

Evaluation of attention in APP/PS1 mice reveals impulsive and compulsive behaviours

Shepherd, A.¹, May, C.¹, Churilov, L.², Adlard, P.A.¹, Hannan, A.J.^{1,3}, Burrows, E.L.^{1,+}

1. Florey Institute of Neuroscience and Mental Health, Kenneth Myer Building, University of Melbourne, Parkville, VIC 3052, Australia

2. Florey Institute of Neuroscience and Mental Health, Melbourne Brain Centre, Heidelberg, VIC 3084, Australia

3. Department of Anatomy and Neuroscience, University of Melbourne, Parkville, VIC 3010, Australia

+corresponding author: Emma L. Burrows, contact: emma.burrows@florey.edu.au, [t+61390356629](tel:+61390356629), [f+61 3 9035 3107](tel:+61390353107)

Abstract

While Alzheimer's disease (AD) is traditionally associated with deficits in episodic memory, early changes in other cognitive domains, such as attention, have been gaining interest. In line with clinical observations, some animal models of AD have been shown to develop attentional deficits, but this is not consistent across all models. The APP^{swe}/PS1 Δ E9 (APP/PS1) mouse is one of the most commonly used AD models and attention has not yet been scrutinised in this model. We set out to assess attention using the 5-choice serial reaction time task (5CSRTT) early in the progression of cognitive symptoms in APP/PS1 mice, using clinically translatable touchscreen chambers. APP/PS1 mice showed no attentional changes across 5CSRTT training or any probes from 9-11 months of age. Interestingly, APP/PS1 mice showed increased impulsive and compulsive responding when task difficulty was high. This suggests that while the APP/PS1 mouse model may not be a good model of attentional changes in AD, it may be useful to study the early changes in impulsive and compulsive behaviour that have been identified in patient studies. As these changes have not previously been reported without attentional deficits in the clinic, the APP/PS1 mouse model may provide a unique opportunity to study these specific behavioural changes seen in AD, including their mechanistic underpinnings and therapeutic implications.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/gbb.12594](https://doi.org/10.1111/gbb.12594)

Introduction

Alzheimer's disease (AD) is the most common form of dementia, affecting approximately 50 million people worldwide¹. The disease appears to start decades prior to symptom development, starting with the accumulation of amyloid-beta ($A\beta$) that begins to form characteristic extracellular plaques. This is followed by the hyperphosphorylation of tau and subsequent formation of neurofibrillary tangles within neurons, and finally full gross atrophy of the brain². These, and other, molecular and cellular changes ultimately lead to progressive cognitive decline, initially presenting with episodic memory changes that are closely followed by executive function and attention deficits³⁻⁵. Ultimately, patients are impaired in all cognitive domains⁵. To date, drugs modifying the cholinergic system are standard care for AD, but these are only symptomatic treatments and furthermore, are not particularly effective⁶. In light of the growing burden of disease and lack of disease-modifying treatments, animal studies are crucial to both better understand the disease and to find more effective treatments.

In most cases of AD, the disease occurs sporadically; however, a small percentage of cases are due to mutations in the amyloid precursor protein (APP) or presenilin 1 or 2 (PS1/PS2) genes⁷. These familial mutations cause over-production of the $A\beta$ peptide and early development of the disease in human patients, and have been leveraged to develop many mouse models of AD, including the APP^{swe}/PS1 Δ E9 (APP/PS1) mouse⁸. This mouse model has two human mutations, the Swedish mutation in the APP proteins and the deletion of exon 9 in the PS1 gene, and is maintained either on a B6 or a B6/C3H background⁹. The APP/PS1 mouse has been characterised extensively using traditional behavioural tasks such as the Morris water maze and the novel object recognition task¹⁰, however attention has not been previously assessed in this model¹¹.

Changes in attention have recently been highlighted as an early symptom in AD and understanding how deficits in attention arise may provide insights into mechanisms that distinguish AD from normal aging¹². There are multiple subtypes of attention, including sustained, selective and divided attention. Sustained attention, sometimes referred to as vigilance, is the ability to pay attention to a task over a period of time¹³. Individuals with mild cognitive impairment (MCI) and early AD patients show deficits in sustained attention tasks^{14,15}. Specifically, mild AD patients have shown a vigilance decrement in a digit discrimination task before deficits in selective attention¹⁶, and MCI and AD patients show increased reaction time in multiple sustained attention tasks¹⁷. Additionally, early AD patients made more errors in a sustained attention to response task¹⁸ and multiple visual attention tasks¹⁹. The severity of these deficits correlated with broad measures of cognitive decline in both studies. Not only can attention deficits predict severity of cognitive decline, but preliminary evidence suggests that this facet of attention may be able to predict conversion of MCI patients to AD²⁰. In a 20-month longitudinal study of 107 control

and MCI participants, only reaction time on a simple sustained attention task declined consistently over the 20-month period. This study comprehensively tested verbal memory, visual memory, executive function and working memory and found them to be stable, indicating that attentional abilities decline prior to other cognitive domains in MCI, and thus may have predictive value for MCI to AD conversion²⁰. Finally, the anti-cholinergic drugs currently used for AD help more with attentional symptoms than memory symptoms²¹. Thus, sustained attention is affected early in AD and may provide information about disease progression, and as such is worth further investigation in animal models.

A well characterised rodent task, the 5-choice serial reaction time task (5CSRTT), has been used extensively in rodents to probe sustained attention as well as impulsive and compulsive behaviour²². In this task, the animal initiates a trial and after a delay, a stimulus is shown in one of five locations. The animal must respond to the location the stimulus was presented in to receive a reward. The task can be made more difficult to tax attention by reducing the stimulus duration or brightness or by changing the delay prior to the stimulus presentation. This test has been adapted for use in rodent touchscreen chambers, enabling high-throughput and standardised assessment of attention in rodents²³. Three AD mouse models have been characterised using the touchscreen version of the 5CSRTT; TgCRND8²⁴, TgTau^{P301L}^{25,26} and 3xTG-AD mice²⁷. These models have A β , tau or both A β and tau pathology respectively. All these models were trained in the 5CSRTT task, followed by a probe in which stimulus duration was reduced to tax attention. Both 4-5-month-old TgCRND8 and 9-month-old 3xTG-AD models showed reduced accuracy compared to WT animals at reduced stimulus durations^{24,27}. No changes in attention were uncovered in the TgTau^{P301L} model up to 17 months of age²⁵. This difference in phenotype expression could possibly indicate that it is the pathological A β protein and not the tau aggregates that may be responsible for attentional deficits in these mouse models, however further testing of this idea is required.

To date, the APP/PS1 mouse model has been widely used to understand mechanisms behind A β accumulation and memory decline, but attention has never been assessed. Attention was assessed in APP/PS1 mice at 9-11 months, an age where early cognitive decline and plaque accumulation has been previously described^{9,28-33}. In addition to taxing their attention by reducing stimulus duration, we also tested whether changing the delay before stimulus presentation or by reducing stimulus brightness would reveal different attentional deficits. These task manipulations can help elucidate impulsive and compulsive behaviour and have not been performed in any AD mouse model previously. Stimulus brightness can also test visual ability on top of probing sustained attention. We hypothesised that the APP/PS1 mouse model would recapitulate the previously described accuracy deficits in other AD mouse models during the stimulus duration probe and would likely show similar deficits across other task manipulations. This study represents the first characterisation of

attention, including impulsive and compulsive behaviour, in the APP/PS1 mouse model of AD.

Methods

Animals

Male APP^{swe}/PS1^{ΔE9} mice (Jax strain: 4462, n=11) and littermate controls (n=10) on a 50:50 C57BL/6;C3H hybrid background were bred on site at the Florey Institute of Neuroscience and Mental Health. Groups of mice were first housed in standard individually ventilated cages with a small shelter on a 7am:7pm light cycle until 6 months of age, after which they were moved to 16.5x31x11cm open top cages (Wiretainers, R.E. Walters, Melbourne, Australia) with a reversed light cycle (7pm:7am). Once acclimatised to this room, animals were weighed for 3 days to obtain a free-feeding weight (FFW), and then food restricted to achieve 85% FFW. Simultaneously, the mice began pre-training for behavioural testing. All procedures were approved by The Florey Institute of Neuroscience and Mental Health Animal Ethics Committee and were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes as described by the National Health and Medical Research Council of Australia.

Apparatus

Animals were tested in automated touchscreen-based operant systems (Campden Instruments Ltd, Loughborough, UK). Task deployment and automated event recordings were managed through the software Whisker Server and ABET II (Lafayette Instruments, Lafayette, IN, USA). Apparatus and task training methods have been published previously²³.

Behavioural procedures

Pre-training

All behavioural procedures were performed during the dark phase of the animals. Pre-training was conducted as described previously²³. Animals were first habituated to the touchscreen chambers with freely available rewards (Iced Strawberry Milk, Nippy's Ltd, Moorook, Australia) for 20 minutes for 2 consecutive days, then trained to nose-poke at a white square stimulus on the screen for the reward. Following this, animals were trained to initiate 30 trials via head entry into the food magazine and finally punished by a light turning on in the chamber for 5s time out, when it touched the screen anywhere but the stimulus square. The reward volume for a correct response at all stages was 7μL, and the inter-trial interval for both correct and incorrect trials was 5s. Animals were 6 months of age at the beginning of pre-training.

5-Choice Serial Reaction Time Task (5CSRTT) training

The 5CSRTT has been described in detail elsewhere²³. Briefly, the animal must voluntarily initiate the trial with a nose poke into the reward magazine, after which there is a delay (baseline=5s) followed by a stimulus presentation in one of five locations. The animal can then respond, either correctly, where it is presented with a tone and reward, or incorrectly, where it is presented with a light and 5s time out where it cannot initiate a new trial. The animal can also not respond (an omission) or respond before the stimulus is presented (premature response), in which case they are also presented with a light and time-out. The task is summarised in Figure 1a. If animals continued to press the screen after an initial response, numbers of touches were recorded as a perseverative response. Animals were presented with 50 trials or allowed to perform for up to 1 hour per session, with sessions performed 6 days a week.

During training, the initial stimulus duration was 32s. The animal was held at this stage until they reached criterion (all 50 trials, >80% correct responses and <20% omissions), at which point the stimulus was halved until they reached a 2s stimulus duration (32, 16, 8, 4 and 2s). This 2s condition was the baseline condition needed to continue into the probe stages of the 5CSRTT. Due to the age of the animals when food restriction was started, FFWs were not consistently 85% until the 2s training stage was reached, thus data to this point was not able to be scrutinised. Animals were 8.5 months of age at the end of training. A timeline of training and probes, including the age of animals during those probes, is presented in Figure 1b.

5CSRTT probes

Once all animals were performing at baseline conditions, they were moved into a series of 4 probes. Probes were performed for 2 consecutive days at a time, with baseline sessions in between to maintain performance levels. During baselining, animals were required to reach criterion for 2 consecutive days (all 50 trials, >80% correct responses and <20% omissions). Animals were repeatedly tested at baseline until the final animal reached criterion for the 're-baseline'. Thus, all animals completed equivalent numbers of baseline sessions between the probes and commence each probe at the same age. During probes, animals were tested 7 days a week.

The stimulus duration probe allowed assessment of sustained, visuospatial attention. Mice were assessed at reduced stimulus durations of 1, 0.8, 0.6, or 0.4s, in a block design for 2 consecutive days. First, mice were tested on different stimulus durations in a semi-randomised Latin square design to control for order effects. The delay probe was presented next and allowed assessment of impulsive and compulsive behaviours by altering the delay prior to stimulus presentation and monitoring premature and perseverative responses. Animals were presented with two 2-day blocks of either 'short delays' (2, 3, 4 and 5s) or 'long delays' (5, 6, 7 and 8s), in

pseudo-randomised order between animals with baseline sessions in between, ultimately completing 4 days of each delay block. The inclusion of the baseline delay value of 5s in each probe, allowed the effect of shortening and lengthening the delay from baseline value. . Presented third, the brightness probe tested sustained, visuospatial attention as well as visual ability by reducing contrast of each stimulus, with the 5 different levels of brightness (100%, 50%, 25%, 12.5% and 6.25%) randomly presented within a session. Finally, the stimulus duration probe was repeated. Animals were 11.5 months at the end of the probe sessions.

Progressive Ratio (PR)

Following all 5CSRTT probes, 11.5-month-old animals were trained on fixed ratio 2, 3 and 5 paradigms, where the animal must touch the stimuli on the screen 2, 3 or 5 times respectively to receive one reward, before moving on to a progressive ratio 4 (PR) task. Here, the number of touches the animal must make to the screen increases by $n+4$ for every reward they collect (e.g. 1 touch, 5 touches, 9, 13, 17 etc.), and the point where the animal stops responding for 5 minutes or the point it reaches after 1 hour, called the breakpoint, taken as a measure of motivation for the reward. Animals were tested for 8 days on the PR task, then average breakpoint calculated both per day and overall. Overall average number of touches to screen and reward collection latencies per day were also analysed.

Statistical Analysis

For progressive ratio data, Student's t-test were used to compare overall breakpoint and touches and 2-way repeated measures ANOVA's were used for reward collection latencies and breakpoint over the 8 days of testing. Progressive ratio data is presented as mean \pm SEM (breakpoint and number of touches) and median \pm 95% CI (reward collection latencies). 5CSRTT data was analysed at the level of trial using generalised linear, latent and mixed models (GLLAMM). GLLAMM were used to implement a random effect analysis, with robust standard error estimation and individual animals treated as random effects. These models were used not only as they enable the analysis on a trial-by-trial basis, but also because they work well on non-normally distributed data and can also account for the fact that some animals may be run for a different number of days or may not complete the same number of trials every day. GLLAMM were run with genotype, day, trial, timestamp, and, if required, the relevant probe variable (stimulus duration, delay block (ie. long delay block and short delay block), or stimulus brightness) as explanatory variables in the model, while the odds of any response, a correct response, a premature response and the expected number of perseverative responses as outcome variables. Two additional explanatory variables investigating the effect of vigilance (trial bin, generated by splitting the number of trials completed into 5 equal bins) and the interaction of trial bin and genotype (trial bin*genotype) were included in the odds of response and odds of correct response model, to investigate sustained attention across a single session. GLLAMM analysis was performed to test if genotype and/or a probe would affect

performance of animals on any outcome variables.

Specifically, logistic regressions were used to analyse the binary variables ‘any response’, ‘correct response’, or ‘premature response’ with corresponding effect sizes reported as adjusted odds ratios (aORs) with respective 95% confidence intervals (95% CIs). To interpret these for genotype effects, an aOR of smaller than 1 indicates that APP/PS1 mice are less likely to have responded to a trial, chosen correctly, or made a premature response, while the corresponding 95% CI not including 1 represents statistical significance ($p < 0.05$). Poisson regressions were used to analyse count data, or the number of perseverative responses, with corresponding effect sizes reported as adjusted incidence rate ratios (aIRRs) with respective 95% CIs. The interpretation for aIRR being larger than 1 is that APP/PS1 mice have a higher expected count of perseverative responses compared to control levels. To interpret the effect of probe, an aOR less than 1 indicates that as stimulus duration/delay/stimulus brightness decrease, the likelihood of a mouse correctly selecting, responding, making a premature response will also reduce. Similarly, an aIRR less than 1 indicates the expected count of perseverative responses is lower. Median regressions were used to analyse continuous variables, such as response and reward collection latencies, with corresponding effect sizes reported as median differences (coefficients) in seconds with respective 95% CIs. For these trial-by-trial analysis, odds ratios and incidence rate ratios are presented as effect sizes with a 95% CI, while median coefficients in seconds with 95% CI. For 5CSRTT mouse level data visualisation purposes only we present mean \pm SEM for accuracy (excluding omitted trials), response rate, premature responses/trial, perseverative responses/trial, response latency and reward collection latency. All statistical analyses were conducted using STATA v13IC (StataCorp, College Station, TX USA).

Results

APP/PS1 mice show equivalent training and baseline performance to WT mice on the 5CSRTT

On the final 2s stage of training, WT and APP/PS1 mice showed no differences in accuracy (Figure 2a, aOR= 1.154 (95% CI: 0.777; 1.713), $p=0.478$) or response rate (Figure 2b, aOR= 1.195 (95% CI: 0.898; 1.590), $p=0.222$). There were also no differences in likelihood of a premature response (Figure 2c, aOR= 0.952 (95% CI: 0.481; 1.883) $p=0.887$) or the expected count of perseverative responses (Figure 2d, aIRR= 1.056 (95% CI: 0.753; 1.481), $p=0.751$). Adjusted odds ratios and incidence rate ratios for these four measures are presented in Figure 2g. Finally, no differences in correct response latency (Figure 2e, Median difference= -0.0560 (95% CI: -0.218; 0.106), $p=0.499$) or reward collection latency (Figure 2f, Median difference= 0.019 (95% CI: -0.076; 0.114), $p=0.693$) were seen, implying no differences in motor performance between APP/PS1 and WT mice. During the probe period, animals

performed numerous baseline 2s sessions. To test for changes over time, accuracy and response rate on baseline days between probes were assessed with no differences seen across the entire 5CSR TT probe period (Supplementary Figure 1; accuracy: aOR= 0.844 (95% CI: 0.549; 1.296), $p=0.438$); response rate: aOR= 1.003 (95% CI: 0.725; 1.388), $p=0.986$). The 3xTG-AD model has shown increased perseverative responding only following a correct, not an incorrect or premature trial in the 5CSR TT previously²⁷. To investigate this, a trial type and genotype interaction variable was generated and run through the perseverative response GLLAMM, with no significant effect of the interaction variable (Supplementary Table 1), indicating that the genotypes had equal counts of perseverative responses following different trial types. Additionally, APP/PS1 and WT mice showed the equivalent number of beam breaks in the touchscreen chambers during the 2s training sessions, signifying that both genotypes are moving in the chamber a similar amount (Supplementary Table 2).

APP/PS1 mice have slightly higher responding during stimulus duration

probe 1

After training, 9-month-old APP/PS1 and WT animals were subjected to a series of probe days where the stimulus duration was shortened from 2s to 1, 0.8, 0.6 or 0.4s to tax sustained attention. Accuracy (Figure 3a and f, aOR= 0.746 (95% CI: 0.720; 0.774), $p<0.001$) and response rate (Figure 3b and f, aOR= 0.766 (95% CI: 0.713; 0.823), $p<0.001$) decreased as stimulus duration was decreased as expected. No effect of genotype was seen on accuracy (Figure 3a and e, aOR= 1.099 (95% CI: 0.818; 1.476), $p=0.531$), but an increase in the odds of response in APP/PS1 mice was seen (Figure 3b and e, aOR= 1.493 (95% CI: 1.021; 2.183), $p=0.039$). Premature and perseverative responses were unchanged by genotype (Figure 3c-e, aOR= 1.822 (95% CI: 0.967; 3.432), $p=0.063$; aIRR= 1.209 (95% CI: 0.868; 1.684), $p=0.262$), and were unchanged and increased respectively when stimulus duration was reduced (Figure 3c-e, aOR= 0.898 (95% CI: 0.797; 1.012), $p=0.077$; aIRR= 1.196 (95% CI: 1.125; 1.271), $p<0.001$). Vigilance across a session was assessed by splitting the trials into 5 bins and assessing the effect of a trial bin*genotype interaction variable on accuracy and response rate across those bins. No effect of genotype was seen across bins for odds of correct, (Supplementary Figure 2; aOR= 1.058 (95% CI: 0.941; 1.187), $p=0.347$), but there was an effect of the interaction variable on the odds of response (aOR= 1.173 (95% CI: 1.071; 1.286), $p=0.001$), in line with the increased odds of response seen across the entire probe. APP/PS1 and WT mice showed similar numbers of beam breaks throughout the stimulus duration probe, indicating equivalent activity (Supplementary Table 2) Perseverative responding following the 3 different trial types (correct, incorrect, and premature) was also assessed, with no effect of the genotype-trial type interaction variable (Supplementary Table 1).

APP/PS1 mice perform equivalently to WT mice when the delay before stimulus presentation is changed

During the delay probe, mice were tested at delays shorter or longer than the baseline 5s. No differences between WT and APP/PS1 mice were seen in accuracy (Figure 4a and g, aOR= 0.992 (95% CI: 0.619; 1.590), p=0.974), response rate (Figure 4b and G, aOR= 1.135 (95% CI: 0.792; 1.626), p=0.491), premature responses (Figure 4c and g, aOR= 1.215 (95% CI: 0.719; 2.052), p=0.466) or perseverative responses (Figure 4d and g aIRR= 1.326 (95% CI: 0.830; 2.117), p=0.238). When assessing the effect of the delay probes on mice regardless of genotype, short delays increased the odds of correct selection (Figure 4e, aOR= 1.438 (95% CI: 1.050; 1.969), p=0.024), and both APP/PS1 and WT mice showed decreased odds of responding (i.e. response rate) (Figure 4e, aOR= 0.700 (95% CI: 0.563; 0.869), p=0.001), as well as less premature responses (Figure 4e, aOR= 0.179 (95% CI: 0.111; 0.290), p<0.001). The expected number of perseverative responses was unchanged by shorter delays (Figure 4e, aOR= 0.858 (95% CI: 0.629; 1.170), p=0.332). When tested at longer delays mice, regardless of genotype, exhibited decreased odds of correct selection (Figure 4f, aOR= 0.645 (95% CI: 0.491; 0.849), p=0.002), increased odds of premature responses (Figure 4f, aOR= 1.424 (95% CI: 1.022; 1.982), p=0.037) and an increased number of perseverative responses (Figure 4f, aIRR= 5.846, 95% CI: 4.017; 8.508, p<0.001). Odds of response were unchanged at longer delays (Figure 4f, aOR= 0.903 (95% CI: 0.730; 1.116), p=0.343). Vigilance across delay probes and perseverative responding following different trial types were also assessed, with no effect of the genotype-interaction variable observed in either case (Supplementary Figure 2; Accuracy: aOR= 0.915 (95% CI: 0.800; 1.047), p=0.195. Response: aOR= 1.088 (95% CI: 0.917; 1.291), p=0.334, and Supplementary Table 1 respectively). WT and APP/PS1 mice showed similar levels of activity throughout the delay probe, as assessed by beam breaks (Supplementary Table 2)

APP/PS1 mice show increased premature and perseverative responding when stimulus brightness is reduced

Stimulus brightness was reduced from 100% to 6.25% to increase attentional load. Approximately chance performance (1 in 5, or 20%) was observed at the 6.25% brightness, indicating that the mice were unable to see the stimulus and were thus responding randomly. Accuracy (Figure 5a and e-f, aOR= 0.614 (95% CI: 0.588; 0.641), p<0.001) and response rate (Figure 5b and e-f, aOR= 0.714 (95% CI= 0.685; 0.744), p<0.001) were dramatically reduced at the 3 lowest levels of stimulus brightness for both APP/PS1 and WT mice, with no effect between genotypes (Figure 5 a-b and e; correct: aOR= 0.779 (95% CI: 0.539; 1.127), p=0.185, response: aOR= 1.109 (95% CI: 0.881; 1.396), p=0.378.), indicating equivalent sustained and visual attentional abilities. Interestingly, both an increased odds of a premature responses

(Figure 5c and f, aOR= 2.872 (95% CI: 1.640; 5.031), $p < 0.001$) and an increased number of perseverative responses (Figure 5d and f, aIRR= 1.417 (95% CI: 1.080; 1.859), $p = 0.012$) were seen in APP/PS1 animals across the probe. This indicated that in high-attentional load situations, APP/PS1 mice are more impulsive and compulsive than their WT counterparts. When assessing the effect of the stimulus brightness probe regardless of genotype, both WT and APP/PS1 mice showed an increased number of perseverative responses as stimulus brightness was decreased, (Figure 5d and f, aIRR= 1.450 (95% CI: 1.368; 1.537), $p < 0.001$), while premature responses were unchanged (Figure 5c and f, aOR= 1.020 (95% CI: 0.957; 1.086), $p = 0.548$). Vigilance across the different stimuli brightness levels was assessed, with a significant interaction between trial bin and genotype on odds of correct response (Supplementary Figure 2; aOR= 0.866 (95% CI: 0.769; 0.975), $p = 0.018$) but not odds of response (Supplementary Figure 2, aOR= 1.098 (95% CI: 0.984; 1.225), $p = 0.095$). Perseverative responding following the 3 different trial types (correct, incorrect, and premature) was also assessed, with no effect of the genotype-trial type interaction variable (Supplementary Table 1). Equivalent numbers of beam breaks between WT and APP/PS1 mice indicate similar levels of activity throughout the brightness probe (Supplementary Table 2).

APP/PS1 mice show equivalent performance on stimulus duration probe 2

At 11 months of age, mice were retested on the stimulus duration probe. As before, APP/PS1 mice and WT mice both show unchanged measures on accuracy (Figure 6a and e, aOR= 0.993 (95% CI: 0.682; 1.446), $p = 0.972$), premature responses (Figure 6c and e, aOR= 0.916 (95% CI: 0.538; 1.560), $p = 0.747$) and perseverative responses (Figure 6d and e, aIRR= 1.096 (95% CI: 0.807; 1.489), $p = 0.557$). In contrast with the increased response rate observed in 9-month-old APP/PS1 mice in the first stimulus duration probe, response rate at 11 months of age was equal between genotypes (Figure 6b and e, aOR= 1.197, (95% CI: 0.872; 1.642) $p = 0.267$). Regardless of genotype, decreasing stimulus duration decreased odds of correct selection (Figure 6a and f, aOR= 0.765 (95% CI: 0.740; 0.792), $p < 0.001$) or response rate (Figure 6b and f, aOR= 0.783 (95% CI: 0.741; 0.828), $p < 0.001$), increased the incidence of perseverative responses (Figure 6d and e, aIRR= 1.220 (95% CI: 1.167; 1.276), $p < 0.001$) and did not change the odds of a premature response (Figure 6c and f, aOR= 0.940 (95% CI: 0.847; 1.043), $p = 0.244$). This mirrors the effect of reducing stimulus duration in the first stimulus duration probe. As with all probes so far, activity was equal between APP/PS1 and WT mice, as assessed by beam breaks (Supplementary Table 2). Vigilance across delays and perseverative responding following different trial types were also assessed, with no effect of genotype observed (Supplementary Figure 2; Accuracy; aOR= 1.047 (95% CI: 0.927; 1.184), $p = 0.459$; Response; aOR= 1.010 (95% CI: 0.899; 1.134), $p = 0.87$) and Supplementary Table 1 respectively). As this probe was a repeat of that performed at 9 months, we compared performance of APP/PS1 and WT mice between the two probes. During the repeated probe, mice showed a reduced likelihood that they would respond to a trial

(Supplementary Figure 3b and e, aOR= 0.766 (95% CI= 0.662; 0.886), $p<0.001$) although this was independent of genotype (Supplementary Figure 3b and f, aOR= 0.76 (95% CI= 0.531; 1.086), $p=0.132$). Furthermore, both groups showed reduced odds of a premature response (Supplementary Figure 3c and e, aOR= 0.657 (95% CI= 0.446; 0.97), $p=0.034$) and decreased expected count of perseverations (Supplementary Figure 3d and e, aIRR= 0.69 (95% CI= 0.594; 0.801), $p<0.001$). Again, both of these changes were genotype independent (Supplementary Figure 3c, d and f, premature: aOR= 0.574 (95% CI= 0.324; 1.017), $p=0.057$, perseverative responses: aIRR= 0.927 (95% CI= 0.645; 1.332), $p=0.682$), indicating both WT and APP/PS1 mice showed the decreased response rate, premature response rate and number of perseverations during the second stimulus duration probe compared to the first stimulus duration probe. The odds of correct response were the same across the repeated probe (Supplementary Figure 3a and e, aOR= 1.094 (95% CI= 0.881; 1.358), $p=0.417$), regardless of genotype (Supplementary Figure 3a and f, aOR= 0.862 (95% CI= 0.603; 1.234), $p=0.418$)

APP/PS1 mice show slightly higher breakpoints than their WT counterparts on the progressive ratio task

To assess level of task participation, mice were subjected to the progressive ratio task at the end of 5CSRRT testing. APP/PS1 mice showed a slightly elevated average breakpoint compared to their WT counterpart (Supplementary Figure 4a, t-test, $t(18)=2.401$, $p=0.027$) although this effect was not significant across the week of testing (Supplementary Figure 4b, two-way ANOVA, $F(1,18)=3.837$, $p=0.066$). Overall, APP/PS1 mice made about 35 more touches during a session compared to their WT counterparts (Supplementary Figure 4c, $t(18)=2.261$, $p=0.036$). Reward collection latency, which has been interpreted as a measure of motivation, was unchanged between the WT and APP/PS1 mice (Supplementary Figure 4d, $F(1,18)=0.0436$, $p=0.837$).

Discussion

In this study, we have shown that APP/PS1 mice do not show attention impairments in the 5CSRRT, at either baseline or when task parameters were manipulated to tax attention. Some subtle changes were evident in APP/PS1 mice, including a small increase in response rate and increased premature and perseverative responding, however these increases were not seen consistently across training and all variations of the task. Additionally, a slight increase in breakpoint in PR was observed. This study represents the first analysis of attention in the APP/PS1 mouse model of AD and suggests that this model is not ideal for studying early attention impairments associated with slow amyloid accumulation.

The lack of change in accuracy and response rate in the 5CSRRT performance by APP/PS1 mice is somewhat surprising, as two amyloid-driven AD mouse models, the

TgCRND8 and 3xTG-AD mice, have shown deficits in this task following plaque build-up at 4-5 months and 9 months of age respectively^{24,27}. It is possible that in this study, APP/PS1 mice were assessed before the onset of attentional impairments. Against this idea, studies scrutinising other facets of executive function in APP/PS1 mice have reported reversal impairments at 6-8 months of age in the T water maze^{28,34} and Morris water maze^{9,30,35,36}, however these are not directly comparable to attention as assessed by 5CSRTT. APP/PS1 mice were tested between 9-11 months, an age deemed to be an equivalent stage of cognitive deficit to the 5CSRTT assessments of the 3xTG-AD and TgCRND8 mice. At 9 months of age, APP/PS1 mice have been shown to have deficits in spatial memory, reversal, fear conditioning and ethologically relevant behaviours like burrowing and nest building as well as increased amyloid beta (plaques, A β 40 and A β 42 oligomers), neuroinflammation and tau hyperphosphorylation in the brain^{9,29,30}. However, age of onset of certain cognitive deficits can be hugely varied across studies; for example, Stewart et al³⁷ compared when spatial memory deficits occurred over various tasks in the Tg2576 model of AD, and showed that only 64% of studies show MWM deficits at 8 months, indicating that the age in which a deficit is seen in one study is not necessarily indicative of when it will occur in another. Furthermore, behavioural results have been shown to vary considerably between laboratories³⁸ and between experimenter^{39,40}, and it is possible that APP/PS1 mice housed in our facility would show impairments in attention later in disease progression. Thus, it is still possible that an older cohort of APP/PS1 mice would mirror attention deficits found previously in other AD mouse models and further testing of attention in animals beyond 12 months of age is warranted. Further to this, it is critical to extend this study to female mice. Various animal studies have indicated that female AD mouse models progress in a divergent way to male AD mice, showing accelerated A β build up and earlier onset of cognitive changes, but a longer lifespan^{41,42}. The choice to use only males in this study was not due to the commonly held misconception that the behaviour of female mice is more variable due to their oestrus cycle⁴³, but because of unavoidable equipment limitations that restricted the study to touchscreen chambers where only male mice are tested. Future studies should be conducted in female mice to better understand how AD affects women and improve pre-clinical translation.

Methodological differences between this study and those scrutinising attention in the 3xTG-AD and TgCRND8 mouse models could also account for differences in their performance⁴⁴. 3xTG-AD and TgCRND8 mouse models, were tested in Med Associates operant chambers, which are square in shape. Conversely, mice in our study were assessed in Bussey-Saksida Campden trapezoid shaped chambers. The trapezoidal design was intended to improve the animals ability to pay attention to the screen, making it easier to learn and perform tasks in these chambers, although to our knowledge this hypothesis has not been directly tested²³. Thus, the trapezoid shaped chamber may have decreased the difficulty of the 5CSRTT and as a result could account for normal attention seen in APP/PS1 animals in the present study. In line with this, TgTau^{P301L} mice^{25,26} were also tested in the trapezoid shaped chambers and

showed no changes in 5CSRTT performance. It is still possible that our original hypothesis, that attentional impairments in the 5CSRTT are linked to amyloid (not tau) pathology, was correct, but this seems less likely given the present study. Another experimental factor that could potentially alter the difficulty of the 5CSRTT is the light phase in which they are tested. While APP/PS1 mice in the present study were tested during their dark phase, TgCRND8 mice were tested during the light phase, when they were less active. Testing during the light phase has been shown to increase error rates on a tone discrimination task in C57BL/6 mice, compared to those tested in the dark phase⁴⁵. Could testing during the light phase unmask a subtle phenotype in TgCRND8 mice? While testing during the dark phase has many advantages including timing cognitive assessment with the animal's active period and not disrupting normal sleeping patterns, mice are exposed to brief intermittent light from both stimuli and chamber illumination following incorrect responses. While there is evidence that dim light can disrupt circadian rhythm in rodents, such disruptions have only been shown when animals were exposed continuously throughout the 12-hour dark phase⁴⁶. It is possible that the brief light exposure in touchscreens in this study could have disrupted circadian rhythms in the APP/PS1 animals, however if this was deleterious, we would have expected to see impairments in accuracy in the task. Differences in the caloric value of the reward has also been shown to affect touchscreen performance⁴⁸, and differences in the reward used for the studies scrutinising attention in AD mice could account for divergent phenotypes. While APP/PS1 and TgTau^{P301L} mice received strawberry milk rewards, 3xTG-AD and TgCRND8 mice received sugar pellets. A study comparing plain and strawberry milk to low fat milk demonstrated that mice not only have a lower breakpoint for low-fat milk, but are slower to learn a task in the touchscreens⁴⁸. The plain and low-fat milk had equal amounts of sugar while the strawberry and plain milk have equal calorie content, indicating that calorie content, rather than flavour, are an important driver of both motivation and learning. Motivation has been shown to impact cognitive performance; one study has shown mice perform better on a sustained attention task when there is a higher likelihood of reward⁴⁹, while another shows that higher value rewards evoke more neuronal activity in the visual cortex of monkeys, and these neurons exhibited a strong attention effect⁵⁰. This indicates that reward value can have a strong effect on performance. As such, strawberry milk may confer more reward value than sugar pellets, making the strawberry milk rewarded version of the 5CSRTT used in this study easier. This could mask any subtle performance differences between APP/PS1 and WT mice.

Inherent differences in the extended behavioural phenotype between the 3 amyloid driven AD mouse models could also explain the disparate results. For example, 3xTG-AD mice have exhibited increased anxiety at similar ages to their attentional deficits⁵¹⁻⁵⁴. Furthermore, hyperactivity has been reported in 3xTG-AD male mice⁵⁵ and TgCRND8 mice⁵⁶, potentially confounding touchscreen performance. Increased hyperactivity or anxiety could easily confound the attentional ability of an animal and are thus important parameters to consider. Despite these

reports, one of the 5CSRTT touchscreen assessments of TgCRND8 mice reported normal activity (measured by total beam breaks²⁴), while it was not assessed in the 3xTG-AD study²⁷. Inconsistencies in the phenotype of AD mouse models are not uncommon^{11,57}. This extends to reports of anxiety in APP/PS1 mice, with some studies showing increased anxiety^{29,58}, while others show no change³² or even decreased anxiety^{59,60}. Hyperactivity has been reported to be increased in APP/PS1 mice at 6 months⁵⁸, however, in another study, hyperactivity was reported at 12 months³¹. In our studies, we showed no changes in total beam breaks between WT and APP/PS1 mice, indicating normal activity in our cohort. These inconsistencies in phenotype descriptions make it difficult to clearly interpret how other behavioural impairments could interact with the expression of attentional abnormalities in touchscreen studies.

Other potential reasons for the absence of attentional deficit in APP/PS1 mice in our study could be intrinsic to touchscreen testing. First, caloric restriction is essential to motivate mice in touchscreen testing, but caloric restriction has considerable effects on longevity and has been shown to improve cognition in AD mouse models⁶¹. Specifically, caloric restriction has been shown to decrease amyloid build up in APP/PS1 mice⁶², 3xTG-AD⁶³ and other animal models of AD⁶⁴⁻⁶⁷. Another confounding element of touchscreen testing is the daily training involved in both learning and performing the multiple probes of the 5CSRTT. Various styles of computerised cognitive training are being investigated as a therapy in AD⁶⁸. Studies in MCI/AD patients have shown that training in a particular domain (e.g. attention) leads to improvements on performance on all tasks in that domain⁶⁹. The touchscreen training employed in this study was extensive, and as mice received daily sessions for 6 months, we can infer the effect of long training. Accuracy was unchanged both across all baseline sessions between probes and between the two stimulus duration probes. Thus, it appears that once animals reached criterion on training their performance was relatively stable. While stable performance is indicative of lack of cognitive improvement with increasing exposure to training on touchscreens, it is still possible that it masked a subtle deficit in APP/PS1 mice or delayed the expression of any impairment.

While no attentional phenotypes were noted in APP/PS1 mice, we did detect a slight increase in premature and perseverative responding during the brightness probe. This increased perseveration phenotype is partially in agreement with previously published results from 3xTG-AD mice, where perseverative responding was increased following correct responses (but not incorrect or premature) in the stimulus duration probe. No changes in perseveration were observed in 5CSRTT assessments of the TgCRND8 model²⁴. Premature and perseverative responding is traditionally measured during the delay probe as a proxy of impulsive and compulsive behaviour. During the delay probe in this study, high performance (>90% accuracy) in both WT and APP/PS1 mice may have obscured any differences in this behaviour. APP/PS1 mice were particularly taxed during the brightness probe, as it was the only case in which

they performed at chance level. It could be that increased perseverative behaviour in this mouse model is unmasked when they are challenged sufficiently. In support of this interpretation is the fact that 3xTG-AD mice showed an impairment in the probe in which increased perseveration was also reported, indicating that the set of task parameters were particularly challenging and thus lead to perseveration. In conflict with this, the TgCRND8 mouse model showed no increase in perseveration despite a significant impairment in attention. However, perseverative responding was not scrutinised at the level of trial type in this study, and thus may not have been identified. Finally, the slight increase in breakpoint in progressive ratio could also be attributed to increased perseverative behaviour. Interestingly, no increases in premature responses in AD models have been noted previously. Premature responding is indicative of impulsive behaviour and, in line with our results, another mouse model expressing multiple APP mutations, APP-KI mice, exhibit increased impulsive and compulsive behaviour before developing attentional deficits in a serial reaction time task⁷⁰. Our findings, together with previous evidence, suggest that perseveration and potentially subtle impulsivity may be caused by some APP mutations. Furthermore, these data demonstrate that when impulsive and compulsive behaviour is expressed in these mouse models, it appears to occur prior to attention deficits.

Impulsive and compulsive behaviour have not been extensively studied in AD patients, but there are some relevant studies. First, impulsivity has been noted in patients, with a lack of premeditation (i.e. to reflect on consequences) shown on an impulsivity scale correlating with the increased number of false alarms on a Go/NoGo task⁷¹. This is analogous to the impulsive behaviour seen in the APP/PS1 mice. Secondly, some level of compulsive behaviour has been noted in AD patients, including repetitive responding to a stimulus⁷², which again is analogous to the behaviour noted in the APP/PS1 mice in this study. Thus, the impulsive and compulsive behaviour seen in the APP/PS1 mice could be clinically relevant to AD patients.

In conclusion, APP/PS1 mice show normal attention at 9-11 months on the 5CSRTT task and probes, with an indication of increased impulsive and compulsive behaviour when attention is taxed. It remains to be determined if this absence of attentional phenotype will be unmasked at later disease progression and also whether any cognitive deficits occur prior to attentional deficits. This warrants further investigation to evaluate the validity of touchscreen testing as a translational tool in rodent models of AD. Furthermore, our findings provide further insights into the utility, and potential limitations, of this particular AD model in preclinical drug trials.

References

1. Alzheimer's Disease International. World Alzheimer Report 2015: The Global Impact of Dementia. <http://www.alz.co.uk/research/world-report-2015>. Published 2015. Accessed November 3, 2015.
2. Villemagne VL, Burnham S, Bourgeat P, et al. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol*. 2013;12(4):357-367. doi:10.1016/S1474-4422(13)70044-9
3. Foldi NS, Lobosco JJ, Schaefer LA. The effect of attentional dysfunction in Alzheimer's disease: theoretical and practical implications. *Semin Speech Lang*. 2002;23(2):139-150. doi:10.1055/s-2002-24990
4. Salmon DP, Bondi MW. Neuropsychological Assessment of Dementia. *Annu Rev Psychol*. 2009;60:257-282. doi:10.1146/annurev.psych.57.102904.190024
5. Weintraub S, Wicklund AH, Salmon DP. The Neuropsychological Profile of Alzheimer Disease. *Cold Spring Harb Perspect Med*. 2012;2(4). doi:10.1101/cshperspect.a006171
6. Yiannopoulou KG, Papageorgiou SG. Current and future treatments for Alzheimer's disease. *Ther Adv Neurol Disord*. 2013;6(1):19-33. doi:10.1177/1756285612461679
7. Musiek ES, Holtzman DM. Three dimensions of the amyloid hypothesis: time, space and "wingmen." *Nat Neurosci*. 2015;18(6):800-806. doi:10.1038/nn.4018
8. Jankowsky JL, Slunt HH, Ratovitski T, Jenkins NA, Copeland NG, Borchelt DR. Co-expression of multiple transgenes in mouse CNS: a comparison of strategies. *Biomol Eng*. 2001;17(6):157-165.
9. Jankowsky JL, Melnikova T, Fadale DJ, et al. Environmental Enrichment Mitigates Cognitive Deficits in a Mouse Model of Alzheimer's Disease. *J Neurosci*. 2005;25(21):5217-5224. doi:10.1523/JNEUROSCI.5080-04.2005
10. Webster SJ, Bachstetter AD, Nelson PT, Schmitt FA, Van Eldik LJ. Using mice to model Alzheimer's dementia: an overview of the clinical disease and the preclinical behavioral changes in 10 mouse models. *Front Genet*. 2014;5:88. doi:10.3389/fgene.2014.00088
11. Shepherd A, Tyebji S, Hannan AJ, Burrows EL. Translational Assays for Assessment of Cognition in Rodent Models of Alzheimer's Disease and

- Dementia. *J Mol Neurosci*. 2016;60(3):371-382. doi:10.1007/s12031-016-0837-1
12. Perry RJ, Hodges JR. Attention and executive deficits in Alzheimer's disease. *Brain*. 1999;122(3):383-404. doi:10.1093/brain/122.3.383
 13. Fortenbaugh FC, DeGutis J, Esterman M. Recent theoretical, neural, and clinical advances in sustained attention research. *Ann N Y Acad Sci*. 2017;1396(1):70-91. doi:10.1111/nyas.13318
 14. Baddeley AD, Baddeley HA, Bucks RS, Wilcock GK. Attentional control in Alzheimer's disease. *Brain J Neurol*. 2001;124(Pt 8):1492-1508. doi:doi:10.1093/brain/124.8.1492
 15. Romberg C, Bussey TJ, Saksida LM. Paying more attention to attention: Towards more comprehensive cognitive translation using mouse models of Alzheimer's disease. *Brain Res Bull*. 2013;92:49-55. doi:10.1016/j.brainresbull.2012.02.007
 16. Berardi AM, Parasuraman R, Haxby JV. Sustained attention in mild Alzheimer's disease. *Dev Neuropsychol*. 2005;28(1):507-537. doi:10.1207/s15326942dn2801_4
 17. Levinoff EJ, Saumier D, Chertkow H. Focused attention deficits in patients with Alzheimer's disease and mild cognitive impairment. *Brain Cogn*. 2005;57(2):127-130. doi:10.1016/j.bandc.2004.08.058
 18. Huntley JD, Hampshire A, Bor D, Owen AM, Howard RJ. The importance of sustained attention in early Alzheimer's disease. *Int J Geriatr Psychiatry*. 2017;32(8):860-867. doi:10.1002/gps.4537
 19. Rizzo M, Anderson SW, Dawson J, Myers R, Ball K. Visual attention impairments in Alzheimer's disease. *Neurology*. 2000;54(10):1954-1959. doi:10.1212/WNL.54.10.1954
 20. Saunders NLJ, Summers MJ. Longitudinal deficits to attention, executive, and working memory in subtypes of mild cognitive impairment. *Neuropsychology*. 2011;25(2):237-248. doi:10.1037/a0021134
 21. Sahakian BJ, Coull JT. Tetrahydroaminoacridine (THA) in Alzheimer's disease: an assessment of attentional and mnemonic function using CANTAB. *Acta Neurol Scand Suppl*. 1993;149:29-35.
 22. Robbins TW. The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. *Psychopharmacology (Berl)*. 2002;163(3-4):362-380. doi:10.1007/s00213-002-1154-7

23. Mar AC, Horner AE, Nilsson SRO, et al. The touchscreen operant platform for assessing executive function in rats and mice. *Nat Protoc.* 2013;8(10):1985-2005. doi:10.1038/nprot.2013.123
24. Romberg C, Horner AE, Bussey TJ, Saksida LM. A touch screen-automated cognitive test battery reveals impaired attention, memory abnormalities, and increased response inhibition in the TgCRND8 mouse model of Alzheimer's disease. *Neurobiol Aging.* 2013;34(3):731-744. doi:10.1016/j.neurobiolaging.2012.08.006
25. Kent BA, Heath CJ, Kim CH, et al. Longitudinal evaluation of Tau - P301L transgenic mice reveals no cognitive impairments at 17 months of age. *Brain Behav.* 2017;8(1). doi:10.1002/brb3.896
26. Bharmal AV, Kent BA, Bussey TJ, Saksida LM. Performance of transgenic TgTau-P301L mice in a 5-choice serial reaction time task (5-CSRTT) as a model of Alzheimer's disease. *Psychiatr Danub.* 2015;27 Suppl 1:S515-525.
27. Romberg C, Mattson MP, Mughal MR, Bussey TJ, Saksida LM. Impaired Attention in the 3xTgAD Mouse Model of Alzheimer's Disease: Rescue by Donepezil (Aricept). *J Neurosci.* 2011;31(9):3500-3507. doi:10.1523/JNEUROSCI.5242-10.2011
28. Filali M, Lalonde R. Age-related cognitive decline and nesting behavior in an APPSwe/PS1 bigenic model of Alzheimer's disease. *Brain Res.* 2009;1292:93-99. doi:10.1016/j.brainres.2009.07.066
29. Janus C, Flores AY, Xu G, Borchelt DR. Behavioral abnormalities in APPSwe/PS1dE9 mouse model of AD-like pathology: comparative analysis across multiple behavioral domains. *Neurobiol Aging.* 2015;36(9):2519-2532. doi:10.1016/j.neurobiolaging.2015.05.010
30. Jiao S-S, Yao X-Q, Liu Y-H, et al. Edaravone alleviates Alzheimer's disease-type pathologies and cognitive deficits. *Proc Natl Acad Sci.* 2015;112(16):5225-5230. doi:10.1073/pnas.1422998112
31. Olesen LØ, Bouzinova EV, Severino M, et al. Behavioural Phenotyping of APPSwe/PS1dE9 Mice: Age-Related Changes and Effect of Long-Term Paroxetine Treatment. *PLOS ONE.* 2016;11(11):e0165144. doi:10.1371/journal.pone.0165144
32. Reiserer RS, Harrison FE, Syverud DC, McDonald MP. Impaired spatial learning in the APPSwe + PSEN1ΔE9 bigenic mouse model of Alzheimer's disease. *Genes Brain Behav.* 2007;6(1):54-65. doi:10.1111/j.1601-183X.2006.00221.x

33. Stover KR, Brown RE. Age-related changes in visual acuity, learning and memory in the APP^{swe}/PS1^{dE9} mouse model of Alzheimer's disease. *Behav Brain Res.* 2012;231(1):75-85. doi:10.1016/j.bbr.2012.02.044
34. Filali M, Lalonde R. The effects of subchronic d-serine on left–right discrimination learning, social interaction, and exploratory activity in APP^{swe}/PS1 mice. *Eur J Pharmacol.* 2013;701(1–3):152-158. doi:10.1016/j.ejphar.2012.12.018
35. Cao D, Lu H, Lewis TL, Li L. Intake of sucrose-sweetened water induces insulin resistance and exacerbates memory deficits and amyloidosis in a transgenic mouse model of Alzheimer disease. *J Biol Chem.* 2007;282(50):36275–82. doi:10.1074/jbc.M703561200
36. Hooijmans CR, Van der Zee CEEM, Dederen PJ, et al. DHA and cholesterol containing diets influence Alzheimer-like pathology, cognition and cerebral vasculature in APP^{swe}/PS1^{dE9} mice. *Neurobiol Dis.* 2009;33(3):482-498. doi:10.1016/j.nbd.2008.12.002
37. Stewart S, Cacucci F, Lever C. Which memory task for my mouse? A systematic review of spatial memory performance in the Tg2576 Alzheimer's mouse model. *J Alzheimers Dis JAD.* 2011;26(1):105-126. doi:10.3233/JAD-2011-101827
38. Crabbe JC, Wahlsten D, Dudek BC. Genetics of mouse behavior: interactions with laboratory environment. *Science.* 1999;284(5420):1670-1672.
39. Chesler EJ, Wilson SG, Lariviere WR, Rodriguez-Zas SL, Mogil JS. Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neurosci Biobehav Rev.* 2002;26(8):907-923.
40. Sorge RE, Martin LJ, Isbester KA, et al. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat Methods.* 2014;11(6):629-632. doi:10.1038/nmeth.2935
41. Jankowsky JL, Zheng H. Practical considerations for choosing a mouse model of Alzheimer's disease. *Mol Neurodegener.* 2017;12(1):89. doi:10.1186/s13024-017-0231-7
42. Rae EA, Brown RE. The problem of genotype and sex differences in life expectancy in transgenic AD mice. *Neurosci Biobehav Rev.* 2015;57:238-251. doi:10.1016/j.neubiorev.2015.09.002
43. Shansky RM. Are hormones a “female problem” for animal research? *Science.* 2019;364(6443):825-826. doi:10.1126/science.aaw7570

44. Burrows EL, Hannan AJ. Towards environmental construct validity in animal models of CNS disorders: optimizing translation of preclinical studies. *CNS Neurol Disord Drug Targets*. 2013;12(5):587-592. doi:10.2174/1871527311312050007
45. Maggi S, Garbugino L, Heise I, et al. A Cross-Laboratory Investigation of Timing Endophenotypes in Mouse Behavior. *Timing Amp Time Percept*. 2014;2(1):35-50. doi:10.1163/22134468-00002007
46. Dauchy RT, Dauchy EM, Tirrell RP, et al. Dark-phase light contamination disrupts circadian rhythms in plasma measures of endocrine physiology and metabolism in rats. *Comp Med*. 2010;60(5):348-356.
47. Roedel A, Storch C, Holsboer F, Ohl F. Effects of light or dark phase testing on behavioural and cognitive performance in DBA mice. *Lab Anim*. 2006;40(4):371-381. doi:10.1258/002367706778476343
48. Kim EW, Phillips BU, Heath CJ, et al. Optimizing reproducibility of operant testing through reinforcer standardization: identification of key nutritional constituents determining reward strength in touchscreens. *Mol Brain*. 2017;10(1):31. doi:10.1186/s13041-017-0312-0
49. Ward RD, Winiger V, Higa KK, et al. The impact of motivation on cognitive performance in an animal model of the negative and cognitive symptoms of schizophrenia. *Behav Neurosci*. 2015;129(3):292-299. doi:10.1037/bne0000051
50. Stănişor L, Togh C van der, Pennartz CMA, Roelfsema PR. A unified selection signal for attention and reward in primary visual cortex. *Proc Natl Acad Sci*. 2013;110(22):9136-9141. doi:10.1073/pnas.1300117110
51. Hebda-Bauer EK, Simmons TA, Sugg A, et al. 3xTg-AD Mice Exhibit an Activated Central Stress Axis during Early-Stage Pathology. *J Alzheimers Dis JAD*. 2013;33(2):407-422. doi:10.3233/JAD-2012-121438
52. Rothman SM, Herdener N, Camandola S, et al. 3xTgAD mice exhibit altered behavior and elevated A β after chronic mild social stress. *Neurobiol Aging*. 2012;33(4):830.e1-830.12. doi:10.1016/j.neurobiolaging.2011.07.005
53. Sterniczuk R, Antle MC, Laferla FM, Dyck RH. Characterization of the 3xTg-AD mouse model of Alzheimer's disease: part 2. Behavioral and cognitive changes. *Brain Res*. 2010;1348:149-155. doi:10.1016/j.brainres.2010.06.011
54. Zhang Y-L, Xing R-Z, Luo X-B, et al. Anxiety-like behavior and dysregulation of miR-34a in triple transgenic mice of Alzheimer's disease. *Eur Rev Med Pharmacol Sci*. 2016;20(13):2853-2862.

55. Sterniczuk R, Dyck RH, Laferla FM, Antle MC. Characterization of the 3xTg-AD mouse model of Alzheimer's disease: part 1. Circadian changes. *Brain Res.* 2010;1348:139-148. doi:10.1016/j.brainres.2010.05.013
56. Ambrée O, Touma C, Görtz N, et al. Activity changes and marked stereotypic behavior precede Aβ pathology in TgCRND8 Alzheimer mice. *Neurobiol Aging.* 2006;27(7):955-964. doi:10.1016/j.neurobiolaging.2005.05.009
57. Shepherd A, Zhang TD, Zeleznikow-Johnston AM, Hannan AJ, Burrows EL. Transgenic Mouse Models as Tools for Understanding How Increased Cognitive and Physical Stimulation Can Improve Cognition in Alzheimer's Disease. *Brain Plast.* 2018;4(1):127-150. doi:10.3233/BPL-180076
58. Cheng D, Logge W, Low JK, Garner B, Karl T. Novel behavioural characteristics of the APP(Swe)/PS1ΔE9 transgenic mouse model of Alzheimer's disease. *Behav Brain Res.* 2013;245:120-127. doi:10.1016/j.bbr.2013.02.008
59. Lalonde R, Kim HD, Fukuchi K. Exploratory activity, anxiety, and motor coordination in bigenic APP^{swe} + PS1/DeltaE9 mice. *Neurosci Lett.* 2004;369(2):156-161. doi:10.1016/j.neulet.2004.07.069
60. Lalonde R, Kim HD, Maxwell JA, Fukuchi K. Exploratory activity and spatial learning in 12-month-old APP(695)SWE/co+PS1/DeltaE9 mice with amyloid plaques. *Neurosci Lett.* 2005;390(2):87-92. doi:10.1016/j.neulet.2005.08.028
61. Van Cauwenberghe C, Vandendriessche C, Libert C, Vandenbroucke RE. Caloric restriction: beneficial effects on brain aging and Alzheimer's disease. *Mamm Genome.* 2016;27(7):300-319. doi:10.1007/s00335-016-9647-6
62. Mouton PR, Chachich ME, Quigley C, Spangler E, Ingram DK. Caloric restriction attenuates amyloid deposition in middle-aged dtg APP/PS1 mice. *Neurosci Lett.* 2009;464(3):184-187. doi:10.1016/j.neulet.2009.08.038
63. Halagappa VKM, Guo Z, Pearson M, et al. Intermittent fasting and caloric restriction ameliorate age-related behavioral deficits in the triple-transgenic mouse model of Alzheimer's disease. *Neurobiol Dis.* 2007;26(1):212-220. doi:10.1016/j.nbd.2006.12.019
64. Patel NV, Gordon MN, Connor KE, et al. Caloric restriction attenuates Aβ-deposition in Alzheimer transgenic models. *Neurobiol Aging.* 2005;26(7):995-1000. doi:10.1016/j.neurobiolaging.2004.09.014

65. Qin W, Chachich M, Lane M, et al. Calorie restriction attenuates Alzheimer's disease type brain amyloidosis in Squirrel monkeys (*Saimiri sciureus*). *J Alzheimers Dis.* 2006;10(4):417-422. doi:10.3233/JAD-2006-10411
66. Schafer MJ, Alldred MJ, Lee SH, et al. Reduction of β -amyloid and γ -secretase by calorie restriction in female Tg2576 mice. *Neurobiol Aging.* 2015;36(3):1293-1302. doi:10.1016/j.neurobiolaging.2014.10.043
67. Wang J, Ho L, Qin W, et al. Caloric restriction attenuates β -amyloid neuropathology in a mouse model of Alzheimer's disease. *FASEB J.* 2005;19(6):659-661. doi:10.1096/fj.04-3182fje
68. Bahar-Fuchs A, Webb S, Bartsch L, et al. Tailored and Adaptive Computerized Cognitive Training in Older Adults at Risk for Dementia: A Randomized Controlled Trial. *J Alzheimers Dis.* 2017;60(3):889-911. doi:10.3233/JAD-170404
69. Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized Cognitive Training in Older Adults With Mild Cognitive Impairment or Dementia: A Systematic Review and Meta-Analysis. *Am J Psychiatry.* 2017;174(4):329-340. doi:10.1176/appi.ajp.2016.16030360
70. Masuda A, Kobayashi Y, Kogo N, Saito T, Saido TC, Itohara S. Cognitive deficits in single App knock-in mouse models. *Neurobiol Learn Mem.* 2016;135(Supplement C):73-82. doi:10.1016/j.nlm.2016.07.001
71. Rochat L, Billieux J, Juillerat Van der Linden A-C, et al. A multidimensional approach to impulsivity changes in mild Alzheimer's disease and control participants: cognitive correlates. *Cortex J Devoted Study Nerv Syst Behav.* 2013;49(1):90-100. doi:10.1016/j.cortex.2011.08.004
72. Nyatsanza S, Shetty T, Gregory C, Lough S, Dawson K, Hodges JR. A study of stereotypic behaviours in Alzheimer's disease and frontal and temporal variant frontotemporal dementia. *J Neurol Neurosurg Psychiatry.* 2003;74(10):1398-1402. doi:10.1136/jnnp.74.10.1398

Figure 1. Schematic of the 5 Choice Serial Reaction Time Task (5CSRTT) and timeline.

(a) An illustration of the 5CSRTT task. First the animal initiated a trial (1) via a nose-poke into the food magazine, followed by a delay (2) and then the stimulus appeared in one of 5 locations (3). The animal then responds to the correct grid position (4a) and receives a reward (5a), or responds prematurely (during the delay) (2a), to the incorrect position (4b) or not at all (i.e. an omission, 4c), all of which caused the house light to turn on and a 5s time out was initiated (5b). (b) Timeline of

the study. Animals were food restricted then began pre-training at 6 months of age. Following this, animals began 5CSRTT training, which took 3-10 weeks. At 9 months of age, animals began the series of probes (first stimulus duration, then delay, then stimulus brightness followed by a repeat of the stimulus duration probe) that took 3 weeks each. Finally, animals were trained to touch a stimulus multiple times for one reward then performed the progressive ratio 4 (PR) task for 8 days. Animals were 12 months at the end of the PR task.

Figure 2. WT and APP/PS1 mice do not differ during the final stage of 5CSRTT training.

WT and APP/PS1 mice show similar, accuracy, response rate, rate of premature responses and rate of perseverative responses as well as similar correct response and reward collection latencies during the final 2s stage of 5CSRTT training. (a) Accuracy excluding omitted trials. (b) Response rate. (c) Number of premature responses per trial. (d) Number of perseverative responses per trial. (e) Latency to make a correct response. (f) Reward collection latency (g) Odds ratio for correct selection, response and premature responses, and incidence rate ratio for perseverative responses during 2s training, showing no effect of genotype. (a-f) are presented as mean \pm SEM with individual animals as dots. (g) presents odds ratio/incidence rate ratio \pm 95% CI, showing the effect of genotype on correct selection, response, premature responses or perseverative responses. (h) presents median coefficients in seconds with 95% CI, showing the effect of genotype on correct response latency and reward collection latency. * = $p < 0.05$; *** = $p < 0.001$

Figure 3. 9-month-old APP/PS1 mice show slightly increased responding on a stimulus duration probe compared to WT littermate controls.

Reducing stimulus duration does not significantly alter correct, premature or perseverative responses across genotype, but APP/PS1 mice do show increased odds to respond to a trial compared to WT. (a) Accuracy excluding omitted trials. (b) Response rate. (c) Number of premature responses per trial. (d) Number of perseverative responses per trial. (e) The effect of genotype on the odds ratio over the stimulus duration probe, with the APP/PS1 mice showing an increased odds of response. (f) The effect of reduced stimulus duration on the odds ratio for correct selection, with the odds of a correct selection, a response and a premature response being significantly decreased, while the incidence of perseverative responses significantly increases. (a-d) are presented as mean \pm SEM while (e-f) represent odds ratios/incidence rate ratios \pm 95% CI, showing the effect of genotype or probe respectively on correct selection, response, premature responses or perseverative responses. * = $p < 0.05$; *** = $p < 0.001$

Figure 4. WT and APP/PS1 mice perform equally well at short and long delays.

Performance measures on a delay probe at 10 months of age do not differ due to genotype. (a) Accuracy excluding omitted trials. (b) Response rate. (c) Number of premature trials per trial. (d) Number of perseverative trials per trial. (e) The effect of reduced delay, showing decreased odds of responses and premature responses, and increased odds of correct responses. (f) The effect of increased delay, showing decreased correct response and increased premature and perseverative responses. (g) The effect of genotype over the delay probe, showing no differences. (a-d) are presented as mean \pm SEM while (e-g) represent odds ratios/incidence rate ratios \pm 95% CI, showing the effect of probes or genotype on correct selection, response, premature responses or perseverative responses. * = $p < 0.05$; *** = $p < 0.001$

Figure 5. APP/PS1 mice make more premature and perseverative responses when stimuli brightness is reduced compared to WT animals.

Reducing stimuli brightness significantly reduced accuracy and response rate equally between WT and APP/PS1 mice, but APP/PS1 mice show elevated premature and perseverative responses compared to WTs. (a) Accuracy excluding omitted trials. (b) Response rate. (c) Number of premature trials per trial. (d) Number of perseverative trials per trial. (e) The effect of genotype over reduced stimulus brightness, showing increased premature and perseverative responses. (f) The effect of decreased stimulus brightness, showing decreased odds of correct selection or response and an increased expected count of perseverative responses. (a-d) are presented as mean \pm SEM while (e-f) represent odds ratios/incidence rate ratios \pm 95% CI, showing the effect of genotype or probe respectively on correct selection, response, premature responses or perseverative responses. * = $p < 0.05$; *** = $p < 0.001$

Figure 6. 11-month-old WT and APP/PS1 mice show no differences in the repeat of the stimulus duration probe.

Reducing stimulus duration for a second time does not significantly alter accuracy, response rate, premature or perseverative responses between genotypes. (a) Accuracy excluding omitted trials. (b) Response rate. (c) Number of premature trials per trial. (d) Number of perseverative trials per trial. (e) The effect of genotype over reduced stimulus duration, showing no changes. (f) The effect of decreased stimulus duration, showing decreased correct selection, response and increased perseverative responses. (a-d) are presented as mean \pm SEM while (e-f) represent odds ratios/incidence rate ratios \pm 95% CI, showing the effect of genotype or probe respectively on correct selection, response, premature responses or perseverative responses. * = $p < 0.05$; *** = $p < 0.001$

Supplementary Figure 1. APP/PS1 and WT mice perform equally across baseline sessions between probes.

APP/PS1 and WT mice show consistent accuracy and response rates across baseline sessions. (a) Accuracy excluding omitted trials, (b) response rate, (c) the effect of genotype over baseline sessions, showing no changes in odds of correct selection, response premature or incidence of perseverative responses. (a-b) are presented as mean \pm SEM while (c) represent odds ratios/incidence rate ratios \pm 95% CI, showing the effect of genotype on correct selection, response, premature responses or perseverative responses. * = $p < 0.05$; *** = $p < 0.001$

Supplementary Figure 2. APP/PS1 mice show intact vigilance across all four 5CSRTT probes.

Trials were split across 5 equal bins, and accuracy and response rate assessed across the 4 different probes. (a) Accuracy excluding omitted trials and (b) response rate across the first stimulus duration probe. (c) Accuracy excluding omitted trials and (d) response rate across the delay probe. (e) Accuracy excluding omitted trials and (f) response rate across the stimulus brightness probe. (g) Accuracy excluding omitted trials and (h) response rate across the second stimulus duration probe. (a-h) are presented as mean \pm SEM. * = $p < 0.05$; *** = $p < 0.001$

Supplementary Figure 3. Both WT and APP/PS1 mice show a decreased response rate, premature response rate and number of perseverative responses on the second stimulus duration probe.

Comparing performance of WT and APP/PS1 mice on the second stimulus duration probe reveals that both genotypes are less likely to make a premature or perseverative response, as well as being less likely to respond at all to a trial. (a) Accuracy excluding omitted trials. (b) Response rate. (c) Number of premature trials per trial. (d) Number of perseverative trials per trial. (e) The effect of repeating the stimulus duration probe, showing a decreased odds of response, premature response or perseverations. (f) The effect of a genotype-probe interaction variable, showing no effect. (a-d) are presented as mean \pm SEM with the second time point in a dashed line. (e-f) represent odds ratios/incidence rate ratios \pm 95% CI, showing the effect of genotype or probe respectively on correct selection, response, premature responses or perseverative responses. * = $p < 0.05$; *** = $p < 0.001$

Supplementary Figure 4 APP/PS1 mice show a slightly higher breakpoint on progressive ratio than their WT counterparts.

APP/PS1 mice show a higher overall breakpoint and touch number than their WT counterparts, but this is not significant over the week of testing. (a) Overall

breakpoint. (b) Breakpoint split over the 8 days of PR. (c) number of touches made. (d) Reward collection latency. a-c are presented as mean \pm SEM while d shows median \pm 95% CI. * = $p < 0.05$; *** = $p < 0.001$

Stage/Probe	IRR	Standard Error	z	p value	95% CI Lower	95% CI Upper
Training	0.899	0.134	-0.71	0.477	0.671	1.205
Stimulus duration 1	0.864	0.09	-1.4	0.161	0.705	1.06
Delay	0.907	0.172	-0.51	0.609	0.626	1.316
Brightness	0.882	0.083	-1.34	0.18	0.733	1.06
Stimulus duration 2	0.949	0.087	-0.57	0.567	0.794	1.135

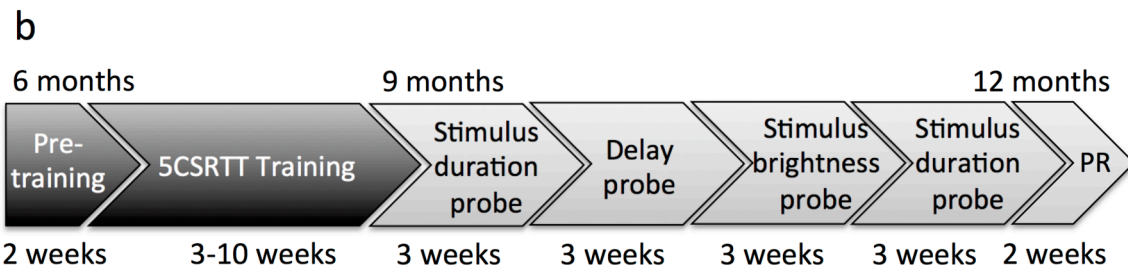
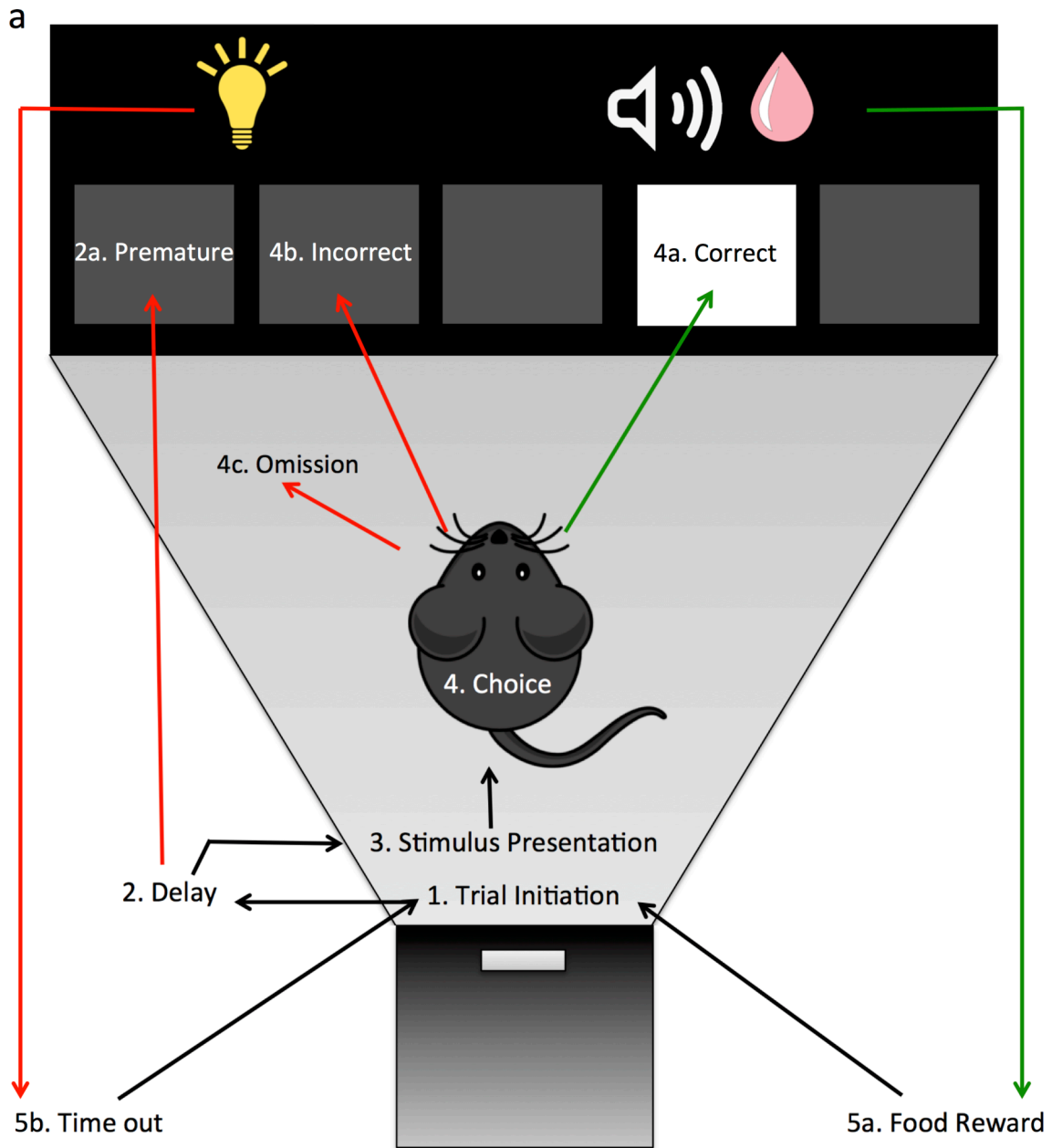
Supplementary Table 1. Effect of trial type and genotype interaction variable on perseverative responses

A genotype and trial type (correct, incorrect, premature) interaction variable was generated and run through a GLLAMM poisson regression model to test the effect of this variable on perseverative responses. The trial type and genotype interaction variable were never significant, indicating the both genotypes had similar perseverative responding patterns following different trial types across all both the final stage of training and probes. IRR= Incidence Rate Ratio

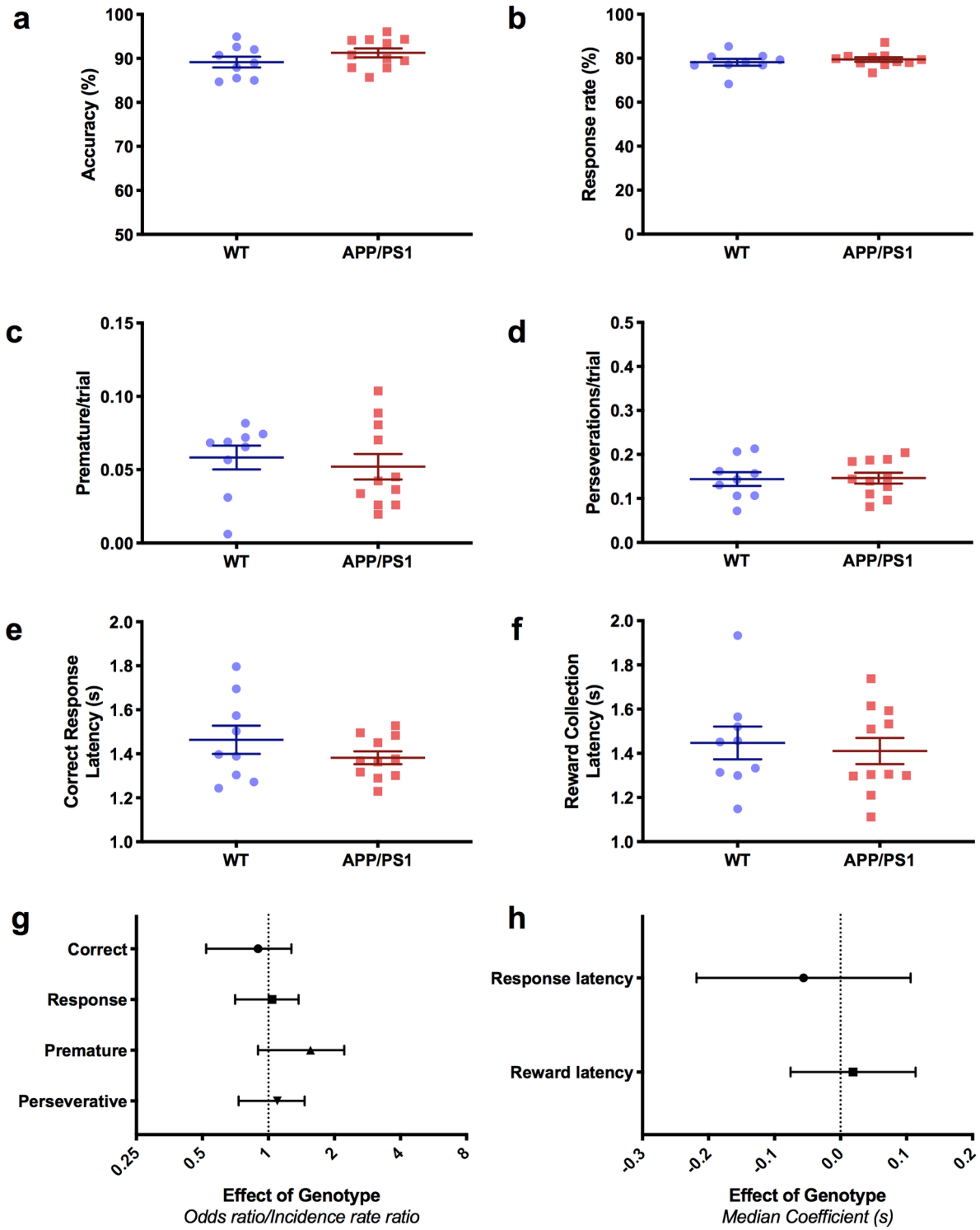
Stage	Beam location	IRR	Standard Error	z	p value	95% CI Lower	95% CI Upper
Training	Front	0.8	0.56	-0.32	0.75	0.2	3.13
	Back	0.95	0.18	-0.29	0.77	0.66	1.37
Stimulus duration 1	Front	0.89	0.78	-0.14	0.89	0.16	4.99
	Back	0.91	0.34	-0.26	0.79	0.44	1.88
Delay	Front	1.03	0.8	0.04	0.97	0.23	4.73
	Back	1	0.48	0	1	0.39	2.58
Brightness	Front	1.09	0.72	0.13	0.9	0.3	4
	Back	1.15	0.56	0.29	0.77	0.44	3
Stimulus duration 2	Front	1.03	0.65	0.05	0.96	0.3	3.56
	Back	1.17	0.55	0.33	0.74	0.46	2.93
Baseline sessions	Front	0.98	0.71	-0.02	0.98	0.24	4.05
	Back	1.01	0.37	0.04	0.97	0.5	2.06

Supplementary Table 2. Genotype has no effect on the number beam breaks at any stage of the 5CSRTT

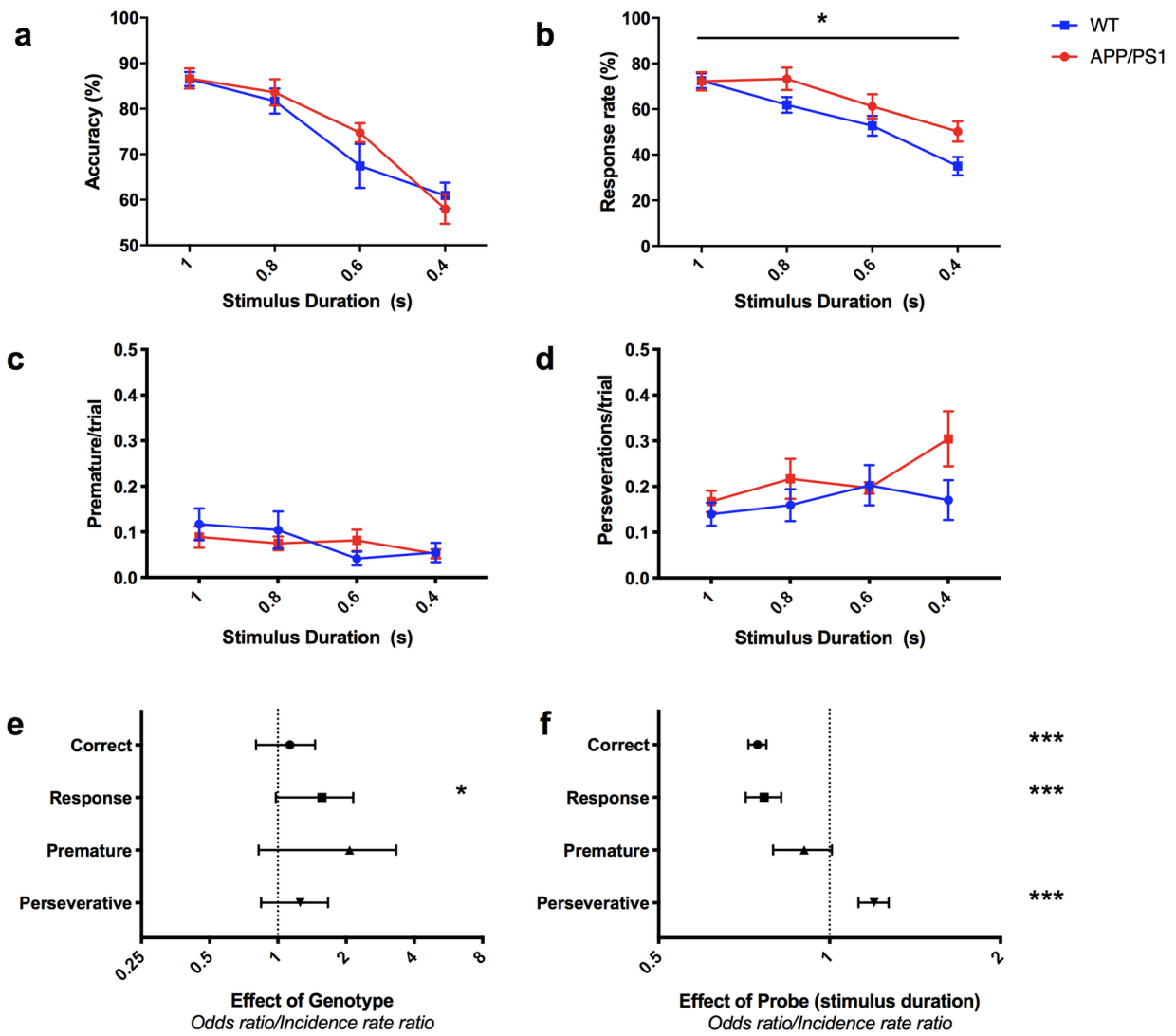
APP/PS1 and WT mice appear to move throughout the touchscreen chambers approximately the same amount at every stage of the 5CSRTT, as measured by the number of beam breaks at the front and back of the chamber. IRR= Incidence Rate Ratio



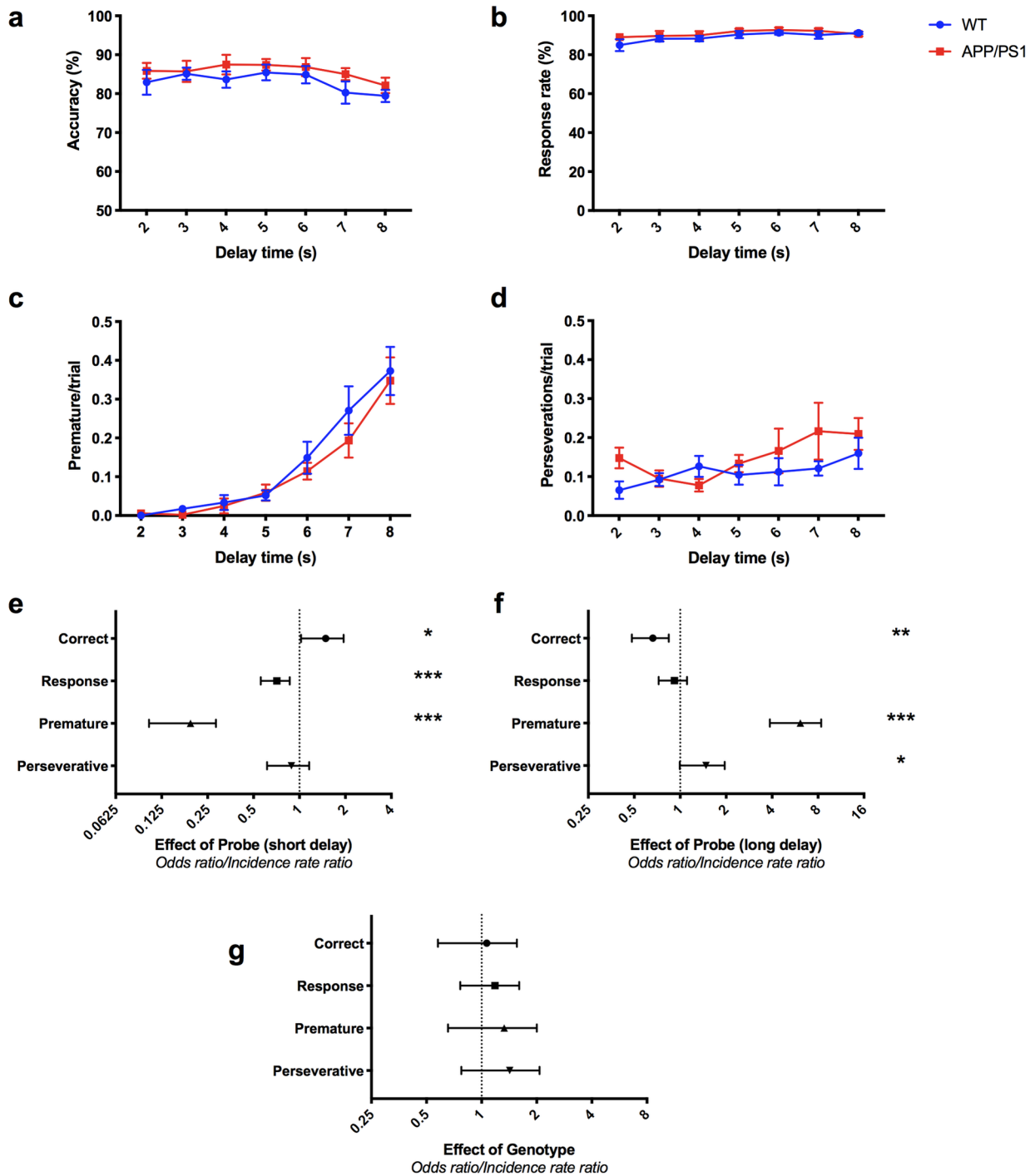
GBB_12594_Figure 1.tiff



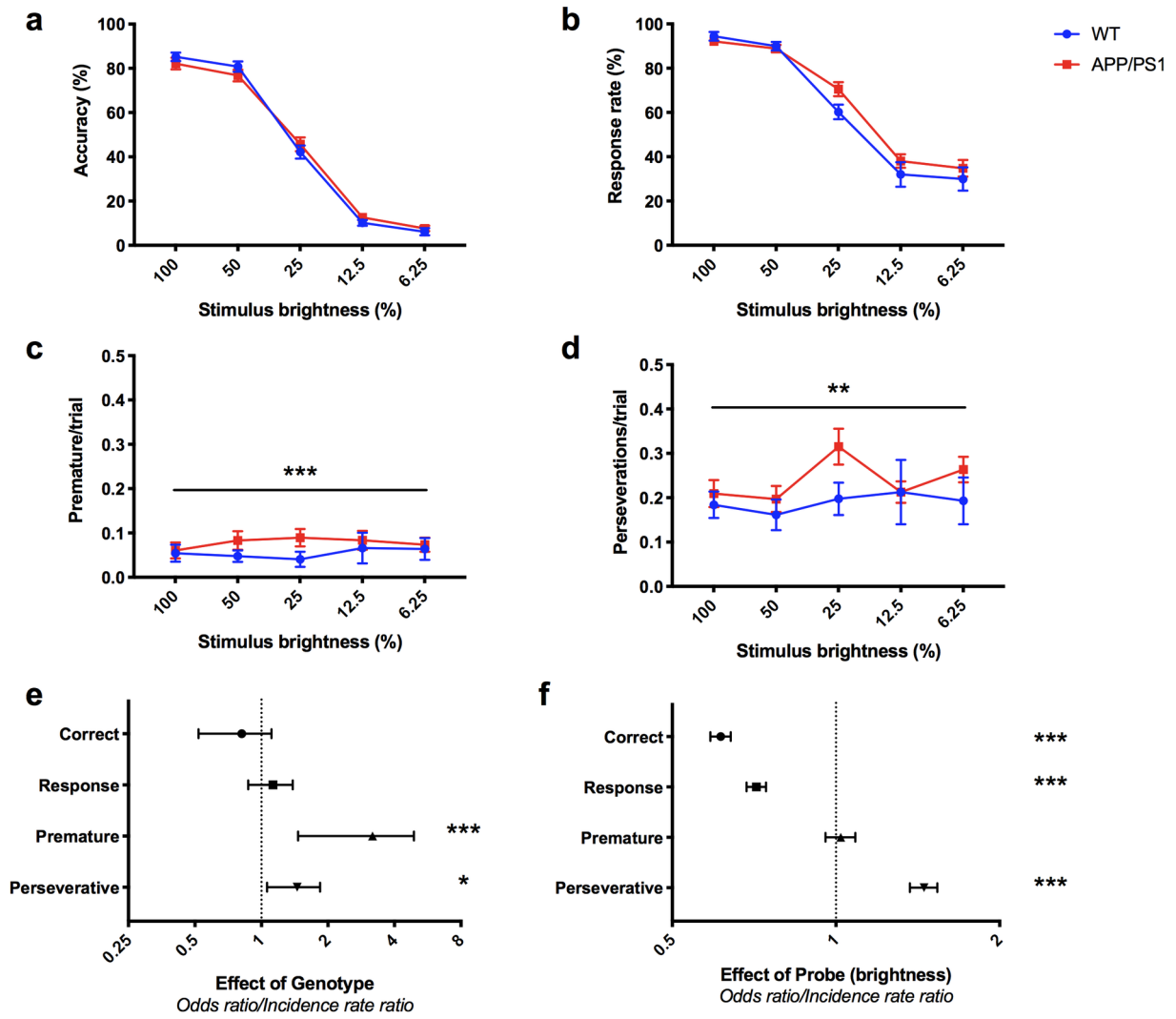
GBB_12594_Figure 2(1).tiff



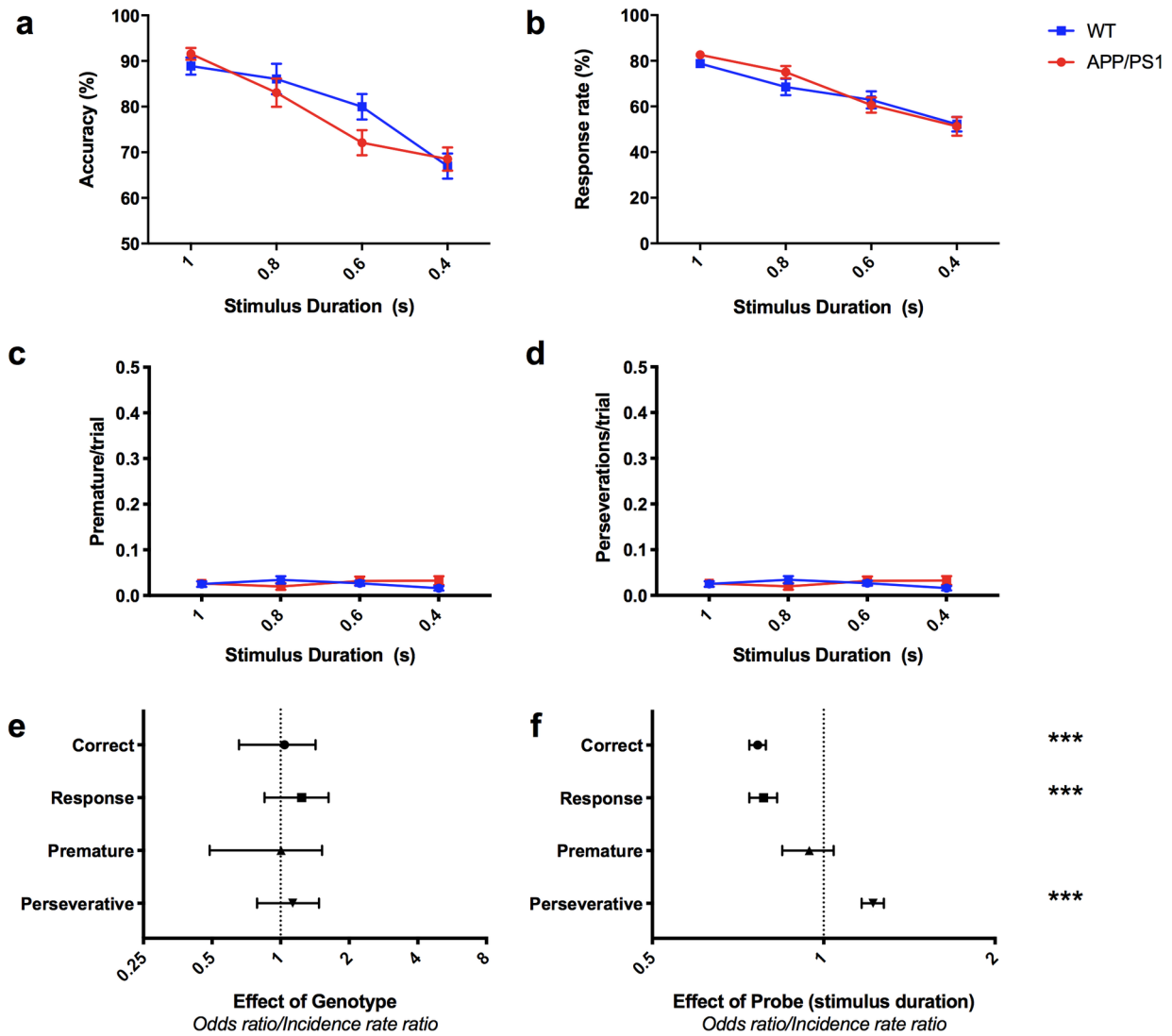
GBB_12594_Figure 3(1).tiff



GBB_12594_Figure 4(1).tiff



GBB_12594_Figure 5(1).tiff



GBB_12594_Figure 6(1).tiff