4 years of testing (Figure A.2c). On average, marmosets made 44.09 ± 4.84 Learning errors in Year 1, significantly fewer Learning errors in Year 2 (35.91 ± 7.25 errors, $B = -1.52 \pm 0.51$, $t(26) = -2.98$, $p = 0.006$) and Year 3 (20.87 ± 3.69 errors, $B = -0.48 \pm 0.22$, $t(26) = -2.14$, $p = 0.042$). In Year 4, the number of Learning errors increased to 50.48 ± 9.79 , but this was not significantly greater than Year 3 ($B = 0.56 \pm 0.29$, $t(26) = 1.92$, $p = 0.065$).

3.1.2 Reversal—Trials to criterion on reversal also followed a quadratic aging trajectory ($B = 57.42 \pm 13.42$, $t(26) = 4.28$, $p < 0.001$; Figure A.2d). On average, marmosets took 505.93 ± 35.18 trials to reach criterion on reversal in Year 1, then significantly improved (377.27 ± 29.69 trials; $B = -186.07 \pm 38.68$, $t(26) = -4.81$, $p < 0.001$) from Year 1 to Year 2 and from Year 2 to Year 3 (363.47 ± 35.57 trials to criterion; $B = -71.23 \pm 19.62$, $t(26) = -3.63$, $p = 0.001$). Trials to criterion increased in Year 4 (464.42 ± 55.93 trials), but this was not significantly different from Year 3.

The errors to criterion followed the same pattern ($B = 18.18 \pm 4.57$, $t(26) = 3.98$, $p < 0.001$, Figure A.2e). Marmosets made an average of 218.70 ± 15.72 errors in Year 1, and fewer errors in Year 2 (171.45 ± 14.99 errors, $B = -65.43 \pm 15.50$, $t(26) = -4.22$, $p < 0.001$) and Year 3 (160.57 ± 16.77 errors, $B = -29.07 \pm 8.36$, $t(26) = -3.48$, $p = 0.002$). In Year 4 marmosets made significantly more errors (186.06 ± 20.17 errors, $B = 7.30 \pm 8.16$, $t(26) = 9.57$, $p < 0.001$) compared to Year 3.

To examine the errors types as a function of age, the total number of errors on reversal was divided into three categories: Perseverative, Chance and Learning (Lai et al., 1995). Perseverative errors occurred when a marmoset was performing significantly below chance level (13/40) based on 1-tailed binomial test ($p < 0.02$). The number of Perseverative errors did not change across aging (Figure A.2f). Chance errors followed a quadratic aging trajectory: marmosets made an average of 82.87 ± 7.89 Chance errors on reversal in Year 1, then made significantly fewer Chance errors in Year 2 (65.49 ± 12.29 errors, $B = -1.35 \pm 0.64$, $t(26) = -2.11$, $p = 0.044$) and Year 3 (51.19 ± 9.51 errors, $B = +0.64 (0.27)$, $t(26) = -2.40$, $p = 0.024$). Marmosets made more Chance errors in Year 4 (65.93 ± 9.21) but this was not a significant increase from Year 3. Learning errors did not change across aging.

3.1.3 Sex Differences—Sex differences were present in the slopes of cognitive performance between Years 1 and Year 3 for trials to criterion on discrimination (Figure 2A; Table A.5). Males needed fewer trials in Year 2 than Year 1 to complete discrimination

whereas females did not show improvement (sex difference in slopes: $t(23) = 2.22$, $p = 0.036$). From Year 2 to Year 3, males had stable cognitive performance whereas females showed decline (sex difference in slopes: $t(23) = 2.45$, $p = 0.022$). There was no sex difference in the trajectory of cognitive performance from Year 3 to Year 4, where both males and females showed cognitive decline by taking more trials to reach criterion. No sex difference was found for the errors to criterion, including Total errors, Chance errors and Learning errors (Figure 2B) suggesting that aging trajectories for errors to criterion were similar for males and females. Cognitive aging trajectories for trials to criterion on reversal did not differ as a function of sex overall. Yet, females had significantly poorer performance than males in Year 3 ($t(24) = -2.09$, $p = 0.047$) and Year 4 ($t(24) = -2.26$, $p = 0.033$; Figure 2C). Females made also more Total errors to criterion on reversal compared to males at the trend level in Year 3 ($t(24) = -1.96$, $p = 0.062$; Figure 2D) and Year 4 ($t(24) = -1.95$, $p = 0.063$). However, there was no sex difference within Perseverative, Chance or Learning errors for reversal.

3.2 Non-cognitive Variables: Aging and Sex Differences

Full results evaluating sex differences in motivation, motor performance, general behavior, and stress reactivity are reported in Table A.6.

3.2.1 Motivation in CANTAB cognitive testing.—To assess whether sex differences in cognitive performance were due to sex differences in motivation, we evaluated the proportion of omitted trials where marmosets declined to respond to any stimulus, as well as the latency to respond to stimuli during discrimination and reversal testing. All marmosets omitted fewer trials across the 4 years of testing ($B = -0.042 \pm 0.014$, $t(21) = -2.99$, $p = 0.007$; Figure 3A), regardless of sex. Response latencies on cognitive trials became faster across the 4 years of testing for both discrimination ($B = -208.46 \pm 45.45$, $t(21) = -4.59$, $p < 0.001$; Figure 3B) and reversal ($B = -161.24 \pm 37.77$, $t(21) = -4.27$, $p < 0.001$; Figure 3C), and did not differ for males and females. Therefore, there was no evidence for sex differences in motivation.

3.2.2 Motor performance in Hill and Valley test.—To complete cognitive tests, marmosets were required to reach with one hand to select a stimulus on the screen. We evaluated whether sex differences in reversal learning were due to sex differences in aging trajectories of motor abilities. We measured accuracy according to the total number of rewards retrieved (maximum 15 points) and latency to complete the motor tasks. We found that motor competence was stable across aging, with no evidence of decline in accuracy (Figure 3D) or latencies (Figure 3E). No sex difference was found for motor accuracy or latency.

3.2.3 Behavior in the homepage.—Calm locomotion, indicative of a relaxed state, increased across aging similarly for all marmosets ($B = 3.34 \pm 0.74$, $t(27) = 4.51$, $p < 0.001$), and did not differ between males and females (Figure 3F). Likewise, there were no significant sex differences in agitated locomotion, indicative of a stressed state, which declined across aging for all marmosets ($B = -0.57 \pm 0.064$, $t(27) = 8.97$, $p < 0.001$; Figure 3G). Males and females initiated social interactions at a similar frequency in Year

1, but females initiated less social behavior across aging. By Year 4 they initiated social behavior significantly less frequently than males ($B = 0.56 \pm 0.27$, $t(25) = 2.09$, $p = 0.047$; Figure 3H). Agonistic displays did not change across aging for males or females and no sex differences were present across years (Figure 3I). Thus, the slight decline in initiating social interactions in females is the only behavior in the homecage that parallels their decline in cognitive performance during the same time period.

3.2.4 Stress reactivity in the Social Separation Test

3.2.4.1 Behavioral Reactivity.: We used three behaviors to assess reactivity: calm locomotion, indicative of a relaxed state, agitated locomotion, an indicator of a stressed state, and inactive alert, indicative of a vigilant state. We calculated the degree to which these behaviors were more or less prevalent during the separation relative to baseline. As expected, calm locomotion was significantly lower during separation compared to baseline. Calm locomotion remained stable across aging (Year 1: $t(27) = -3.99$, $p = 0.001$, Year 2: $t(27) = -5.66$, $p < 0.001$, Year 3: $t(27) = -3.92$, $p < 0.001$; Figure 4A), and we found no sex differences. Also as expected, agitated locomotion was significantly higher during separation relative to baseline for Year 1 ($t(27) = 2.68$, $p = 0.012$) and Year 2 ($t(27) = 3.92$, $p < 0.001$). Agitated locomotion was similar to baseline during Year 3 ($t(27) = 0.993$, $p = 0.33$). No sex differences were detected (Figure 4B). Alert behavior was significantly higher during separation relative to baseline, remained stable across aging (Year 1: $t(27) = 4.67$, $p = 0.004$, Year 2: $t(27) = 4.26$, $p < 0.001$, Year 3: $t(27) = 2.11$, $p = 0.044$; Figure 4C) and did not differ between males and females.

3.2.4.2 Physiological Reactivity and Regulation.: The assays for urinary cortisol were different across years and therefore we could not assess within-individual trajectories of change in basal cortisol across aging. We analyzed cortisol reactivity, calculated as the percentage increase in cortisol at the peak level during separation compared to baseline for each individual. As expected, all marmosets had significantly elevated urinary cortisol during separation (Year 1: $t(26) = 7.73$, $p < 0.001$, Year 2: $t(26) = 10.67$, $p < 0.001$, Year 3: $t(27) = 8.18$, $p < 0.001$; Figure 4D), independent of sex. We also calculated cortisol regulation, or the degree to which marmosets returned to baseline cortisol levels the day following separation, by calculating the percentage increase in cortisol from urine the day following separation compared to baseline. Urinary cortisol was significantly elevated the day following separation compared to baseline each year across aging (Year 1: $t(27) = 3.65$, $p = 0.001$, Year 2: $t(27) = 5.36$, $p < 0.001$, Year 3: $t(27) = 3.63$, $p = 0.001$, Figure 4E) independent of sex.

4. Discussion

We longitudinally assessed reversal learning performance in marmosets across four years as they transitioned from middle (~ 5 years old) to old age (~ 9 years old). Reversal learning is a benchmark task of executive function, a cognitive domain that substantially declines with age in all primates studied (Lacreuse et al., 2020) and across a range of other taxa and paradigms (Izquierdo et al., 2017). Marmosets performed better on the discrimination compared to the reversal phase, indicated by fewer trials to criterion and fewer errors, which

is typical for reversal learning in a wide range of primate (Lai et al., 1995; Rapp, 1990; Voytko, 1999) and non-primate species (dogs: Boutet et al., 2005; rats: Brushfield et al., 2008). We found changes across aging in both discrimination and reversal performance. Specifically, we observed cognitive decline after a period of initial improvement in cognitive performance, a practice effect characteristic of longitudinal studies of cognitive aging (Salthouse, 2010). Marmosets improved on both discrimination and reversal in the second year compared to the first. Performance then remained stable for discrimination, whereas marmosets continued to improve on reversals. Performance worsened in the final year, as shown by a decline on discrimination and a lack of practice effect on reversal. These results suggest that we captured the beginning of cognitive decline in the final year of the study, corresponding to age 8 to 9, which is consistent with the timing of decline identified by a cross-sectional study in young and older marmosets (Sadoun et al., 2019). We found a greater effect of age on discrimination whereas Sadoun (2019) found greater age-related decline on reversal. These discrepancies could be due to experimental differences including the overall approach, as cross-sectional studies can overestimate aging effects, while longitudinal designs are more sensitive to change within individuals across aging (Salthouse, 2014).

Importantly, sex differences pointed to an earlier cognitive impairment in female marmosets. Unlike males, female marmosets did not show an initial improvement in discrimination performance. This is potentially important as lack of practice effect in older humans has been associated to future cognitive decline and dementia diagnosis (Jutten et al., 2020). In addition, females declined on discrimination starting in middle-age, during the second year of the study, whereas males did not decline until the third year. For reversals, the aging trajectory for males and females are more similar as both display initial improvement and maintenance followed by decline. However, females' age-related decline in reversal performance was more prominent as they obtained significantly poorer scores than males in Years 3 and 4, without specific increases in perseverative, chance or learning errors. In a previous publication from the same study (Workman et al., 2018), we also report evidence of female impairment when we assessed changes in reversal learning performance from the first two years of the study on a subset of animals (n=18) and using a different analytic approach (i.e. repeated measures ANOVA considering reversal learning phases of discrimination and reversal as a within-subject variable). This initial report also found practice effects with improved performance in Year 2, but the sex difference was more prominent for the reversals, as opposed to discrimination performance. In agreement with Workman et al. (2018), the present paper also finds that females obtain poorer scores than males on reversals in Year 1 and Year 2, but the sex difference is not significant with a larger number of animals and an analysis considering each phase of reversal learning independently.

We assessed several non-cognitive factors that could contribute to age-related decline and sex differences. Interestingly, motivation, as assessed by latencies to respond and omitted trials in cognitive testing, improved across aging. Motor ability, as assessed by the Hill and Valley test, likewise did not decline, which was unexpected as slowing of motor function has been reliably observed in aging NHPs (Lacreuse et al., 2005, 2014; Zhang et al., 2000). These discrepancies could be due to task differences or the age range studied.

Behaviors in the home cage suggested that all marmosets, regardless of sex, displayed less agitated locomotion together with more calm locomotion as they age. We also found that all marmosets maintained species-typical responses to social separation with elevated behavioral and physiological reactivity compared to pre-separation baseline (French et al., 2007). Based on these results, we do not believe that changes in behavior profiles or stress reactivity significantly contributed to the sex differences in cognitive decline. However, since the Social Separation Test was not conducted in the final year, more data are needed to completely rule out the role of these factors in age-related cognitive decline. Males and females only differed in the initiation of social interactions with the opposite-sex partner, with females initiating social interaction less frequently than males across aging. This very interesting finding is worth investigating further, as changes in social interactions have been shown to co-occur with cognitive decline and to be indicative of β -amyloid deposition in cognitively normal older people (Biddle et al., 2019).

A caveat to this study is that we did not measure steroid hormones. However, endocrine differences between males and females are likely to contribute to sex differences observed in cognitive aging. Female marmosets, unlike women, do not experience menopause and continue to cycle to the end of life (Abbott et al., 2003), whereas male marmosets, like men, experience a decline in testosterone levels with age (Tardif et al., 2008). In a previous study, we found that estradiol treatment impaired cognitive performance in ovariectomized female marmosets such that they made more perseverative errors during serial reversals (Lacreuse et al., 2014). Recent work in rats also points to estrogens impairing sensitivity to reward contingencies and overall cognitive flexibility (Schoenberg et al., 2019). In contrast, testosterone administration had no effect on serial reversal acquisition in gonadectomized male marmosets (LaClair & Lacreuse, 2016). Additional studies will be needed to clarify the role of sex steroids in modulating patterns of cognitive aging in male and female marmosets.

The observed sex differences in aging trajectories may not be specific of cognitive flexibility. Indeed, there was no evidence for an increase in perseverative errors overall in any year of the study in either sex. In addition, the sex difference in performance was not limited to the reversals, but extended to discrimination. Discrimination and reversals involve somewhat distinct processes, with reversals being unique in requiring flexibility, the ability to inhibit a prepotent response and select a stimulus that was previously non-rewarded. Although associative learning is involved in both phases, it has been suggested that discrimination and reversals may involve a different task space (Costa et al., 2015; Jang et al., 2015) or representation of task demands (Wilson et al., 2014), where the probability for a change of contingencies occurs only in the reversals (see Izquierdo et al., 2017 for a detailed discussion; Harris et al., in press). One interpretation of our finding is that females experienced earlier and greater age-related decline than males in associative learning, that impacted both the initial learning stages (discrimination) and learning of the reversed reward-stimulus associations. Alternatively, we cannot rule out that the sex difference in the reversals is more related to the cognitive representations of the task. Studies investigating the neural bases of reversal learning in each sex are needed to clarify these results.

It is important to note that we did not track age-related changes in sensory abilities such as vision, audition or olfaction, which can decline across aging for marmosets (Ross, 2019).

We cannot completely rule out a contribution of age-related visual impairments to cognitive decline as difficulty in discriminating between the two visual stimuli would have led to longer acquisition times in the discrimination phase. However, at the end of this phase, all monkeys reached the 90% correct criterion, demonstrating their ability to discriminate between the two visual stimuli. Therefore, visual impairment should only have a limited impact on the ability to switch to the alternate stimulus in the reversal phase. Because cognitive decline was observed in both phases, the development of visual impairment with age cannot solely explain this pattern of results. Yet, we recognize the crucial importance of carefully assessing visual and other sensory functions in future investigations, as these changes can precede major cognitive changes in humans (Lindenberger & Baltes, 1994) and have been suggested as early indicators of AD.

Interestingly, we reported remarkably large and robust sex differences in resting state functional connectivity (rsFC) in this cohort of marmosets, measured on a subset of animals during Year 1 (Laclair et al., 2019) and Year 2 (Nephew et al., 2020). At both timepoints males were found to have stronger rsFC compared to females, especially within the PFC and striatum, and this pattern of connectivity correlated with cognitive performance. Additionally, males had greater glutamine + glutamate (Glx) levels in PFC in Year 1, as assessed by magnetic resonance spectroscopy (Lacreuse et al., 2018). These findings suggest that sex differences in brain networks and neurochemistry may shape differential cognitive processing between the sexes. Future studies should use neuroimaging to track changes in brain structure and brain activity over time, along with measuring gonadal steroids and cognitive performance.

Neurocognitive aging in humans is highly variable and can range from normal decline to pathological decline characteristic of dementia. Biological sex seems to influence both healthy and pathological aging in humans; sex differences have been found in healthy, cognitively normal adults across aging, but there are only a few studies and the findings are inconsistent (Levine et al., 2021; McCarrey et al., 2016). With regards to pathological aging, women are considered more susceptible to AD as they make up two thirds of patients diagnosed with dementia (Alzheimer's Association, 2019). However, the epidemiological evidence is highly debated (Ferretti et al., 2018), due in part to longer lifespans for women (Lemaître et al., 2020). Nevertheless, the clinical evidence points to women experiencing greater AD symptomatology, faster rate of cognitive decline and brain atrophy than men (Ferretti et al., 2018). Our study suggests that female marmosets may likewise be more susceptible to earlier and steeper cognitive decline. Interestingly, marmosets have the opposite of pattern of longevity than humans, with females having a shorter lifespan (10 years) than males (12 years) (Nishijima et al., 2012). Although we did not find evidence of sex differences in frailty in our animals it will be important to determine whether cognitive decline in female marmoset is related to reproductive burden and overall shorter lifespan (Nishijima et al., 2012).

Beyond biological sex, individual trajectories of neurocognitive aging have been found to become more variable across older age in healthy aging adults (Nyberg et al., 2020). Likewise, we found highly variable cognitive trajectories, both within and across sexes, with presence of individuals that were cognitively stable and other that declined sharply. At

this time, it is unclear whether the aging trajectories that we observed are representative of healthy aging, pathological aging or a mix of both processes. In order to leverage the true value of the marmoset as a model for neurocognitive aging, ongoing work will link cognitive aging trajectories to neuropathological burden within the same individuals. Including sex in these preclinical studies (Shansky & Murphy, 2021; Waters & Laitner, 2021) will be a crucial step to understand how sex shapes neurocognitive trajectories in our own species (Ferretti et al., 2018).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by R01 AG 046266 to AL and F32 AG064925 to ER. We thank Drs Nicole Gervais and Matthew LaClair and many undergraduate students for assistance in collecting data. This long-term study would not have been possible without the excellent support of the UMass Animal Care and Veterinary Staff headed by Dr. Paul Spurlock. We also thank the Psychology shop staff, headed by Joe Bergman, and the UMass Center for Research on Families, especially Dr. Holly Laws, for assistance with data analysis.

References

- Abbott DH, Barnett DK, Colman RJ, Yamamoto ME, & Schultz-Darken NJ (2003). Aspects of Common Marmoset Basic Biology and Life History Important for Biomedical Research. *Comparative Medicine*, 53(4), 339–350. [PubMed: 14524409]
- Alzheimer's Association. (2019). 2019 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 15(3), 321–387. 10.1016/j.jalz.2019.01.010
- Arnsten AFT, Datta D & Preuss TM (2021). Studies of aging nonhuman primates illuminate the etiology of early-stage Alzheimer's-like neuropathology: An evolutionary perspective. *American Journal of Primatology*, e23254. 10.1002/ajp.23254 [PubMed: 33960505]
- Avila JF, Vonk MJ, Verney SP, Witkiewitz K, Arce Rentería M, Schupf N, Mayeux R, & Manly JJ (2019). Sex/gender differences in cognitive trajectories vary as a function of race/ethnicity. *Alzheimer's and Dementia*, 15(12), 1516–1523. 10.1016/j.jalz.2019.04.006
- Baxter MG (2001). Cognitive Aging in Nonhuman Primates. In *Functional Neurobiology of Aging* (pp. 407–419). Elsevier. 10.1016/b978-012351830-9/50028-7
- Biddle KD, d'Oleire Uquillas F, Jacobs HIL, Zide B, Kirn DR, Rentz DM, Johnson KA, Sperling RA, & Donovan NJ (2019). Social Engagement and Amyloid- β -Related Cognitive Decline in Cognitively Normal Older Adults. *American Journal of Geriatric Psychiatry*, 27(11), 1247–1256. 10.1016/j.jagp.2019.05.005
- Bihel E, Pro-Sistiaga P, Letourneur A, Toutain J, Saulnier R, Insausti R, Bernaudin M, Roussel S, & Touzani O (2010). Permanent or transient chronic ischemic stroke in the non-human primate: Behavioral, neuroimaging, histological, and immunohistochemical investigations. *Journal of Cerebral Blood Flow and Metabolism*, 30(2), 273–285. 10.1038/jcbfm.2009.209 [PubMed: 19794396]
- Boutet I, Ryan M, Kulaga V, McShane C, Christie LA, Freedman M, & Milgram NW (2005). Age-associated cognitive deficits in humans and dogs: A comparative neuropsychological approach. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 29(3), 433–441. 10.1016/J.PNPBP.2004.12.015 [PubMed: 15795052]
- Brushfield A, Luu T, Callahan B, & Gilbert P (2008). A comparison of discrimination and reversal learning for olfactory and visual stimuli in aged rats. *Behavioral Neuroscience*, 122(1), 54–62. 10.1037/0735-7044.122.1.54 [PubMed: 18298249]
- Burkart JM, & van Schaik CP (2020). Marmoset prosociality is intentional. *Animal Cognition*, 23(3), 581–594. 10.1007/s10071-020-01363-6 [PubMed: 32107657]

- Cacioppo JT, Cacioppo S, Capitanio JP, & Cole SW (2015). The Neuroendocrinology of Social Isolation. *Annual Review of Psychology*, 66(1), 733–767. 10.1146/annurev-psych-010814-015240
- Caine NG, Osorio D, & Mundy NI (2010). A foraging advantage for dichromatic marmosets (*Callithrix geoffroyi*) at low light intensity. *Biology Letters*, 6(1), 36–38. 10.1098/rsbl.2009.0591 [PubMed: 19740895]
- Cavanaugh J, Carp SB, Rock CM, & French JA (2016). Oxytocin modulates behavioral and physiological responses to a stressor in marmoset monkeys. *Psychoneuroendocrinology*, 66, 22–30. 10.1016/j.psyneuen.2015.12.027 [PubMed: 26771946]
- Clarke HF, Walker SC, Crofts HS, Dalley JW, Robbins TW, & Roberts AC (2005). Prefrontal Serotonin Depletion Affects Reversal Learning But Not Attentional Set Shifting. 10.1523/JNEUROSCI.3690-04.2005
- Clarke HF, Hill GJ, Robbins TW, & Roberts AC (2011). Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. *Journal of Neuroscience*, 31(11), 4290–4297. 10.1523/JNEUROSCI.5066-10.2011 [PubMed: 21411670]
- Collins P, Wilkinson LS, Everitt BJ, Robbins TW, & Roberts AC (2000). The effect of dopamine depletion from the caudate nucleus of the common marmoset (*Callithrix jacchus*) on tests of prefrontal cognitive function. *Behavioral Neuroscience*, 114(1), 3–17. 10.1037/0735-7044.114.1.3 [PubMed: 10718258]
- Costa VD, Tran VL, Turchi J, & Averbeck BB (2015). Reversal Learning and Dopamine: A Bayesian Perspective. *Journal of Neuroscience*, 35(6), 2407–2416. 10.1523/JNEUROSCI.1989-14.2015 [PubMed: 25673835]
- Eslamboli A, Baker HF, Ridley RM, & Annett LE (2003). Sensorimotor deficits in a unilateral intrastriatal 6-OHDA partial lesion model of Parkinson's disease in marmoset monkeys. *Experimental Neurology*, 183(2), 418–429. 10.1016/S0014-4886(03)00139-0 [PubMed: 14552882]
- Ferretti MT, Iulita MF, Cavedo E, Chiesa PA, Schumacher Dimech A, Santucci Chadha A, Baracchi F, Girouard H, Misoch S, Giacobini E, Depypere H, & Hampel H (2018). Sex differences in Alzheimer disease — the gateway to precision medicine. *Nature Reviews Neurology*, 14(8), 457–469. 10.1038/s41582-018-0032-9 [PubMed: 29985474]
- Fischer KE, & Austad SN (2011). The Development of Small Primate Models for Aging Research. *ILAR Journal*, 52(1), 78–88. 10.1093/ilar.52.1.78 [PubMed: 21411860]
- French JA, Fite JE, Jensen H, Oparowski K, Rukstalis MR, Fix H, Jones B, Maxwell H, Pacer M, Power Mi. L., & Schulkin J (2007). Treatment With CRH-1 Antagonist Antalarmin Reduces Behavioral and Endocrine Responses to Social Stressors in Marmosets (*Callithrix kuhlii*). *American Journal of Primatology*, 69, 877–889. 10.1002/ajp.20385 [PubMed: 17397037]
- Geula C, Nagykeri N, & Wu C-K (2002). Amyloid- β deposits in the cerebral cortex of the aged common marmoset (*Callithrix jacchus*): incidence and chemical composition. *Acta Neuropathologica*, 103(1), 48–58. 10.1007/s004010100429 [PubMed: 11837747]
- Harlow HF (1949). The formation of learning sets. *Psychological Review*, 56(1), 51–65. 10.1037/h0062474 [PubMed: 18124807]
- Harris C, Aguirre C, Kolli S, Das K, Izquierdo A, & Soltani A (2021). Unique features of stimulus-based probabilistic reversal learning. *BioRxiv*, 2020.09.24.310771. 10.1101/2020.09.24.310771
- Hennessy MB (1997). Hypothalamic-pituitary-adrenal responses to brief social separation. *Neuroscience and Biobehavioral Reviews*, 21(1), 11–29. 10.1016/S0149-7634(96)00013-9 [PubMed: 8994206]
- Hook MA, & Rogers LJ (2000). Development of hand preferences in marmosets (*Callithrix jacchus*) and effects of aging. *Journal of Comparative Psychology*, 114(3), 263–271. 10.1037/0735-7036.114.3.263 [PubMed: 10994842]
- Huber L, & Voelkl B (2009). Social and Physical Cognition in Marmosets and Tamarins. In *The Smallest Anthropoids* (pp. 183–201). Springer US. 10.1007/978-1-4419-0293-1_10
- Izquierdo A, Brigman JL, Radke AK, Rudebeck PH, & Holmes A (2017). The neural basis of reversal learning: An updated perspective. In *Neuroscience* (Vol. 345, pp. 12–26). Elsevier Ltd. 10.1016/j.neuroscience.2016.03.021 [PubMed: 26979052]

- Jackson SAW, Horst NK, Axelsson SFA, Horiguchi N, Cockcroft GJ, Robbins TW, & Roberts AC (2018). Selective Role of the Putamen in Serial Reversal Learning in the Marmoset. *Cerebral Cortex*, 1–14. 10.1093/cercor/bhy276 [PubMed: 29253248]
- Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, & Averbeck BB (2015). The Role of Frontal Cortical and Medial-Temporal Lobe Brain Areas in Learning a Bayesian Prior Belief on Reversals. *Journal of Neuroscience*, 35(33), 11751–11760. 10.1523/JNEUROSCI.1594-15.2015 [PubMed: 26290251]
- Jutten RJ, Grandoit E, Foldi NS, Sikkes SAM, Jones RN, Choi S, Lamar ML, Loudon DKN, Rich J, Tommet D, Crane PK, & Rabin LA (2020). Lower practice effects as a marker of cognitive performance and dementia risk: A literature review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 12(1), e12055. 10.1002/dad2.12055
- Kaas JH (2021). Comparative Functional Anatomy of Marmoset Brains. *ILAR Journal*, 2021, 1–14. 10.1093/ilar/ilaa026
- Kangas BD, Bergman J, & Coyle JT (2016). Touchscreen assays of learning, response inhibition, and motivation in the marmoset (*Callithrix jacchus*). *Animal Cognition*, 19(3), 673–677. 10.1007/s10071-016-0959-4 [PubMed: 26846231]
- Laclair M, Febo M, Nephew B, Gervais NJ, Poirier G, Workman K, Chumachenko S, Payne L, Moore MC, King JA, & Lacreuse A (2019). Sex differences in cognitive flexibility and resting brain networks in middle-aged marmosets. *ENeuro*, 6(4). 10.1523/ENEURO.0154-19.2019
- LaClair M, & Lacreuse A (2016). Reversal learning in gonadectomized marmosets with and without hormone replacement: are males more sensitive to punishment? *Animal Cognition*, 19(3), 619–630. 10.1007/s10071-016-0966-5 [PubMed: 26909674]
- Lacreuse A, Chang J, Metevier CM, LaClair M, Meyer JS, & Ferris CM (2014). Oestradiol Modulation of Cognition in Adult Female Marmosets (*Callithrix jacchus*). *Journal of Neuroendocrinology*, 26(5), 296–309. 10.1111/jne.12147 [PubMed: 24617856]
- Lacreuse A, & Herndon JG (2009). Nonhuman Primate Models of Cognitive Aging. In *Animal Models of Human Cognitive Aging* (pp. 1–30). Humana Press. 10.1007/978-1-59745-422-3_2
- Lacreuse A, Diehl MM, Goh MY, Hall MJ, Volk AM, Chhabra RK, & Herndon JG (2005). Sex differences in age-related motor slowing in the rhesus monkey: behavioral and neuroimaging data. *Neurobiology of Aging*, 26(4), 543–551. 10.1016/J.NEUROBIOLAGING.2004.05.007 [PubMed: 15653182]
- Lacreuse A, Moore CM, LaClair M, Payne L, & King JA (2018). Glutamine/glutamate (Glx) concentration in prefrontal cortex predicts reversal learning performance in the marmoset. *Behavioural Brain Research*, 346, 11–15. 10.1016/j.bbr.2018.01.025 [PubMed: 29378291]
- Lacreuse A, Raz N, Schmidtke D, Hopkins WD, & Herndon JG (2020). Age-related decline in executive function as a hallmark of cognitive ageing in primates: an overview of cognitive and neurobiological studies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 375(1811), 20190618. 10.1098/rstb.2019.0618 [PubMed: 32951543]
- Lacreuse A, Russell JL, Hopkins WD, & Herndon JG (2014). Cognitive and motor aging in female chimpanzees. *Neurobiology of Aging*, 35(3), 623–632. 10.1016/J.NEUROBIOLAGING.2013.08.036 [PubMed: 24112794]
- Lai ZC, Moss MB, Killiany RJ, Rosene DL, & Herndon JG (1995). Executive system dysfunction in the aged monkey: Spatial and object reversal learning. *Neurobiology of Aging*. 10.1016/0197-4580(95)02014-4
- Lemaître JF, Ronget V, Tidière M, Allainé D, Berger V, Cohas A, Colchero F, Conde DA, Garratt M, Liker A, Marais GAB, Scheuerlein A, Székely T, & Gaillard JM (2020). Sex differences in adult lifespan and aging rates of mortality across wild mammals. *Proceedings of the National Academy of Sciences of the United States of America*, 117(15), 8546–8553. 10.1073/pnas.1911999117 [PubMed: 32205429]
- Levine DA, Gross AL, Briceño EM, Tilton N, Giordani BJ, Sussman JB, Hayward RA, Burke JF, Hingtgen S, Elkind MSV, Manly JJ, Gottesman RF, Gaskin DJ, Sidney S, Sacco RL, Tom SE, Wright CB, Yaffe K, & Galecki AT (2021). Sex Differences in Cognitive Decline Among US Adults. *JAMA Network Open*, 4(2), e210169. 10.1001/jamanetworkopen.2021.0169 [PubMed: 33630089]

- Lindenberger U, & Baltes PB (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging*, 9(3), 339–355. 10.1037/0882-7974.9.3.339 [PubMed: 7999320]
- Maclean CJ, Baker HF, Ridley RM, & Mori H (2000). Naturally occurring and experimentally induced β -amyloid deposits in the brains of marmosets (*Callithrix jacchus*). *Journal of Neural Transmission*, 107(7), 799–814. 10.1007/s007020070060 [PubMed: 11005545]
- Marshall JWB, & Ridley RM (2003). Assessment of cognitive and motor deficits in a marmoset model of stroke. In *ILAR Journal* (Vol. 44, Issue 2, pp. 153–160). Institute for Laboratory Animal Research. 10.1093/ilar.44.2.153 [PubMed: 12652010]
- McCarrey AC, An Y, Kitner-Triolo MH, Ferrucci L, & Resnick SM (2016). Sex differences in cognitive trajectories in clinically normal older adults. *Psychology and Aging*, 31(2), 166–175. 10.1037/pag0000070 [PubMed: 26796792]
- McQuail JA, Dunn AR, Stern Y, Barnes CA, Kempermann G, Rapp PR, Kaczorowski CC, & Foster TC (2021). Cognitive Reserve in Model Systems for Mechanistic Discovery: The Importance of Longitudinal Studies. *Frontiers in Aging Neuroscience*, 12, 532. 10.3389/fnagi.2020.607685
- Miller CT (2017). Why Marmosets? *Developmental Neurobiology*, 77(3), 237–243. 10.1002/dneu.22483 [PubMed: 28170158]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41(1), 49–100. 10.1006/cogp.1999.0734 [PubMed: 10945922]
- Moss MB, Killiany RJ, & Herndon JG (1999). Age-Related Cognitive Decline in the Rhesus Monkey (pp. 21–47). Springer, Boston, MA. 10.1007/978-1-4615-4885-0_2
- Munger EL, Takemoto A, Raghanti MA, & Nakamura K (2017). Visual discrimination and reversal learning in aged common marmosets (*Callithrix jacchus*). *Neuroscience Research*, 124, 57–62. 10.1016/j.neures.2017.06.002 [PubMed: 28602996]
- Nakamura K, Koba R, Miwa M, Yamaguchi C, Suzuki H, & Takemoto A (2018). A Method to Train Marmosets in Visual Working Memory Task and Their Performance. *Frontiers in Behavioral Neuroscience*, 12, 46. 10.3389/fnbeh.2018.00046 [PubMed: 29615876]
- Nephew BC, Febo M, Cali R, Workman KP, Payne L, Moore CM, King JA, & Lacreuse A (2020). Robustness of sex-differences in functional connectivity over time in middle-aged marmosets. *Scientific Reports*, 10(1), 16647. 10.1038/s41598-020-73811-9 [PubMed: 33024242]
- Nishijima K, Saitoh R, Tanaka S, Ohsato-Suzuki M, Ohno T, & Kitajima S (2012). Life span of common marmoset (*Callithrix jacchus*) at CLEA Japan breeding colony. *Biogerontology*, 13, 439–443. 10.1007/s10522-012-9388-1 [PubMed: 22752736]
- Nyberg L, Boraxbekk CJ, Sörman DE, Hansson P, Herlitz A, Kauppi K, Ljungberg JK, Lövheim H, Lundquist A, Adolfsson AN, Oudin A, Pudas S, Rönnlund M, Stiernstedt M, Sundström A, & Adolfsson R (2020). Biological and environmental predictors of heterogeneity in neurocognitive ageing: Evidence from *Betula* and other longitudinal studies. In *Ageing Research Reviews* (Vol. 64, p. 101184). Elsevier Ireland Ltd. 10.1016/j.arr.2020.101184 [PubMed: 32992046]
- Pessoa DMA, Tomaz C, & Pessoa VF (2005). Color vision in marmosets and tamarins: behavioral evidence. *American Journal of Primatology*, 67(4), 487–495. 10.1002/ajp.20202 [PubMed: 16342075]
- Phillips KA, Ross CN, Spross J, Cheng CJ, Izquierdo A, Biju KC, Chen C, Li S, & Tardif SD (2017). Behavioral phenotypes associated with MPTP induction of partial lesions in common marmosets (*Callithrix jacchus*). *Behavioural Brain Research*, 325, 51–62. 10.1016/j.bbr.2017.02.010 [PubMed: 28219749]
- Phillips KA, Watson CM, Bearman A, Knippenberg AR, Adams J, Ross C, & Tardif SD (2019). Age-related changes in myelin of axons of the corpus callosum and cognitive decline in common marmosets. *American Journal of Primatology*, 81(2), e22949. 10.1002/ajp.22949 [PubMed: 30620098]
- Pryce CR, Dettling AC, Spengler M, Schnell CR, & Feldon J (2004). Deprivation of parenting disrupts development of homeostatic and reward systems in marmoset monkey offspring. *Biological Psychiatry*, 56(2), 72–79. 10.1016/j.biopsych.2004.05.002 [PubMed: 15231438]

- Rahman A, Jackson H, Hristov H, Isaacson RS, Saif N, Shetty T, Etingin O, Henchcliffe C, Brinton RD, & Mosconi L (2019). Sex and Gender Driven Modifiers of Alzheimer's: The Role for Estrogenic Control Across Age, Race, Medical, and Lifestyle Risks. In *Frontiers in Aging Neuroscience* (Vol. 11, p. 315). Frontiers Media S.A. 10.3389/fnagi.2019.00315 [PubMed: 31803046]
- Rapp PR (1990). Visual discrimination and reversal learning in the aged monkey (*Macaca mulatta*). *Behavioral Neuroscience*, 104(6), 876–884. 10.1037//0735-7044.104.6.876 [PubMed: 2285486]
- Raudenbush SW, & Byrk AS (2002). *Hierarchical Linear Models: Applications and data analysis methods*. Sage. <https://us.sagepub.com/en-us/nam/hierarchical-linear-models/book9230>
- Raz N, & Lindenberger U (2011). Only Time Will Tell: Cross-Sectional Studies Offer No Solution to the Age-Brain-Cognition Triangle: Comment on Salthouse (2011). 10.1037/a0024503
- Ridley RM, Haystead TAJ, & Baker HF (1981). An analysis of visual object reversal learning in the marmoset after amphetamine and haloperidol. *Pharmacology, Biochemistry and Behavior*, 14(3), 345–351. 10.1016/0091-3057(81)90401-9
- Roberts AC, Robbins TW, & Everitt BJ (1988). The Effects of Intradimensional and Extradimensional Shifts on Visual Discrimination Learning in Humans and Non-human Primates. *The Quarterly Journal of Experimental Psychology Section B*, 40(4), 321–341. 10.1080/14640748808402328
- Roberts AC, Robbins TW, Everitt BJ, Jones GH, sirkia TE, Wilkinson J, & Page K (1990). The effects of excitotoxic lesions of the basal forebrain on the acquisition, retention and serial reversal of visual discriminations in marmosets. *Neuroscience*, 34(2), 311–329. 10.1016/0306-4522(90)90142-Q [PubMed: 2110326]
- Rodriguez-Callejas JD, Fuchs E, & Perez-Cruz C (2016). Evidence of Tau Hyperphosphorylation and Dystrophic Microglia in the Common Marmoset. *Frontiers in Aging Neuroscience*, 8, 315. 10.3389/fnagi.2016.00315 [PubMed: 28066237]
- Ross CN (2019). Marmosets in Aging Research. *The Common Marmoset in Captivity and Biomedical Research*, 355–376. 10.1016/B978-0-12-811829-0.00021-2
- Ross CN, Davis K, Dobek G, & Tardif SD (2012). Aging phenotypes of common marmosets (*Callithrix jacchus*). *Journal of Aging Research*, 2012. 10.1155/2012/567143
- Rothwell ES, Freire-Cobo C, Varghese M, Edwards M, Janssen WGM, Hof PR, & Lacreuse A (2021). The marmoset as an important primate model for longitudinal studies of neurocognitive aging. *American Journal of Primatology*, e23271. 10.1002/ajp.23271 [PubMed: 34018622]
- Sadoun A, Rosito M, Fonta C, & Girard P (2019). Key periods of cognitive decline in a nonhuman primate model of cognitive aging, the common marmoset (*Callithrix jacchus*). *Neurobiology of Aging*, 74, 1–14. 10.1016/J.NEUROBIOLAGING.2018.10.003 [PubMed: 30391909]
- Salthouse TA (2010). Influence of Age on Practice Effects in Longitudinal Neurocognitive Change. *Neuropsychology*, 24(5), 563–572. 10.1037/a0019026 [PubMed: 20804244]
- Salthouse TA (2014). Why Are There Different Age Relations in Cross-Sectional and Longitudinal Comparisons of Cognitive Functioning? *Current Directions in Psychological Science*, 23(4), 252–256. 10.1177/0963721414535212 [PubMed: 25382943]
- Saltzman W, Schultz-Darken NJ, Wegner FH, Wittwer DJ, & Abbott DH (1998). Suppression of cortisol levels in subordinate female marmosets: Reproductive and social contributions. *Hormones and Behavior*, 33(1), 58–74. 10.1006/hbeh.1998.1436 [PubMed: 9571014]
- Schoenberg HL, Sola EX, Seyller E, Kelberman M, & Toufexis DJ (2019). Female rats express habitual behavior earlier in operant training than males. *Behavioral Neuroscience*, 133(1), 110–120. 10.1037/bne0000282 [PubMed: 30359063]
- Schubiger MN, Kissling A, & Burkart JM (2019). Does opportunistic testing bias cognitive performance in primates? Learning from drop-outs. *PLOS ONE*, 14(3), e0213727. 10.1371/journal.pone.0213727 [PubMed: 30893340]
- Shansky RM, & Murphy AZ (2021). Considering sex as a biological variable will require a global shift in science culture. *Nature Neuroscience*, 1–8. 10.1038/s41593-021-00806-8
- Spinelli S, Pennanen L, Dettling AC, Feldon J, Higgins GA, & Pryce CR (2004). Performance of the marmoset monkey on computerized tasks of attention and working memory. *Cognitive Brain Research*, 19(2), 123–137. 10.1016/j.cogbrainres.2003.11.007 [PubMed: 15019709]

- Strasser A, & Burkart JM (2012). Can We Measure Brain Efficiency? An Empirical Test with Common Marmosets (*Callithrix jacchus*). *Brain, Behavior and Evolution*, 80(1), 26–40. 10.1159/000338014
- Takemoto A, Miwa M, Koba R, Yamaguchi C, Suzuki H, & Nakamura K (2015). Individual variability in visual discrimination and reversal learning performance in common marmosets. *Neuroscience Research*, 93, 136–143. 10.1016/j.neures.2014.10.001 [PubMed: 25315397]
- Tardif SD, Araujo A, Arruda MF, French JA, Sousa MBC, & Yamamoto ME (2008). Reproduction and Aging in Marmosets and Tamarins. In *Primate Reproductive Aging* (Vol. 36, pp. 29–48). KARGER. 10.1159/000137678
- Voytko M Lou. (1999). Impairments in acquisition and reversals of two-choice discriminations by aged rhesus monkeys. *Neurobiology of Aging*, 20(6), 617–627. 10.1016/S0197-4580(99)00097-4 [PubMed: 10674427]
- Voytko ML, & Tinkler GP (2004). Cognitive function and its neural mechanisms in nonhuman primate models of aging, Alzheimer’s disease, and menopause. In *Frontiers in Bioscience* (Vol. 9, pp. 1899–1914). *Frontiers in Bioscience*. 10.2741/1370 [PubMed: 14977596]
- Waters A, & Laitner MH (2021). Biological sex differences in Alzheimer’s preclinical research: A call to action. *Alzheimer’s & Dementia: Translational Research & Clinical Interventions*, 7(1), e12111. 10.1002/trc2.12111
- Wilson RC, Takahashi YK, Schoenbaum G, & Niv Y (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, 81(2), 267–279. 10.1016/J.NEURON.2013.11.005 [PubMed: 24462094]
- Workman KP, Healey B, Carlotto A, & Lacreuse A (2018). One-year change in cognitive flexibility and fine motor function in middle-aged male and female marmosets (*Callithrix jacchus*). *American Journal of Primatology*, e22924. 10.1002/ajp.22924 [PubMed: 30281810]
- Zhang Z, Andersen A, Smith C, Grondin R, Gerhardt G, & Gash D (2000). Motor slowing and Parkinsonian signs in aging rhesus monkeys mirror human aging. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 55(10). 10.1093/gerona/55.10.B473

Highlights:

- First longitudinal study of cognition in aging marmosets (*Callithrix jacchus*)
- Practice effects preceded age-related cognitive decline
- Females exhibited cognitive decline at earlier age than males
- No sex differences in aging for motor abilities, behavior or social stress reactivity

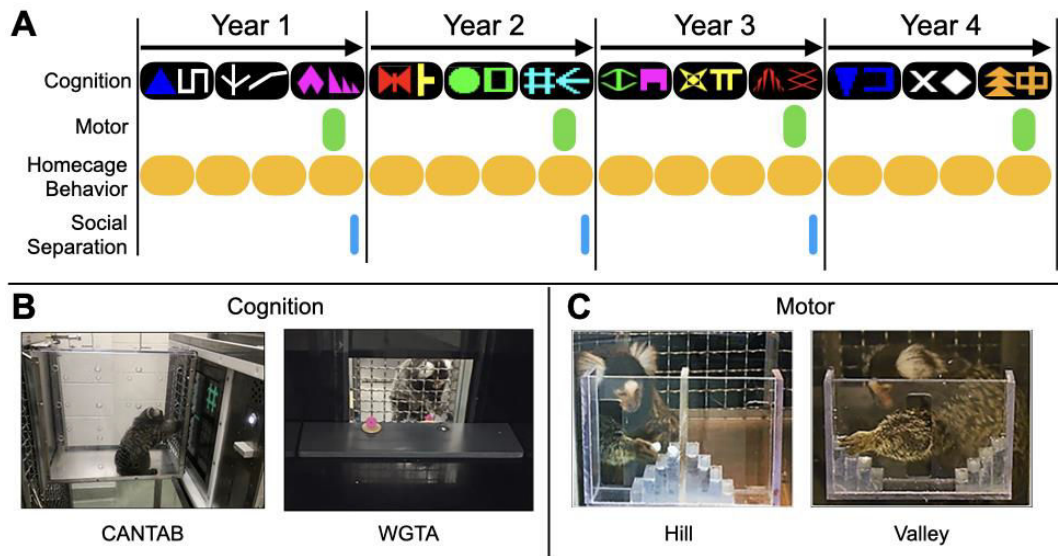


Figure 1. Study Timeline (A) and Apparatuses (B, C).

A: Cognitive testing in the format of Discrimination and Reversal testing occurred all 4 years with 3 stimuli sets per year. Motor testing was administered each year following completion of cognitive testing. Social separation testing occurred on one day per year during years 1 to 3 and did not overlap with any other testing. Behavior was observed weekly in the home cage. B: Cognitive apparatuses: CANTAB touchscreen (left) and WGTA (right); C: Motor apparatuses: Hill (left) and Valley (right).

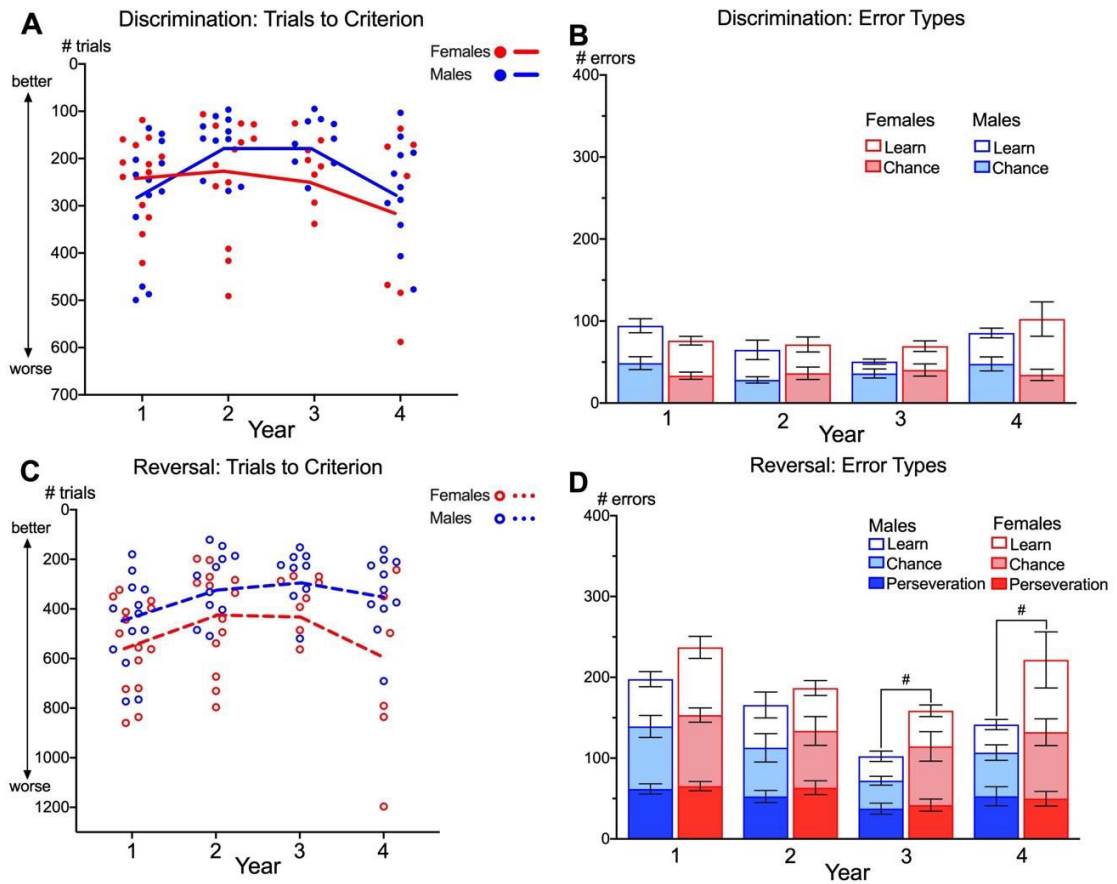


Figure 2. Sex differences in cognitive aging trajectories in marmosets.

Trials to criterion for each year of testing for males (blue) and females (red) on discrimination (A&B) and reversal (C&D). Points represent averaged scores for individual marmosets within each year. Lines represent average aging trajectories from multilevel growth models reported in Table A.5. Error types for discrimination (B) and reversal (D) are presented as means and SEM. # $p < 0.065$. Y-axes are inverted for A & C for readability.

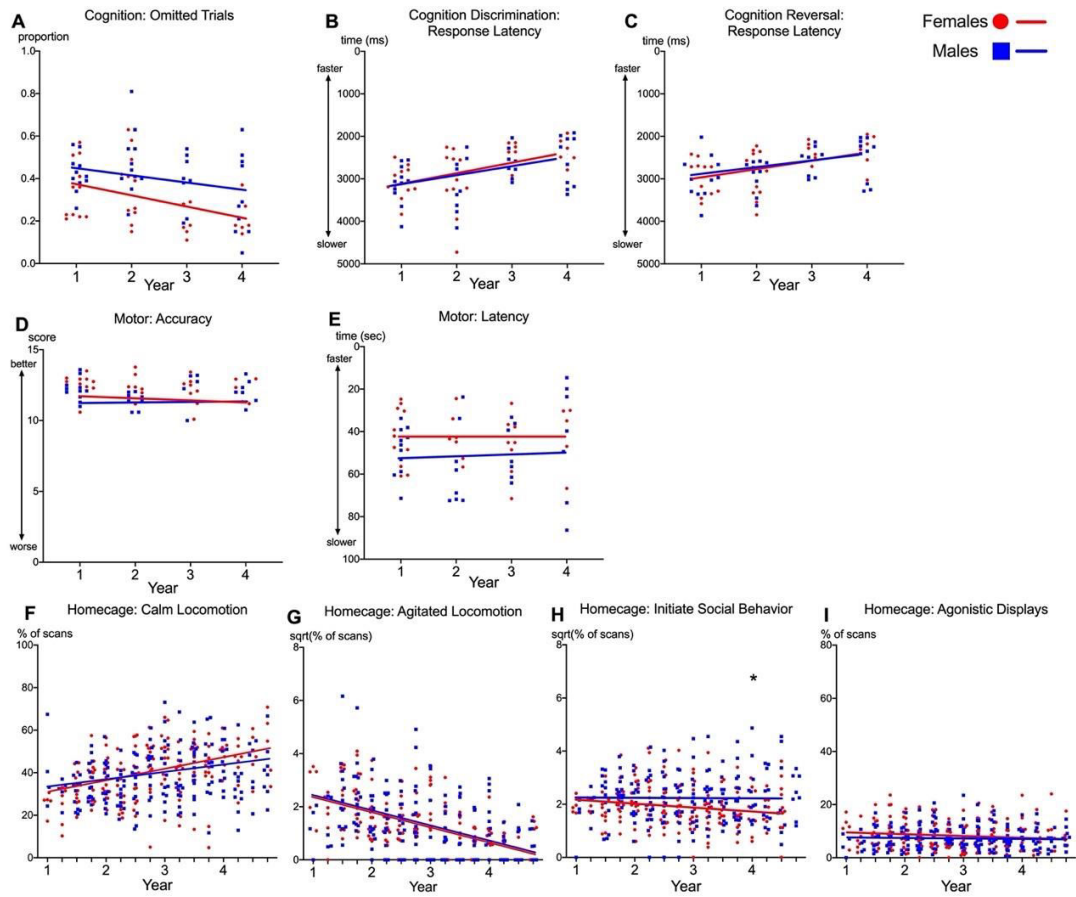


Figure 3. Aging trajectories for non-cognitive variables in marmosets.

Aging Trajectories in males (blue) and females (red) for: the proportion of omissions on all cognitive trials (A), response times for discrimination (B) and reversal (C) (Top panel); Motor accuracy and latency (D-E, Middle panel); Homeage behaviors (F-I, Bottom panel). Points represent averaged scores for individual marmosets. Lines represent average aging trajectories from multilevel growth models reported in Table A.6. Y-axes are inverted for B, C & E for readability.

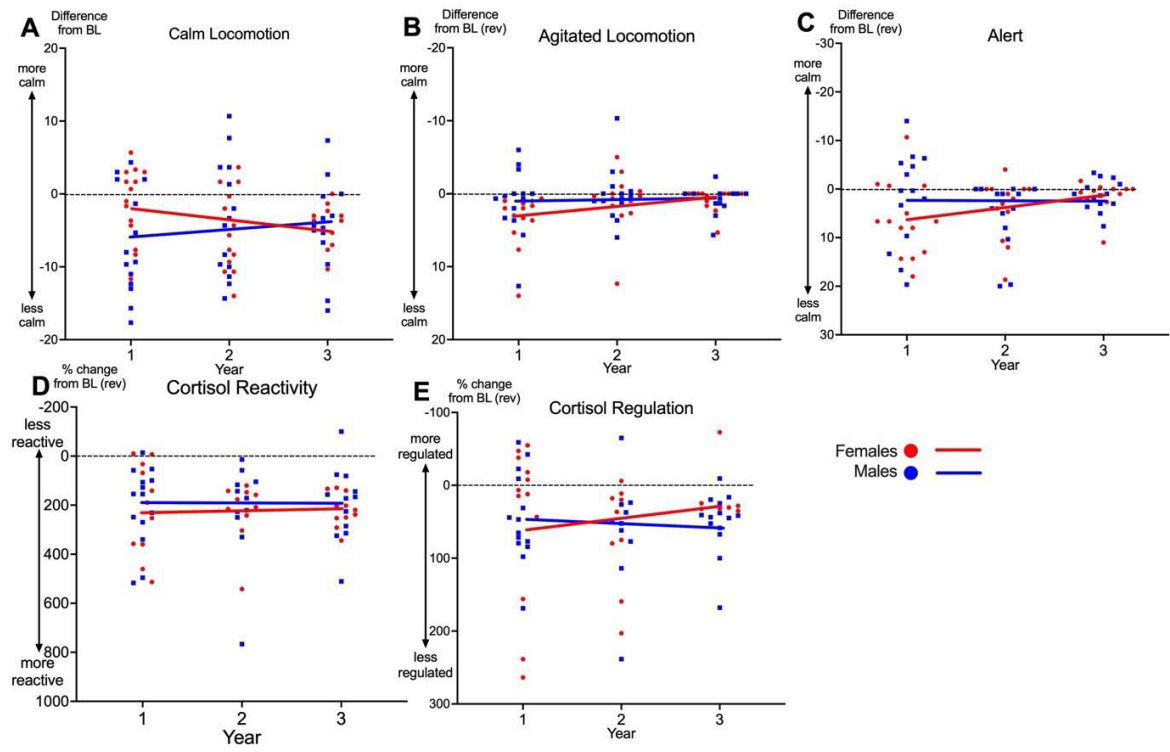


Figure 4. Behavioral and Physiological Responses to Social Separation Test.

Aging trajectories for males (blue) and females (red) for measures of behavioral (A-C) and physiological (D-E) reactivity, calculated as difference between pre-separation baseline (BL) and total separation score. Positive values represent increase from BL during separation and negative values represent decrease from BL. Zero represents no change from BL, which is indicated with a dotted line. Points represent averages scores for individual marmosets within each year. Lines represent average aging trajectories from multilevel growth models reported in Table A.6. Y-axes are inverted for B, C, D & E for readability.