Assessing Cognitive Flexibility in Mice Using a Custom-Built Touchscreen Chamber

Rui C. Pais[1*], Ali Goldani[1], Jayden Hutchison[1], Amirhossein Mazrouei[1], Mostafa Khavaninzadeh[1], Leonardo A. Molina[3], Bruce L. McNaughton[1,4], Robert J. Sutherland[1], Majid H. Mohajerani[1,2*]

[1]Canadian Centre for Behavioural Neuroscience, Department of Neuroscience, University of Lethbridge, Lethbridge, AB, Canada
[2]Department of Psychiatry, Douglas Hospital Research Centre, McGill University, Montréal, QC, Canada
[3]Hotchkiss Brain Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[4]Center for the Neurobiology of Learning and Memory, University of California Irvine, Irvine, California, USA

*** Correspondence:**
Rui C. Pais
r.pais@uleth.ca
Majid H. Mohajerani
majid.mohajerani@mcgill.ca

Keywords: touchscreen chamber, raspberry pi, serial reversal learning, cognitive flexibility

Abstract

Automated touchscreen systems have become increasingly prevalent in rodent model screening. This technology has significantly enhanced cognitive and behavioral assessments in mice and has bridged the translational gap between basic research using rodent models and human clinical research. Our study introduces a custom-built touchscreen operant conditioning chamber powered by a Raspberry Pi and a commercially available computer tablet, which effectively addresses the significant cost barriers traditionally associated with this technology. In order to test our prototype, we decided to train C57BL/6 mice on a visual discrimination serial reversal task, and both C57BL/6 and App^{NL-G-F} mice on a new location discrimination serial reversal task. The results demonstrated a clear progression towards asymptotic performance, particularly in the location discrimination task, which also revealed potential genotype-specific deficits, with App^{NL-G-F} mice displaying an increase in the average number of errors in the first reversal as well as in perseverative errors, compared to wild-type mice. These results validate the practical utility of our touchscreen apparatus and underline its potential to provide insights into the behavioral and cognitive markers of neurobiological disorders.

1. Introduction

The evolution of behavioral tasks in Neuroscience, from traditional mazes to touchscreen paradigms, has yielded profound insights about the dynamic interplay between brain and behavior.

The development and refinement of rodent touchscreen chambers, as well as the wide variety of tasks developed for this platform over the years, has been remarkably successful in evaluating different cognitive skills in both wild-type and genetically modified rodent strains, as well as in the ability to investigate potential behavioral and neurophysiological changes resulting from pharmacological interventions (Timothy J Bussey, Muir, and Robbins 1994; T J Bussey, Saksida, and Rothblat 2001; Timothy J. Bussey, Everitt, and Robbins 1997; S. R. O. Nilsson, Saksida, and Bussey 2016; Hvoslef-Eide et al. 2016; Talpos et al. 2009; Mar et al. 2013; Horner et al. 2013).

Ever since Skinner's groundbreaking work in the context of reflexive physiology introduced automated training in the 1930's, researchers have uncovered a plethora of tools for understanding learning processes (Staddon and Cerutti 2003; B, F Skinner 1986; B. F. Skinner 1937). The development of operant conditioning boxes allowed for the precise manipulation of contextual contingencies and the measurement of behavior over a specified period of time, and significantly reduced the interaction between the experimenters and the animal subjects (Sakagami and Lattal 2016; Ferster 1953; Mar et al. 2013; Weiss 1972; Staddon and Cerutti 2003; Wetzel 1986; Pinkston 2022). By using levers or buttons the animals can press or peck in order to obtain a reinforcement (e.g. water, food pellets among others), the involvement of the experimenter during training is minimized, in favour of an auto-shaping process whereby the animals can learn the desired behaviours independently.

These operant conditioning apparatuses continued to evolve, and over time researchers started to incorporate computer screens where different images were displayed, and eventually touchscreen systems, which allowed the animals to directly interact with the displayed images in order to make a choice. The touchscreen chambers, which were initially developed to be used with pigeons, as well as human and non-human primates, were eventually adapted for rodents in the mid-nineties, and have become an invaluable tool in cognitive and behavioral neuroscience research since then (Markham, Butt, and Dougher 1996; Mar et al. 2013; Sakagami and Lattal 2016; Timothy J Bussey, Muir, and Robbins 1994; T J Bussey, Saksida, and Rothblat 2001; Winters, Saksida, and Bussey 2008; S. R. O. Nilsson, Saksida, and Bussey 2016; Timothy J. Bussey et al. 2008; Alicia Izquierdo et al. 2006; Sullivan 2022; Phillips et al. 2017; Nithianantharajah, Mckechanie, et al. 2015; Wright et al. 1988). In comparison to more traditional approaches to rodent phenotyping methods, which require multiple tests in different environments such as open-fields, mazes or conventional operant conditioning boxes, the touchscreen technology offers a controlled setting that closely mimics human cognitive assessment. This allows not only for more accurate data collection, but also for a significantly less stressful experience for the animals (Dumont, Salewski, and Beraldo 2021; O'Leary et al. 2018; Sullivan 2022).

Over the years, researchers have developed multiple tasks that cover a wide range of cognitive functions, such as visual discrimination, object-location paired-associations, visual-category learning, working memory, rule-switching, or pattern separation tasks (Wang et al. 2022; Hvoslef-Eide et al. 2015; 2016; M. Kim et al. 2016; C. H. Kim et al. 2015; Kwak, Lim, and Kaang 2016; 2015; Creighton et al. 2019; Barnard et al. 2021). In addition, the touchscreen

chamber enables high throughput testing by allowing multiple animals to be tested simultaneously; effectively streamlining the efficiency of the experimental procedures and allowing experiments to be conducted as required. With its high degree of automation, similarities to human-based cognitive assessments, and the standardization of touchscreen tasks, this behavioral apparatus has enhanced the translatability of preclinical models, leading to its widespread adoption across multiple research institutions. These include universities, biotechnological firms, and pharmaceutical companies, particularly as mice have become the preferred model organism in basic and preclinical research, due to the widespread availability of transgenic lines and the continuous refinement of genetic and molecular tools that enable in-vivo recordings and circuit labeling (Dumont, Salewski, and Beraldo 2021; Hvoslef-Eide et al. 2016; Horner et al. 2013; Dickson et al. 2013).

Among the different applications of this technology, reversal learning tasks have emerged as an important tool for assessing cognitive flexibility. These tasks require multiple executive functions such as attention, working memory or response inhibition, and depend on the subjects' adaptability to changing rewards or feedback (Van den Broeck et al. 2019; Cools et al. 2002; Dickson et al. 2013; Fowler et al. 1980; Bryce and Howland 2015; Marquardt, Sigdel, and Brigman 2017; Odland, Sandahl, and Andreasen 2021; A. Izquierdo et al. 2017). Serial reversal paradigms further test the ability to learn, maintain, and then re-learn behavioral rules over multiple iterations, as each change requires the suppression of previously reinforced behaviors and the subsequent adaptation to new rules, thus engaging executive functions such as inhibitory control, cognitive flexibility and attentional processes to an even greater extent. (Dickson et al. 2013; Kosaki and Watanabe 2012; Boulougouris, Dalley, and Robbins 2007; Castañé Anna, Theobald, and Robbins 2010; A. Izquierdo et al. 2017).

Reversal learning studies were among the first to adopt touchscreen technology for both human and non-human primates, whereas rodent studies typically relied on either spatial or non-visual cues - a discrepancy that stemmed from automation challenges and difficulties in standardizing experiments across species. However, touchscreen technology has bridged this gap and enabled standardized tasks that could be adapted and used across various species, while maintaining the underlying focus on adaptive responses and rule switching (Hvoslef-Eide et al. 2016; Timothy J. Bussey, Everitt, and Robbins 1997; T J Bussey, Saksida, and Rothblat 2001; S. R. O. Nilsson, Saksida, and Bussey 2016; Talpos et al. 2009; Hvoslef-Eide et al. 2015; Nithianantharajah, McKechanie, et al. 2015).

Despite its longstanding use, reversal learning remains an important behavioural paradigm, especially when it comes to identifying learning and cognitive flexibility deficits in neuropsychiatric disorders, such as schizophrenia, obsessive-compulsive disorder (OCD), depression, autism, Parkinson’s, and Alzheimer’s disease (Guarino et al. 2019; Lafleche and Albert 1995; Millan et al. 2012; Monni et al. 2023; Gruner and Pittenger 2017; D’Cruz et al. 2013; Valerius et al. 2008; Jara-Rizzo et al. 2020; Marazziti et al. 2010; Wobrock et al. 2009). Concurrently, cross-species studies have also highlighted the role of the prefrontal cortex - specifically, the orbitofrontal (OFC) and medial prefrontal (mPFC) cortices – as well as subcortical regions such as the dorsal striatum and amygdala, in facilitating these tasks (Clatworthy et al. 2009; Chudasama and Robbins 2003; Brigman, Graybeal, and Holmes 2010; Graybeal et al. 2011; Alicia Izquierdo et al. 2006; Alsiö et al. 2015; Dias, Robbins, and Roberts 1996; Cools et al. 2002; Hampshire and Owen 2006; Hornak et al. 2004; Lucantonio, Caprioli, and Schoenbaum 2014; Alicia Izquierdo and Jentsch 2012; A. Izquierdo et al. 2017).

While the benefits of touchscreen-based tasks for assessing cognitive and behavioral skills in rodents, and more specifically mice, are clear, especially in bridging the gap between species through standardized procedures, the adoption of these technologies is not without its challenges. Despite its numerous advantages, the main concern regarding the adoption of rodent touchscreen chambers has remained relatively unchanged over the years, and that is the considerable financial investment required. The expenses associated with acquiring even a single exemplar of these touchscreen chambers can be prohibitively high, which effectively hinders an even more widespread adoption and a swifter integration into the arsenal of behavioural assessment tools in basic research. Even though this technology has become progressively less expensive, the large financial outlay has led different research groups to develop their own alternatives to circumvent this issue (O'Leary et al. 2018; Eleftheriou et al. 2023; Wiesbrock, Musall, and Kampa 2022; Pineño 2014). This is particularly notable considering the accessibility of modern touchscreens as well as the different components required for the assembly and functioning of a similar product, which allow for the development and programming of various touchscreen-based tasks tailored to specific research needs.

Driven by the evolving demands of cognitive and behavioral neuroscience for automated and adaptable experimental tools, alongside the practical challenges of high equipment costs, and the need to collect behaviorally relevant data on both wild-type and Alzheimer's disease mouse models, we set out to develop a custom touchscreen apparatus for mice. To validate this approach, we designed and implemented two distinct touchscreen tasks with a specific focus on cognitive flexibility: a visual discrimination serial-reversal task, and a location discrimination serial-reversal task.

Our efforts reflect a need to develop versatile and accessible technologies to advance research in rodent cognitive flexibility, and ultimately contribute to a broader comprehension of both normal and pathological brain functions.

2. Materials And Methods

2.1 Hardware

The touchscreen apparatus was designed using computer-aided design software (SOLIDWORKS 2023 SP 3.0, Dassault Systèmes) and was adapted from specifications detailed in prior studies. The inner chamber featured a trapezoidal behavioral area, or more accurately, a triangle with rounded corners, optimized to focus on both the touchscreen and the reward area. Specific dimensions were 80 mm wide at the reward area, 260 mm wide at the screen, and a trapezoidal length of 240 mm, with a working area height of 190 mm and wall thicknesses of 10 mm. The walls were 3D printed using black PLA to minimize external light interference and enhance visual contrast during experiments. The lid and floor of the chamber were constructed from 6.5 mm thick black plexiglass to facilitate cleaning and maintain durability.

For the touchscreen interface we selected a Samsung Galaxy Tab A 8.0 SM-T350, with a resolution of 1024x768 pixels, mounted horizontally opposite the reward area and accessed through a 163 mm x 125 mm aperture. This tablet not only recorded touch interactions but also managed the experimental flow, communicating with a Raspberry Pi (RPi) 4 Model B (8 GB RAM). The Raspberry Pi was enclosed in a custom 3D-printed case attached to the touchscreen wall, designed with apertures for cable management and component interconnection.

Reward delivery was managed using a 5V solenoid valve connected to medical-grade silicone tubing (HelixMark Standard Silicone Tubing, Freudenberg Medical), which extended to a metal tube. This tube, protruding 10 mm from the wall, was 3 mm in diameter and dispensed approximately 2.5 µl of 10% sucrose water. The sucrose solution was stored in a 60 ml syringe, functioning as the reservoir for the system. The availability of the reward was signaled by a blue LED visible through a 3 mm round aperture, positioned 10 mm above the reward tube, and auditory cues that varied by the type of response were emitted through the tablet's speakers.

Videos were recorded by a small camera (Raspberry Pi Camera Module 2), positioned on top of the lid, to capture detailed activity within the chamber, and enhanced by an array of infrared LEDs for consistent illumination under low lighting conditions. This setup not only allowed the videos to be recorded locally on the RPi for later analysis, but also enabled the hosting of a local live stream from inside the chamber as soon as the trial software started. This annotated live stream allowed experimenters to supervise real-time activity within the chamber and address any issues that might interfere with the flow of the experiment.

2.2 Software

To give researchers the ability to create and control task parameters, we used an XML schema to define each experiment's specifications. An XML configuration file for an experiment is structured with tags that define different functions and sections of the experiment. Each function or parameter is enclosed in <tags> and may have various attributes. The general outline of a configuration file is shown in Figure 2.

There are five main functions within each configuration file for setting up the experimental environment. The *prepare* function allows experimenters to specify key parameters: 1) overall duration, which dictates that the experiment continues until either completion or the specified duration elapses; 2) background color, which defines the visual setting of the experiment; 3) number and size of sections, determining whether the active touching area is divided into two or four sections; 4) section dividers, specifying both the presence and color of dividers between sections; 5) initial reward cues, including the presence, number, and timing interval between these cues; 6) touch time-out, setting the duration before a time-out is triggered when the wrong image/3D object or side of the screen is touched; 7) image pre-loading, which minimizes the image/3D object load times during the experiment.

Within the *main* function, experimenters can specify the number of trials, setting it to a predetermined amount based on their experimental design. In the *reward* function, users can specify a text for logging in the final reports whenever the reward is triggered, adjust the frequency and duration of the tone played, and control the opening and closing durations of the solenoid valve. Similar to the *reward* function, the *time-out* function allows for the display of a time-out alert by filling the entire screen with a bright color for a specified duration. Users can also determine the sections where the correct and incorrect images appear; if not specified, experimenters can choose to randomize the location for each trial.

Finally, in the *trial* function, experimenters can define each trial's parameters. For visual discrimination tasks, they can select a single image or 3D virtual object or allow a random choice from a series of images for both rewarded (S+) and unrewarded (S-) categories. In location discrimination tasks, the settings allow for a cue to be set to static or blinking, with adjustable frequency. This configuration syntax enables experimenters to create a diverse range of touchscreen tasks tailored to their research needs.

The software deployed on the Samsung tablet is a Unity application developed with Unity Game Engine (Unity Technologies, 2024). Through Unity, we could easily develop the logic of the software and, using its tools for building Graphical User Interface (GUI), create the interface that best suits the experimenter's needs. To communicate with the RPi and to be able to control the hardware modules, we implemented a socket communication system so the tablet can send commands to the RPi through a wireless network. RPi's built-in GPIO4 and Picamera5 libraries were used for communication with the hardware. The software is developed as a state machine with main components working in their own evet loops. An overall view of the software components is shown in Figure 3.

The software running on the RPi is a python program that hosts a socket server and accepts connections from the tablet running the Unity app. Through this socket communication, commands from the tablet are sent with minimum delay to control hardware components connected to the RPi. For example, when the socket server receives the command "reward", it turns on the blue LED and opens the Solenoid Valve for a split second to deliver reward fluid. Screenshots from the Unity app can be found in Figure 4. The source code for the software part of this project can be found on our GitHub page.

2.3 Experimental Flow

The experiments performed with this software follow a general pattern. Each experiment starts with initial reward deliveries to give some satiation to the animals before the actual trials start. One can select multiple or no initial rewards. Then the program proceeds to execute the trials as defined by the user; they can be any kind of trial explainable by the options provided in XML configuration files. All the activities of the subject are recorded from this point, any interaction with trial objects that results in a feedback response, will be logged in a .CSV report file, accessible at the end of the experiment. Furthermore, the video recording will capture all the ongoing events within the experiment box and contains timestamps of the screen interactions along with their respective outcomes (time-out or rewarded), as well as trial number. The flow of the experiment can be seen more clearly in Figure 5.

2.4 Subjects

A total of 27 adult mice, bred in-house, were used in this study: 9 C57/BL6 mice (23 – 31 g, 6 – 8 months old, 3 males and 6 females) for the object reversal learning task, and 18 mice, comprising 9 C57/BL6 mice (26 – 31 g, 8 – 9 months old, 4 males and 5 females) and 9 $App^{NL\text{-}G\text{-}F}$ knock-in mice (25 – 32 g, 8 – 10 months old, 4 males and 5 females) for the location-reversal task.

All animals were housed in groups of 2 to 4 individuals, in standard mouse cages. The room temperature was maintained at 24 °C under a 12 h light/dark cycle with the lights on at 7:30 AM and free access to food and water before the beginning of the behavioural training. All procedures were in accordance with the guidelines established by the Canadian Council on Animal care and with the protocols approved by the Animal Welfare Committee of the University of Lethbridge.

Mice were water deprived throughout the duration of the behavioural training. During this period mice were given a daily ad libitum access to water for 30 minutes in their home cages

30 minutes after the last training session, and their weight was maintained to at least 85% of the baseline.

2.4.1 Alzheimer's disease mouse model

Alzheimer's Disease (AD) is the most prevalent form of dementia, and it is characterized by the progressive aggregation of amyloid-β (Aβ) and formation of neurofibrillary tangles, which lead to memory loss, cognitive impairments, and overall decline in quality of life (McAllister et al. 2020; Mehla et al. 2019; Braak and Braak 1991; Folch et al. 2018; Ettcheto et al. 2018). Central to AD pathogenesis is the spread of Aβ, resulting in neuroinflammation, plaque deposition, and tau hyperphosphorylation, which eventually causes brain atrophy (Harper and Lansbury 1997; Bloom 2014; Walker, Lynn, and Chernoff 2018).

The App^{NL-G-F} mouse model used in this study, incorporates humanized murine Aβ sequences with three specific mutations: Swedish (NL), Beyreuther/Iberian (F), and Arctic (G) (Saito et al. 2014; P. Nilsson, Saito, and Saido 2014). Unlike other App transgenic lines, the App^{NL-G-F} model avoids artifacts introduced by App overexpression by using a knock-in approach to express App at wild-type levels, thus ensuring that any observed pathologies are a direct result of pathogenic Aβ rather than App overexpression (Saito et al. 2014; Guardia-Laguarta et al. 2010; Shin et al. 2010). This mouse model expresses App with familial Alzheimer's disease-associated mutations which promote Aβ toxicity, an increase in total Aβ production, the Aβ42/ Aβ40 ratio, as well as promoting Aβ aggregation (P. Nilsson, Saito, and Saido 2014; Saito et al. 2014). In addition, this model reproduces several pathologies associated with AD including amyloid plaques, synaptic loss, and neuroinflammation - specifically microgliosis and astrocytosis around plaques - while also displaying age-associated cognitive impairments (Saito et al. 2014; Latif-Hernandez et al. 2019; Upīte et al. 2020; Mehla et al. 2023; Lacoursiere et al. 2022; Latif-Hernandez et al. 2020).

2.5 *Experimental Design*

2.5.1 Visual discriminating serial reversal task

This task is based on the classic touchscreen pairwise discrimination task described in previous studies (Mar et al. 2013; Horner et al. 2013), with some slight modifications. Briefly, in this task mice need to choose between 2 images, or virtual objects, appearing on each side of the screen, by touching the surface of the touchscreen where the virtual objects are displayed. Before the pairwise discrimination takes place, the animals need to undergo some form of pretraining, where they learn the basic rules of the task in a progressive stepwise manner. The pre-training sessions were divided into four different stages: (1) *Habituation*, in which mice are introduced to the touchscreen chamber for 10 and 30 minutes, for 2 consecutive days, followed by 2 daily sessions of 60 minutes each, where the screen is OFF and the reward is delivered in 10 second intervals; (2) *Image Presentation*, where the rewarded (S+) image is introduced and paired with a tone and the reward delivery in 10 second intervals, for a total of 60 minutes; (3) *Touchscreen Interaction*, where the animals must learn to touch the area on the screen where the object appears in order to trigger the release of the reward for a total of 30 trials or 60 minute duration; (4) *Time-Out*, where mice are introduced to a small time-out on commission of an error, if the screen is touched anywhere besides where the S+ image appears, with the passing criteria defined as 80% correct responses or 24 out of 30 trials for 2 consecutive sessions. Finally, in the

Acquisition stage, the S- image is introduced, and mice must make a choice between the S+ and S- images which can appear on either the left or right side of the screen in a pseudo-random manner. After completing this stage, the reward contingencies are then reversed, and the S+ becomes the new S- and vice-versa. This cycle is then repeated 5 times, with an upper limit of 60 sessions per reversal.

2.5.2 Location discrimination serial reversal task

The location discrimination reversal task we developed differs from the one used in previous studies (Saifullah et al. 2020; C. H. Kim et al. 2015), in the sense that it essentially functions as the mirror image of the visual discrimination task. Instead of using a two-phase task with low and high degrees of separation between stimuli comprised of bright squares, we decided to take advantage of the animals' tendency to persevere after a correct choice. In other words, instead of having several within-session location-reversals, we opted for having a reversal-learning scheme across sessions, where we allowed mice to essentially become "sided" and then once the passing criteria is reached (>80% correct responses), we reverse the contingency, making the previously unrewarded side of the screen (S-), the new S+. In this task we also used the same images used in the visual discrimination task, but now they serve as distractors which mice need to ignore and focus only on the side of the screen that correspond to the S+. The pretraining sessions followed a similar structure to the the visual discrimination task, with a few notable differences.

The task starts with the (1) *Habituation* stage, which follows the same parameters described in the visual discrimination task. In the (2) *Cue Presentation* stage, a blinking cue (1x per second) appears on either the left or right side of the screen (depending on the starting location determined a priori by the experimenter) signaling the S+ location. The following pre-training stages – *Touchscreen Interaction* (3) and *Time-Out* (4) – follow the exact same criteria outlined in the previous task. In the 4th and the last stage of pre-training (*Pre-acquisition*), the blinking cue is eliminated, and we introduce 2 distractor images, the same ones used in the visual discrimination task, but here, only one of them can appear in a pseudo-random fashion, on each trial. The animals must ignore the distractor image and continue to touch the same side of the screen to obtain the reward. Finally, in the *Acquisition* stage, both distractor images are presented on either side of the screen in a pseudo-random manner across trials. The objective is for the animals to consistently select the S+ side of the screen. The contingencies are then reversed 5 times, with the S+ and S- switching between the right and left side of the screen at each reversal, with the passing criteria remaining at 80% correct responses.

2.6 *Data Analysis*

Behavioral performance was monitored through post-session video analysis. The data from each session were automatically saved as .CSV files, organized in Microsoft Excel (Office 2021), analyzed using GraphPad Prism (GraphPad Software Inc. Version 10.2.3), and the figures prepared using Adobe Illustrator (Adobe Systems Inc. Version 27.8.1).

Statistical analyses were conducted using ANOVAs, with a significance threshold set at $p < .05$. A paired t-test was specifically employed to compare error types in the visual discrimination serial reversal task. For post-hoc analyses, Tukey's multiple comparison test was

the primary choice, except for analyzing error types in the first reversal stage in both C57/BL6 and App^{NL-G-F} mice, where Fisher's LSD test was used.

3. Results

3.1 Visual Discrimination task

Figure 6 displays the group average learning curve across days for the visual discrimination task. While all mice successfully passed each learning stage and successfully completed the task, the group average learning curve remains below the 80% correct response threshold required for passing each reversal stage. This discrepancy can be attributed to the variability in the timing of stage completion among individual mice, with each starting and concluding each stage at different times. Additionally, the average performance is also influenced by the fact that for most of the task, individual mice frequently scored below the 80% criterion, except during the critical sessions where they met the passing threshold of 80% correct responses in two consecutive sessions. Therefore, while individual performances at specific points met the required threshold to progress, the aggregated data across all sessions and mice reflects a lower overall average.

3.1.1 Average number of sessions

The number of sessions required to complete the experiment varied across learning stages, with means and standard deviations as follows: Acquisition (Acq.) phase had a mean (M) of 21.44 (SD = 10.13), while Reversal 1 (R1) increased to M = 40.44 (SD = 10.30), with subsequent learning stages (R2 through R5) showing a gradual decrease in session counts. Specifically, R2 had an M = 32.11 (SD = 7.39), R3 an M = 24.89 (SD = 11.24), R4 an M = 22.67 (SD = 9.08), and R5 an M = 20.22 (SD = 8.45).

In order to assess performance differences across stages, a Repeated Measures One-Way ANOVA with Geisser-Greenhouse correction ($\varepsilon = 0.6690$), revealed significant variability among the session means, $F(3.345, 26.76) = 7.942$, $p < .001$. Tukey's multiple comparison test further identified significant differences between the Acquisition phase (Acq) and the first Reversal stage (R1), $p = .008$, and between R1 and R3 ($p = .021$), R4 ($p = .007$), and R5 ($p = .008$). All other comparisons between stages did not show significant differences ($p > 0.05$).

3.1.2 Average number of errors

When examining the average number of errors across the different learning stages, a clear trend of decreasing errors also emerged: Acq. (M = 230, SD = 145.2), followed by a peak at R1 (M = 603.6, SD = 174), with subsequent reductions seen in R2 (M = 456.8, SD = 111), R3 (M = 341.7, SD = 183.2), R4 (M = 291.7, SD = 134.1), and R5 (M = 270.8, SD = 146.3). A Repeated Measures ANOVA, conducted without assuming sphericity ($\varepsilon = 0.6932$), showed significant differences in the average number of errors across learning stages, $F(3.466, 27.73) = 10.49$, $p < 0.0001$; Post-hoc comparisons identified significant variations between Acq. and R1 ($p = 0.001$), and less pronounced yet significant differences between R1 and R3 ($p = 0.008$), R1 and R4 ($p = 0.006$), and R1 and R5 ($p = 0.007$). All other comparisons did not yield any significant differences between errors across different learning stages ($p > 0.05$).

3.1.3 Type of error

We conducted a focused analysis on errors during sessions where mice followed the response rule from the previous learning stage (Fig. 7C). Reversal 1 (R1) was selected as the primary stage for this examination due to its high incidence of response errors. We established a cutoff of 45% correct responses to categorize the errors: those occurring in sessions with ≤ 45% correct responses were classified as perseverance errors, and errors in sessions with performance above 45% were classified as learning errors. This classification includes those errors made in sessions where mice adopted a "win-stay, lose-switch" strategy, which typically occur around 50% correct responses. A paired t-test revealed no significant difference between perseveration errors (M = 277.9, SD = 120.9) and learning errors (M = 325.7, SD = 139.5) in the first reversal stage, despite a slight increase in learning errors, as observed in Figure 7 C ($t(8) = 0.7368$, $p = 0.4823$).

3.2 Location Discrimination task

All mice used in this study were able to learn the location discrimination reversal task. Figure 8 shows the average performance of the two groups, C57 and $App^{NL\text{-}G\text{-}F}$ across days. Both groups exhibited a sharp decline in performance at the onset of R1, dropping bellow 40% correct responses. However, C57 mice appeared to lean at a slightly faster rate than the $App^{NL\text{-}G\text{-}F}$ group, consistently maintaining performance above the 80% cutoff until the conclusion of R5. Additionally, C57 mice completed the task three days earlier than the $App^{NL\text{-}G\text{-}F}$ group, indicating a swifter acquisition of the task rules.

3.2.1 Average Number of sessions

Even though there were individual as well as group differences in the amount of time necessary for the animals to complete the task, the general tendency was to converge towards the minimum number of sessions required to pass each stage (2 consecutive sessions). A Two-Way Repeated Measures ANOVA indicated a significant interaction between Learning Stage and Genotype on the number of sessions to reach the passing criteria (≥ 80% correct responses in two consecutive sessions), $F(5, 80) = 4.935$, $p < .001$. Additionally, we found a significant main effect of Learning Stage, ($F(2.994, 47.90) = 50.78$, $p < .001$), and Genotype ($F(1, 16) = 7.806$, $p = .013$). No significant variability was attributed to individual differences among subjects, ($F(16, 80) = 1.218$, $p = .273$).

Post-hoc comparisons revealed a significant difference between C57 and $App^{NL\text{-}G\text{-}F}$ mice in the first Reversal stage (R1), with C57 mice showing a mean (M) of 4.11 errors (Standard Deviation, SD = 0.60) compared to $App^{NL\text{-}G\text{-}F}$ mice (M = 5.88, SD = 1.453), $p = 0.006$. No significant differences were observed in other stages, including the Acquisition phase (Acq) where both C57 and APPNL-GF mice recorded M = 2 errors (SD = 0), and subsequent Reversal stages: R2 (C57: M = 2.77, SD = 0.83; APPNL-GF: M = 3.33, SD = 1), R3 (C57: M = 2.44, SD = 0.88; APPNL-GF: M = 2.22, SD = 0.44), R4 (C57: M = 2.11, SD = 0.33; APPNL-GF: M = 2.33, SD = 0.70), and R5 (C57: M = 2, SD = 0; APPNL-GF: M = 2.11, SD = 0.33), all yielding $p > .05$.

Within-group analysis revealed distinct patterns of significant differences in the average number of sessions spent across learning stages for both C57 and App^{NL-G-F} mice. For C57 mice, comparisons between R1 and all other stages, except R2, showed significant differences: R1 vs Acq. ($p < 0.001$), R1 vs R3 ($p = 0.017$), R1 vs R4 ($p = 0.001$), and R1 vs R5 ($p < 0.001$). In contrast, the comparison between R1 and R2 only approached significance ($p = .055$), suggesting a less pronounced difference between these reversal stages.

In the App^{NL-G-F} group, R1 showed significant differences when compared to all other learning stages, highlighting a consistent pattern: R1 vs Acq. ($p < 0.001$), R1 vs R2 ($p = 0.008$), R1 vs R3 ($p < 0.001$), R1 vs R4 ($p = 0.001$), and R1 vs R5 ($p = 0.001$). Additionally, statistical analysis also identified significant differences between R2 and Acq. ($p = 0.032$), and between R2 and R5 ($p = 0.043$).

3.2.2 Average Number of errors

A similar trend was observed in terms of the average number of errors between C57 and App^{NL-G-F} mice across the different learning stages. A Two-way Repeated Measures ANOVA highlighted significant effects for the interaction between Learning Stage and Genotype, ($F(5, 80) = 5.405$, $p < .001$). Significant main effects were observed for Learning Stage, ($F(2.309, 36.95) = 75.72$, $p < .001$), and for Genotype, ($F(1, 16) = 7.037$, $p = .017$). Additionally, variability attributed to individual mice was also significant, ($F(16, 80) = 1.803$, $p = .045$). The only statistically significant difference between groups, was once again observed in R1 (C57: M = 43.77, SD = 11.98; APP: M = 71.33, SD = 24.28; $p = 0.010$).

Conversely, the comparisons revealed no significant differences in the Acquisition stage (C57: M = 3.88, SD = 2.47; APP: M = 2.55, SD = 1.74, $p = 0.206$) R2 (C57: M = 10.88, SD = 6.86; APP: M = 21.11, SD = 15.22, $p = 0.093$) R3 (C57: M = 9.55, SD = 10.86; APP: M = 8, SD = 9.02, $p = 0.746$) R4 (C57: M = 4.66, SD = 3.27; APP: M = 9.77, SD = 11.23, $p = 0.221$) and R5 (C57: M = 4.44, SD = 3.12; APP: M = 5.88, SD = 3.75, $p = 0.389$).

Within group comparisons showed once again, differences between R1 and every other learning stage for control mice (R1 vs Acq.: $p < 0.001$; R1 vs R2: $p < 0.001$; R1 vs R3: $p < 0.001$; R1 vs R4: $p < 0.001$; R1 vs R5: $p < 0.001$), whereas for App^{NL-G-F} mice differences were found between Acq. and R2 ($p = 0.031$), and R1 versus the remaining learning stages (R1 vs Acq.: $p < 0.001$; R1 vs R2: $p = 0.001$; R1 vs R3: $p < 0.001$; R1 vs R4: $p = 0.002$; R1 vs R5: $p < 0.001$).

3.2.3 Type of error

A Two-Way Repeated Measures ANOVA revealed no significant interaction between Error Type and Genotype, ($F(1, 16) = 1.378$, $p = .258$). However, significant main effects were observed for both Error Type, ($F(1, 16) = 25.74$, $p < .001$), and Genotype, ($F(1, 16) = 9.321$, $p = .008$), with no significant differences attributed to individual subjects, ($F(16, 16) = 0.6530$, $p = .798$). Post-hoc analysis using Fisher's LSD revealed a statistically significant difference in perseverative errors between the C57 and App^{NL-G-F} groups (C57: M = 32.77, SD = 11.13; APP: M = 53.11, SD = 24.55; $p = 0.008$), but not in learning errors (C57: M = 11, SD = 6.946; APP: M = 18.22, SD = 12.35; $p = 0.322$). Statistically significant within-group differences in terms of error type were also observed in both groups (C57: $p = 0.014$; APP: $p < 0.001$).

4. Discussion

4.1 Behavioural tasks

In the visual discrimination task, the C57BL/6 mice displayed a trend in the average number of errors across learning stages, which was also reflected in the average number of sessions, revealing a progressive improvement in performance over time. The lack of significant differences between R1 and R2, and among subsequent reversal stages, suggests that despite the initial struggle mice gradually adapt to the new reward contingencies. And although there was a slight increase in the average number of learning errors in comparison with perseveration errors, the difference was not statistically significant.

On the other hand, in the location discrimination serial reversal task, despite the considerable gap in terms of both the average number of sessions and average number of errors across the different learning stages, both wild-type and $App^{NL\text{-}G\text{-}F}$ mice showed a clear progression toward asymptotic performance. Mirroring the performance of the C57 mice in the visual discrimination task, both groups experienced significant challenges when first adjusting to reversed reward contingencies, reflecting the difficulty in overriding previously learned associations.

Both C57 and $App^{NL\text{-}G\text{-}F}$ mice revealed significant differences in terms of both average number of sessions and errors, particularly in R1. This suggests a stark contrast in cognitive flexibility between genotypes and better adaptability, with C57 mice adjusting more quickly to the reversals and making fewer mistakes compared to the $App^{NL\text{-}G\text{-}F}$ cohort. Furthermore, when examining the specific type of errors (perseverative versus learning errors), significant differences emerged between the genotypes, with $App^{NL\text{-}G\text{-}F}$ mice generally committing more perseverative errors. These findings underscore potential genotype-specific challenges in shifting strategies after rule changes, and overall cognitive flexibility, which could reflect broader implications in neurological or cognitive research, particularly in understanding conditions such as Alzheimer's disease (Braak and Braak 1991; Knopman et al. 2021; Llinas and Moreno 2017; Sasaguri et al. 2022; Guarino et al. 2019; Allegri, Harris, and Drake 2000; Walker, Lynn, and Chernoff 2018; McAllister et al. 2020).

The discrepancies observed in these tasks might stem from the extended time needed to establish and reverse the association between specific visual inputs, such as virtual objects or images, and a reward. Although the number of sessions required for the animals to learn the new reward contingency in the visual discrimination task decreased over time, perseverative behavior persisted until the fifth reversal. In contrast, in our location discrimination reversal task, such behavior was observed only from the first to the third reversals, with a significantly shorter timescale for cessation. By the fourth reversal, almost all mice had reached a performance asymptote, typically requiring just two sessions to meet the passing criterion.

Our findings suggest that further research is needed to fully understand the behavioral dynamics between these two tasks. Our version of the location discrimination task, differing from those reported in previous studies by employing "across session" instead of "within-session" reversals, presents unique challenges in terms of overwriting the previously acquired rules. This is not only due to the considerable number of individual trials required to meet the passing criterion, which strengthens the association between the rules and outcomes, but also due to the presence of distractor images that could influence decision-making. Interestingly, animals in the location discrimination task tended to ignore the visual cues and consistently choose a

specific side, suggesting that in this context, visual stimuli do not significantly impact their behavior. This observation raises important questions about the relative influence of spatial versus visual cues in shaping behavioral strategies.

Lastly, it is also important to acknowledge the length of the training procedures, which can be quite onerous for both the animals and the experimenters. This was particularly evident in the visual discrimination serial reversal task, with some animals taking up to 4 months to complete the task, conducting two to three daily sessions each consisting of thirty trials. The lengthy nature of this experiment was also noticeable when some mice began to lose motivation despite being water restricted, a condition that can lead to suboptimal performance levels.

4.2 Touchscreen apparatus

One of the most important aspects of any scientific endeavor is exploration, and while it is crucial to standardize behavioral procedures in research, allowing for experimentation and the expansion of methods is equally vital. This requires different labs experimenting with various hardware and software configurations for a comprehensive assessment of cognitive functions, as it is important to determine whether certain elements or steps in behavioral tasks, especially in touchscreen tasks, are indispensable features, or if they are subject to improvement or even elimination.

Our group tested various configurations before adopting a design inspired by the original touchscreen chambers, however, other groups have introduced their own designs without significantly deviating from the outcomes observed with standard setups (O'Leary et al. 2018; Eleftheriou et al. 2023; Wiesbrock, Musall, and Kampa 2022; Pineño 2014).

Among the configurations we tested, placing the reward tube directly below the screen, worked surprisingly well for most mice, provided that the inter-stimulus interval (ITI) allowed images to appear before the animal finished the reward. This means that animals don't necessarily need to initiate each trial if the ITI is time-based; that is, determined by the amount of time it takes the mouse to collect the reward, unlike the standard setup where the ITI starts once the animal collects the reward. Although the original settings described in several publications provide valid information, setting a fixed and specifically tailored ITI (5 seconds in our tasks) can also be a valid approach, as long as it ensures that images are displayed in time for a clear view upon approach.

Another modification to the standard touchscreen task setup was the exclusion of correction trials. Although our user interface has the option to select correction trials, after testing them during pilot tasks, we determined that they did not enhance animal performance or reduce the time spent in each learning stage, including pre-training. In fact, we observed that with correction trials, the animals used in the pilot experiments tended to lose interest in the task, despite being water restricted. This was particularly noticeable during the first reversal stage. Therefore, we decided to train the animals twice per day – morning and afternoon – with each session consisting of only the required 30 trials. Despite the lengthy training period in the visual discrimination task and the inter-subject variability in terms of learning capabilities, all mice were ultimately able to learn the task.

We also experimented with the apparatus layout. Despite choosing a trapezoidal shape with dimensions similar to those of the original mouse touchscreen chambers, we initially tried a square-shaped chamber, akin to traditional operant conditioning boxes. Although it featured

ample unused space, which could distract animals during pre-training, it actually proved to be technically sufficient for mouse training.

As for the adoption of a Samsung SM-T350 tablet as a touchscreen for our experimental setup, our choice was driven by both pragmatic considerations and the results of comprehensive testing. We evaluated multiple 7 and 4-inch touchscreens commercially available for Raspberry Pi devices, and despite all the screens being touch-capacitive, we found that their touch sensitivity was subpar. Most of these touchscreens failed to accurately register the rodent's touch input, resulting in an inability for the animals to learn task rules and associate specific behaviors with outcomes, rendering them unsuitable for our research needs.

Furthermore, while infrared frames or touch panels are standard in commercially available rodent touchscreen chambers, sourcing companies that can manufacture these to precise specifications can be challenging and incurs substantial costs. These factors undermine one of our main objectives with this study, which was to develop a cost-effective alternative. In contrast, using the SM-T350 tablets allowed us greater flexibility in developing the behavioral tasks and creating a library of virtual objects using on an Android platform. This approach not only maintained low costs but also ensured the reliability and sensitivity required for accurate behavioral research.

Lastly, regarding the visual stimuli, we opted for two colored images or virtual objects, unlike most studies, which use simple 2D black and white images. It is crucial to emphasize that the choice of stimuli should be determined primarily by the researchers' needs. This means any stimulus parameter – such as brightness, color, 2D versus 3D, stationary versus rotating/moving stimuli, fully visible versus partially occluded, and image/object size, among others – should be explored and modified according to the research objectives. In our case, we found the color dimension to be irrelevant for the serial-reversal design of the visual discrimination task we employed. Nevertheless, we have created a library with multiple images and virtual objects, including both color and black and white options, to serve the varying needs and goals of specific experiments. Given the growing interest in this technology, it is important to focus on flexibility, continuous experimentation, and innovation regarding its critical features, including hardware and software aspects as well as affordability.

Additionally, exploring new research avenues, such as integrating touchscreen technology directly into animals' home cages, holds promise for significant advancements. This strategy could not only mitigate stress from exposure to unfamiliar environments but also substantially reduce human-animal interaction, therefore minimizing the introduction of confounding variables that could skew results despite the standardization of experimental protocols. Allowing for the assessment of ethologically relevant behavior, while virtually eliminating experimenter involvement could represent a step forward in creating more humane and precise behavioral research methodologies.

5. Conclusion

Our custom-built touchscreen apparatus for mice has proven to be both practical and cost-effective, offering a viable alternative to more expensive commercial systems. By leveraging commercially available computer tablets integrated with a Raspberry Pi, our system not only reduces equipment costs but also provides detailed insights into cognitive flexibility and behavioral strategies. Through this approach, we developed both visual discrimination and location discrimination tasks with five reversals each, which allowed us to observe distinct

performance patterns. Despite similarities in their overall design, the two tasks require varying levels of cognitive flexibility, underscoring the need for further research into the specific mechanisms underlying these differences, and their implications for understanding cognitive and behavioral processes in different mouse models, and a broader comprehension of both normal and pathological brain functions.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Rui C. Pais: Conceptualization, Methodology, Investigation, Formal analysis, Writing – Original draft, Visualization, Project administration. **Ali Goldani:** Methodology, Software, Data Curation, Writing – Original draft, Visualization. **Jayden Hutchison:** Investigation. **Amirhossein Mazrouei:** Software. **Mostafa Khavaninzadeh:** Methodology. **Leonardo A. Molina**: Software. **Bruce L. McNaughton:** Resources. **Robert J. Sutherland:** Conceptualization, Supervision. **Majid H. Mohajerani:** Conceptualization, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Funding

This work was supported by Alberta Innovate and Natural Sciences and Engineering Research Council of Canada (grant 390930), awarded to Majid H. Mohajerani.

Acknowledgements

We thank Mike Zhou and Cameron Chin from the Engineering Department at the University of Waterloo, as well as Hardeep Singh Ryat from the University of Lethbridge, for their assistance with the design and assembly of the touchscreen chamber. We also extend our gratitude to Di Shao, Isabelle Gauthier, and all members of the Animal Welfare Committee at the University of Lethbridge for enabling our animal experiments.

Data Availability Statement

The dataset and config files generated for this study can be found in the project's GitHub repository [https://github.com/Mohajerani-Lab/touchscreen-chamber-unity].

References

Allegri, R F, P Harris, and M Drake. 2000. "La Evaluación Neuropsicológica En La Enfermedad de Alzheimer." *Rev Neurol Arg*, no. 1431, 11–15.

Alsiö, J., S. R.O. Nilsson, F. Gastambide, R. A.H. Wang, S. A. Dam, A. C. Mar, M. Tricklebank, and T. W. Robbins. 2015. "The Role of 5-HT2C Receptors in Touchscreen Visual Reversal Learning in the Rat: A Cross-Site Study." *Psychopharmacology* 232 (21–22): 4017–31. https://doi.org/10.1007/s00213-015-3963-5.

Barnard, Ilne L., Timothy J. Onofrychuk, Dan L. McElroy, and John G. Howland. 2021. "The Touchscreen-Based Trial-Unique, Nonmatching-To-Location (TUNL) Task as a Measure of Working Memory and Pattern Separation in Rats and Mice." *Current Protocols* 1 (9). https://doi.org/10.1002/cpz1.238.

Bloom, George S. 2014. "Amyloid-β and Tau: The Trigger and Bullet in Alzheimer Disease Pathogenesis." *JAMA Neurology* 71 (4): 505–8. https://doi.org/10.1001/jamaneurol.2013.5847.

Boulougouris, Vasileios, Jeffrey W. Dalley, and Trevor W. Robbins. 2007. "Effects of Orbitofrontal, Infralimbic and Prelimbic Cortical Lesions on Serial Spatial Reversal Learning in the Rat." *Behavioural Brain Research* 179 (2): 219–28. https://doi.org/10.1016/j.bbr.2007.02.005.

Braak, H, and E Braak. 1991. "Neuropathological Stageing of Alzheimer-Related Changes." *Acta Neuropathologica* 82 (June):239–59. https://doi.org/10.1007/BF00308809.

Brigman, Jonathan L., Carolyn Graybeal, and Andrew Holmes. 2010. "Predictably Irrational: Assaying Cognitive Inflexibility in Mouse Models of Schizophrenia." *Frontiers in Neuroscience* 4 (MAY): 19–28. https://doi.org/10.3389/neuro.01.013.2010.

Broeck, Lore Van den, Pierre Hansquine, Zsuzsanna Callaerts-Vegh, and Rudi D'hooge. 2019. "Impaired Reversal Learning in APPPS1-21 Mice in the Touchscreen Visual Discrimination Task." *Frontiers in Behavioral Neuroscience* 13 (April). https://doi.org/10.3389/fnbeh.2019.00092.

Bryce, Courtney A., and John G. Howland. 2015. "Stress Facilitates Late Reversal Learning Using a Touchscreen-Based Visual Discrimination Procedure in Male Long Evans Rats." *Behavioural Brain Research* 278 (February):21–28. https://doi.org/10.1016/j.bbr.2014.09.027.

Bussey, T J, L M Saksida, and L A Rothblat. 2001. "Discrimination of Computer-Graphic Stimuli by Mice: A Method for the Behavioral Characterization of Transgenic and Gene-Knockout Models." *Behavioral Neuroscience* 115 (4): 957–60. https://doi.org/10.1037/0735-7044.115.4.957.

Bussey, Timothy J., Barry J. Everitt, and Trevor W. Robbins. 1997. "Dissociable Effects of Cingulate and Medial Frontal Cortex Lesions on Stimulus-Reward Learning Using a Novel Pavlovian Autoshaping Procedure for the Rat: Implications for the Neurobiology of Emotion." *Behavioral Neuroscience* 111 (5): 908–19. https://doi.org/10.1037/0735-7044.111.5.908.

Bussey, Timothy J, Janice L Muir, and Trevor W Robbins. 1994. "A Novel Automated Touchscreen Procedure for Assessing Learning in the Rat Using Computer Graphic Stimuli." *Neuroscience Research Communications* 15 (2): 103–10.

Bussey, Timothy J., Tina L. Padain, Elizabeth A. Skillings, Boyer D. Winters, A. Jennifer Morton, and Lisa M. Saksida. 2008. "The Touchscreen Cognitive Testing Method for Rodents: How to Get the Best out of Your Rat." *Learning and Memory* 15 (7): 516–23. https://doi.org/10.1101/lm.987808.

Castañé Anna, A., David E.H. Theobald, and Trevor W. Robbins. 2010. "Selective Lesions of the Dorsomedial Striatum Impair Serial Spatial Reversal Learning in Rats." *Behavioural Brain Research* 210 (1): 74–83. https://doi.org/10.1016/j.bbr.2010.02.017.

Chudasama, Y, and Trevor W Robbins. 2003. "Behavioral/Systems/Cognitive Dissociable Contributions of the Orbitofrontal and Infralimbic Cortex to Pavlovian Autoshaping and

Discrimination Reversal Learning: Further Evidence for the Functional Heterogeneity of the Rodent Frontal Cortex."

Clatworthy, Philip L., Simon J.G. Lewis, Laurent Brichard, Young T. Hong, David Izquierdo, Luke Clark, Roshan Cools, et al. 2009. "Dopamine Release in Dissociable Striatal Subregions Predicts the Different Effects of Oral Methylphenidate on Reversal Learning and Spatial Working Memory." *Journal of Neuroscience* 29 (15): 4690–96. https://doi.org/10.1523/JNEUROSCI.3266-08.2009.

Cools, Roshan, Luke Clark, Adrian M Owen, and Trevor W Robbins. 2002. "Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging." www.mrc-cbu.cam.ac.uk/imaging.

Creighton, Samantha D., Heather A. Collett, Paula M. Zonneveld, Raiva A. Pandit, Andrew E. Huff, Kristen H. Jardine, Bruce L. McNaughton, and Boyer D. Winters. 2019. "Development of an 'Object Category Recognition' Task for Mice: Involvement of Muscarinic Acetylcholine Receptors." *Behavioral Neuroscience* 133 (5): 527–36. https://doi.org/10.1037/bne0000331.

D'Cruz, Anna Maria, Michael E. Ragozzino, Matthew W. Mosconi, Sunil Shrestha, Edwin H. Cook, and John A. Sweeney. 2013. "Reduced Behavioral Flexibility in Autism Spectrum Disorders." *Neuropsychology* 27 (2): 152–60. https://doi.org/10.1037/a0031721.

Dias, R, T W Robbins, and A C Roberts. 1996. "Dissociation in Prefrontal Cortex of Affective and Attentional Shifts." *Nature* 380 (7): 69–72. https://doi.org/10.1038/380069a0.

Dickson, Price E., Beau Corkill, Eric McKimm, Mellessa M. Miller, Michele A. Calton, Daniel Goldowitz, Charles D. Blaha, and Guy Mittleman. 2013. "Effects of Stimulus Salience on Touchscreen Serial Reversal Learning in a Mouse Model of Fragile X Syndrome." *Behavioural Brain Research* 252 (September):126–35. https://doi.org/10.1016/j.bbr.2013.05.060.

Dumont, Julie R., Ryan Salewski, and Flavio Beraldo. 2021. "Critical Mass: The Rise of a Touchscreen Technology Community for Rodent Cognitive Testing." *Genes, Brain and Behavior* 20 (1). https://doi.org/10.1111/gbb.12650.

Eleftheriou, Constantinos, Thomas Clarke, V. Poon, Marie Zechner, and Ian Duguid. 2023. "Visiomode: An Open-Source Platform for Building Rodent Touchscreen-Based Behavioral Assays." *Journal of Neuroscience Methods* 386 (February). https://doi.org/10.1016/j.jneumeth.2022.109779.

Ettcheto, Miren, Sonia Abad, Dmitry Petrov, Ignacio Pedrós, Oriol Busquets, Elena Sánchez-López, Gemma Casadesús, et al. 2018. "Early Preclinical Changes in Hippocampal CREB-Binding Protein Expression in a Mouse Model of Familial Alzheimer's Disease." *Molecular Neurobiology* 55 (6): 4885–95. https://doi.org/10.1007/s12035-017-0690-4.

Ferster, Charles B. 1953. "The Use of the Free Operant in the Analysis of Behavior." *Psychological Bulletin* 50 (4): 189–208. https://doi.org/10.1037/h0032956.

Folch, J., M. Ettcheto, D. Petrov, S. Abad, I. Pedrós, M. Marin, J. Olloquequi, and A. Camins. 2018. "Review of the Advances in Treatment for Alzheimer Disease: Strategies for Combating β-Amyloid Protein." *Neurología (English Edition)* 33 (1): 47–58. https://doi.org/10.1016/j.nrleng.2015.03.019.

Fowler, K, G Handelmann, S Mitchell, M Mishkin, C Anderson, J Argerson, T Cox, et al. 1980. "Object Discrimination by Rats : The Role of Frontal and Hippocampal Systems in Retention and Reversal I" 24:33–38.

Graybeal, Carolyn, Michael Feyder, Emily Schulman, Lisa M. Saksida, Timothy J. Bussey, Jonathan L. Brigman, and Andrew Holmes. 2011. “Paradoxical Reversal Learning Enhancement by Stress or Prefrontal Cortical Damage: Rescue with BDNF.” *Nature Neuroscience* 14 (12): 1507–9. https://doi.org/10.1038/nn.2954.

Gruner, Patricia, and Christopher Pittenger. 2017. “Cognitive Inflexibility in Obsessive-Compulsive Disorder.” *Neuroscience*. Elsevier Ltd. https://doi.org/10.1016/j.neuroscience.2016.07.030.

Guardia-Laguarta, Cristina, Marta Pera, Jordi Clarimón, José Luis Molinuevo, Raquel Sánchez-Valle, Albert Lladó, Mireia Coma, et al. 2010. “Clinical, Neuropathologic, and Biochemical Profile of the Amyloid Precursor Protein I716F Mutation.” *Journal of Neuropathology & Experimental Neurology* 69 (1): 53–59. https://doi.org/10.1097/NEN.0b013e3181c6b84d.

Guarino, Angela, Francesca Favieri, Ilaria Boncompagni, Francesca Agostini, Micaela Cantone, and Maria Casagrande. 2019. “Executive Functions in Alzheimer Disease: A Systematic Review.” *Frontiers in Aging Neuroscience*. Frontiers Media S.A. https://doi.org/10.3389/fnagi.2018.00437.

Hampshire, Adam, and Adrian M. Owen. 2006. “Fractionating Attentional Control Using Event-Related FMRI.” *Cerebral Cortex* 16 (12): 1679–89. https://doi.org/10.1093/cercor/bhj116.

Harper, James D, and Peter T Lansbury. 1997. “Models of Amyloid Seeding in Alzheimer’s Disease and Scrapie: Mechanistic Truths and Physiological Consequences of the Time-Dependent Solubility of Amyloid Proteins.” *Annu. Rev. Biochem* 66:385–407. https://doi.org/10.1146/annurev.biochem.66.1.385.

Hornak, J, J O’doherty, J Bramham, E T Rolls, R G Morris, P R Bullock, and C E Polkey. 2004. “Reward-Related Reversal Learning after Surgical Excisions in Orbito-Frontal or Dorsolateral Prefrontal Cortex in Humans.” http://direct.mit.edu/jocn/article-pdf/16/3/463/1934771/089892904322926791.pdf.

Horner, Alexa E., Christopher J. Heath, Martha Hvoslef-Eide, Brianne A. Kent, Chi Hun Kim, Simon R.O. Nilsson, Johan Alsiö, et al. 2013. “The Touchscreen Operant Platform for Testing Learning and Memory in Rats and Mice.” *Nature Protocols* 8 (10): 1961–84. https://doi.org/10.1038/nprot.2013.122.

Hvoslef-Eide, M., A. C. Mar, S. R.O. Nilsson, J. Alsiö, C. J. Heath, L. M. Saksida, T. W. Robbins, and T. J. Bussey. 2015. “The NEWMEDS Rodent Touchscreen Test Battery for Cognition Relevant to Schizophrenia.” *Psychopharmacology*. Springer Verlag. https://doi.org/10.1007/s00213-015-4007-x.

Hvoslef-Eide, M., S. R.O. Nilsson, L. M. Saksida, and T. J. Bussey. 2016. “Cognitive Translation Using the Rodent Touchscreen Testing Approach.” In *Current Topics in Behavioral Neurosciences*, 28:423–47. Springer Verlag. https://doi.org/10.1007/7854_2015_5007.

Izquierdo, A., J. L. Brigman, A. K. Radke, P. H. Rudebeck, and A. Holmes. 2017. “The Neural Basis of Reversal Learning: An Updated Perspective.” *Neuroscience*. Elsevier Ltd. https://doi.org/10.1016/j.neuroscience.2016.03.021.

Izquierdo, Alicia, and J. David Jentsch. 2012. “Reversal Learning as a Measure of Impulsive and Compulsive Behavior in Addictions.” *Psychopharmacology*. https://doi.org/10.1007/s00213-011-2579-7.

Izquierdo, Alicia, Lisa M. Wiedholz, Rachel A. Millstein, Rebecca J. Yang, Timothy J. Bussey, Lisa M. Saksida, and Andrew Holmes. 2006. “Genetic and Dopaminergic Modulation of Reversal Learning in a Touchscreen-Based Operant Procedure for Mice.” *Behavioural Brain Research* 171 (2): 181–88. https://doi.org/10.1016/j.bbr.2006.03.029.

Jara-Rizzo, María F., Juan F. Navas, Jose A. Rodas, and José C. Perales. 2020. “Decision-Making Inflexibility in a Reversal Learning Task Is Associated with Severity of Problem Gambling Symptoms but Not with a Diagnosis of Substance Use Disorder.” *BMC Psychology* 8 (1): 120. https://doi.org/10.1186/s40359-020-00482-6.

Kim, Chi H., Carola Romberg, Martha Hvoslef-Eide, Charlotte A. Oomen, Adam C. Mar, Christopher J. Heath, Andrée Anne Berthiaume, Timothy J. Bussey, and Lisa M. Saksida. 2015. “Trial-Unique, Delayed Nonmatching-to-Location (TUNL) Touchscreen Testing for Mice: Sensitivity to Dorsal Hippocampal Dysfunction.” *Psychopharmacology* 232 (21–22): 3935–45. https://doi.org/10.1007/s00213-015-4017-8.

Kim, Myeongwon, Chuljung Kwak, Nam Kyung Yu, and Bong Kiun Kaang. 2016. “Optimization of the Touchscreen Paired-Associate Learning (PAL) Task for Mice and Its Dorsal Hippocampal Dependency.” *Animal Cells and Systems* 20 (5): 229–36. https://doi.org/10.1080/19768354.2016.1221855.

Knopman, David S., Helene Amieva, Ronald C. Petersen, Gäel Chételat, David M. Holtzman, Bradley T. Hyman, Ralph A. Nixon, and David T. Jones. 2021. “Alzheimer Disease.” *Nature Reviews Disease Primers* 7 (1). https://doi.org/10.1038/s41572-021-00269-y.

Kosaki, Yutaka, and Shigeru Watanabe. 2012. “Dissociable Roles of the Medial Prefrontal Cortex, the Anterior Cingulate Cortex, and the Hippocampus in Behavioural Flexibility Revealed by Serial Reversal of Three-Choice Discrimination in Rats.” *Behavioural Brain Research* 227 (1): 81–90. https://doi.org/10.1016/j.bbr.2011.10.039.

Kwak, Chuljung, Chae Seok Lim, and Bong Kiun Kaang. 2016. “Assessments of Cognitive Abilities in a Mouse Model of Parkinson’s Disease with a Touch Screen Test.” *Behavioural Brain Research* 301 (March):63–71. https://doi.org/10.1016/j.bbr.2015.12.016.

Kwak, Chuljung, Chae-Seok Lim, and Bong-Kiun Kaang. 2015. “Development of a Touch-Screen-Based Paradigm for Assessing Working Memory in the Mouse.” *Experimental Neurobiology* 24 (1): 84–89. https://doi.org/10.5607/en.2015.24.1.84.

Lacoursiere, Sean G, Jiri Safar, David, Westaway, Majid H Mohajerani, and Robert J Sutherland. 2022. “The Effect of Aβ Seeding Is Dependent on the Presence of Knock-in Genes in the AppNL-G-F Mice.” *Frontiers in Dementia*, September. https://doi.org/10.3389/frdem.2022.941879.

Lafleche, Ginette, and Marilyn S Albert. 1995. “Executive Function Deficits in Mild Alzheimer’s Disease.” *Neuropsychology*. Vol. 9.

Latif-Hernandez, Amira, Victor Sabanov, Tariq Ahmed, Katleen Craessaerts, Takashi Saito, Takaomi Saido, and Detlef Balschun. 2020. “The Two Faces of Synaptic Failure in App NL-G-Fknock-in Mice.” *Alzheimer’s Research and Therapy* 12 (1). https://doi.org/10.1186/s13195-020-00667-6.

Latif-Hernandez, Amira, Disha Shah, Kathleen Craessaerts, Takaomi Saido, Takashi Saito, Bart De Strooper, Annemie Van der Linden, and Rudi D’Hooge. 2019. “Subtle Behavioral Changes and Increased Prefrontal-Hippocampal Network Synchronicity in APP NL−G−F Mice before Prominent Plaque Deposition.” *Behavioural Brain Research* 364 (May):431–41. https://doi.org/10.1016/j.bbr.2017.11.017.

Llinas, Rodolfo, and Herman Moreno. 2017. “Perspective on Calcium and Alzheimer Disease.” *Alzheimer’s & Dementia* 13 (January): 1–2. https://doi.org/10.1016/j.jalz.2017.01.004.

Lucantonio, Federica, Daniele Caprioli, and Geoffrey Schoenbaum. 2014. “Transition from ‘model-Based’ to ‘Model-Free’ Behavioral Control in Addiction: Involvement of the

Orbitofrontal Cortex and Dorsolateral Striatum." *Neuropharmacology*. Elsevier Ltd. https://doi.org/10.1016/j.neuropharm.2013.05.033.
Mar, Adam C., Alexa E. Horner, Simon R.O. Nilsson, Johan Alsiö, Brianne A. Kent, Chi Hun Kim, Andrew Holmes, Lisa M. Saksida, and Timothy J. Bussey. 2013. "The Touchscreen Operant Platform for Assessing Executive Function in Rats and Mice." *Nature Protocols* 8 (10): 1985–2005. https://doi.org/10.1038/nprot.2013.123.
Marazziti, Donatella, Giorgio Consoli, Michela Picchetti, Marina Carlini, and Luca Faravelli. 2010. "Cognitive Impairment in Major Depression." *European Journal of Pharmacology*. https://doi.org/10.1016/j.ejphar.2009.08.046.
Markham, Kichael R, Allen E Butt, and Michael J Dougher. 1996. "A Computer Touch-Screen Apparatus for Training Visual Discriminations in Rats." *Training* 65 (I): 173–82. https://doi.org/10.1901/jeab.1996.65-173.
Marquardt, Kristin, Rahul Sigdel, and Jonathan L. Brigman. 2017. "Touch-Screen Visual Reversal Learning Is Mediated by Value Encoding and Signal Propagation in the Orbitofrontal Cortex." *Neurobiology of Learning and Memory* 139 (March):179–88. https://doi.org/10.1016/j.nlm.2017.01.006.
McAllister, Brendan B., Sean G. Lacoursiere, Robert J. Sutherland, and Majid H. Mohajerani. 2020. "Intracerebral Seeding of Amyloid-β and Tau Pathology in Mice: Factors Underlying Prion-like Spreading and Comparisons with α-Synuclein." *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2020.01.026.
Mehla, Jogender, Scott H. Deibel, Hadil Karem, Nancy S. Hong, Shakhawat R. Hossain, Sean G. Lacoursiere, Robert J. Sutherland, Majid H. Mohajerani, and Robert J. McDonald. 2023. "Repeated Multi-Domain Cognitive Training Prevents Cognitive Decline, Anxiety and Amyloid Pathology Found in a Mouse Model of Alzheimer Disease." *Communications Biology* 6 (1). https://doi.org/10.1038/s42003-023-05506-6.
Mehla, Jogender, Sean G. Lacoursiere, Valerie Lapointe, Bruce L. McNaughton, Robert J. Sutherland, Robert J. McDonald, and Majid H. Mohajerani. 2019. "Age-Dependent Behavioral and Biochemical Characterization of Single APP Knock-in Mouse (APPNL-G-F/NL-G-F) Model of Alzheimer's Disease." *Neurobiology of Aging* 75 (March):25–37. https://doi.org/10.1016/j.neurobiolaging.2018.10.026.
Millan, Mark J., Yves Agid, Martin Brüne, Edward T. Bullmore, Cameron S. Carter, Nicola S. Clayton, Richard Connor, et al. 2012. "Cognitive Dysfunction in Psychiatric Disorders: Characteristics, Causes and the Quest for Improved Therapy." *Nature Reviews Drug Discovery*. https://doi.org/10.1038/nrd3628.
Monni, Alessandra, Michele Scandola, Sébastien Hélie, and L. Francesca Scalas. 2023. "Cognitive Flexibility Assessment with a New Reversal Learning Task Paradigm Compared with the Wisconsin Card Sorting Test Exploring the Moderating Effect of Gender and Stress." *Psychological Research* 87 (5): 1439–53. https://doi.org/10.1007/s00426-022-01763-y.
Nilsson, Per, Takashi Saito, and Takaomi C. Saido. 2014. "New Mouse Model of Alzheimer's." *ACS Chemical Neuroscience*. American Chemical Society. https://doi.org/10.1021/cn500105p.
Nilsson, S R O, L M Saksida, and T J Bussey. 2016. "Cognitive Translation Using the Rodent Touchscreen Testing Approach," 423–47. https://doi.org/10.1007/7854.
Nithianantharajah, J, A G Mckechanie, T J Stewart, and M Johnstone. 2015. "Bridging the Translational Divide : Identical Cognitive Touchscreen Testing in Mice and Humans

Carrying Mutations in a Disease- Relevant Homologous Gene." *Nature Publishing Group*, no. February, 3–7. https://doi.org/10.1038/srep14613.
Nithianantharajah, J., A. G. McKechanie, T. J. Stewart, M. Johnstone, D. H. Blackwood, D. St Clair, S. G.N. Grant, T. J. Bussey, and L. M. Saksida. 2015. "Bridging the Translational Divide: Identical Cognitive Touchscreen Testing in Mice and Humans Carrying Mutations in a Disease-Relevant Homologous Gene." *Scientific Reports* 5 (February): 3–7. https://doi.org/10.1038/srep14613.
Odland, Anna U, Rune Sandahl, and Jesper T Andreasen. 2021. "Sequential Reversal Learning: A New Touchscreen Schedule for Assessing Cognitive Flexibility in Mice." *Psychopharmacology* 238:383–97. https://doi.org/10.1007/s00213-020.
O'Leary, James D., Olivia F. O'Leary, John F. Cryan, and Yvonne M. Nolan. 2018. "A Low-Cost Touchscreen Operant Chamber Using a Raspberry PiTM." *Behavior Research Methods* 50 (6): 2523–30. https://doi.org/10.3758/s13428-018-1030-y.
Phillips, Benjamin U, Christopher J Heath, Zofia Ossowska, Timothy J Bussey, and Lisa M Saksida. 2017. "Optimisation of Cognitive Performance in Rodent Operant (Touchscreen) Testing : Evaluation and Effects of Reinforcer Strength." https://doi.org/10.3758/s13420-017-0260-7.
Pineño, Oskar. 2014. "ArduiPod Box: A Low-Cost and Open-Source Skinner Box Using an IPod Touch and an Arduino Microcontroller." *Behavior Research Methods* 46 (1): 196–205. https://doi.org/10.3758/s13428-013-0367-5.
Pinkston, Jonathan W. 2022. "Operant Responding: Beyond Rate and Interresponse Times." *Brain Research Bulletin* 186 (August):79–87. https://doi.org/10.1016/j.brainresbull.2022.05.009.
Saifullah, Md Ali Bin, Okiru Komine, Yutao Dong, Kazuya Fukumoto, Akira Sobue, Fumito Endo, Takashi Saito, Takaomi C. Saido, Koji Yamanaka, and Hiroyuki Mizoguchi. 2020. "Touchscreen-Based Location Discrimination and Paired Associate Learning Tasks Detect Cognitive Impairment at an Early Stage in an App Knock-in Mouse Model of Alzheimer's Disease." *Molecular Brain* 13 (1). https://doi.org/10.1186/s13041-020-00690-6.
Saito, Takashi, Yukio Matsuba, Naomi Mihira, Jiro Takano, Per Nilsson, Shigeyoshi Itohara, Nobuhisa Iwata, and Takaomi C. Saido. 2014. "Single App Knock-in Mouse Models of Alzheimer's Disease." *Nature Neuroscience* 17 (5): 661–63. https://doi.org/10.1038/nn.3697.
Sakagami, Takayuki, and Kennon A. Lattal. 2016. "The Other Shoe: An Early Operant Conditioning Chamber for Pigeons." *Behavior Analyst* 39 (1): 25–39. https://doi.org/10.1007/s40614-016-0055-8.
Sasaguri, Hiroki, Shoko Hashimoto, Naoto Watamura, Kaori Sato, Risa Takamura, Kenichi Nagata, Satoshi Tsubuki, et al. 2022. "Recent Advances in the Modeling of Alzheimer's Disease." *Frontiers in Neuroscience*. Frontiers Media S.A. https://doi.org/10.3389/fnins.2022.807473.
Shin, Jong-Yeon, Saet-Byeol Yu, Un-Young Yu, Sang-Mee Ahnjo, and Jung-Hyuck Ahn. 2010. "Swedish Mutation within Amyloid Precursor Protein Modulates Global Gene Expression towards the Pathogenesis of Alzheimer's Disease." *BMB Reports* 43 (10): 704–9. https://doi.org/10.5483/bmbrep.2010.43.10.704.
Skinner, B. F. 1937. "Two Types of Conditioned Reflex: A Reply to Konorski and Miller." *Journal of General Psychology* 16 (1): 272–79. https://doi.org/10.1080/00221309.1937.9917951.

Skinner, B, F. 1986. “Some Thoughts about the Future.” *Journal Of The Experimental Analysis Of Behavior* 45 (March):229–35.
Staddon, J. E.R., and D. T. Cerutti. 2003. “Operant Conditioning.” *Annual Review of Psychology*. https://doi.org/10.1146/annurev.psych.54.101601.145124.
Sullivan, Jacqueline A. 2022. “Novel Tool Development and the Dynamics of Control: The Rodent Touchscreen Operant Chamber as a Case Study.” *Philosophy of Science* 89 (5): 1203–12. https://doi.org/10.1017/psa.2022.63.
Talpos, J. C., B. D. Winters, R. Dias, L. M. Saksida, and T. J. Bussey. 2009. “A Novel Touchscreen-Automated Paired-Associate Learning (PAL) Task Sensitive to Pharmacological Manipulation of the Hippocampus: A Translational Rodent Model of Cognitive Impairments in Neurodegenerative Disease.” *Psychopharmacology* 205 (1): 157–68. https://doi.org/10.1007/s00213-009-1526-3.
Upīte, Jolanta, Inga Kadish, Thomas van Groen, and Baiba Jansone. 2020. “Subchronic Administration of Auranofin Reduced Amyloid-β Plaque Pathology in a Transgenic APPNL-G-F/NL-G-F Mouse Model.” *Brain Research* 1746 (November). https://doi.org/10.1016/j.brainres.2020.147022.
Valerius, Gabriele, Anne Lumpp, Anne-Katrin Kuelz, Tobias Freyer, and Ulrich Voderholzer. 2008. “Reversal Learning as a Neuropsychological Indicator for the Neuropathology of Obsessive Compulsive Disorder? A Behavioral Study.” *The Journal of Neuropsychiatry and Clinical Neurosciences*. Vol. 20. http://neuro.psychiatryonline.org.
Walker, Lary C., David G. Lynn, and Yury O. Chernoff. 2018. “A Standard Model of Alzheimer’s Disease?” *Prion*. Taylor and Francis Inc. https://doi.org/10.1080/19336896.2018.1525256.
Wang, Hao, Na Sun, Xinyue Wang, Jinyuan Han, Yongxiang Zhang, Yan Huang, and Wenxia Zhou. 2022. “A Touchscreen-Based Paradigm to Measure Visual Pattern Separation and Pattern Completion in Mice.” *Frontiers in Neuroscience* 16 (August). https://doi.org/10.3389/fnins.2022.947742.
Weiss, Stanley J. 1972. “Stimulus Compounding in Free-Operant and Classical Conditioning. A Review and Analysis.” *Psychological Bulletin* 78 (3): 189–208. https://doi.org/10.1037/h0032956.
Wetzel, Mary C. 1986. “Operant Conditioning in Motor and Neural Integration.” *Neuroscience & Biobehavioral Reviews* 10 (4): 387–429. https://doi.org/10.1016/0149-7634(86)90004-7.
Wiesbrock, Christopher, Simon Musall, and Björn M. Kampa. 2022. “A Flexible Python-Based Touchscreen Chamber for Operant Conditioning Reveals Improved Visual Perception of Cardinal Orientations in Mice.” *Frontiers in Cellular Neuroscience* 16 (October). https://doi.org/10.3389/fncel.2022.866109.
Winters, Boyer D., Lisa M. Saksida, and Timothy J. Bussey. 2008. “Object Recognition Memory: Neurobiological Mechanisms of Encoding, Consolidation and Retrieval.” *Neuroscience and Biobehavioral Reviews* 32 (5): 1055–70. https://doi.org/10.1016/j.neubiorev.2008.04.004.
Wobrock, Thomas, Ullrich K.H. Ecker, Harald Scherk, Thomas Schneider-Axmann, Peter Falkai, and Oliver Gruber. 2009. “Cognitive Impairment of Executive Function as a Core Symptom of Schizophrenia.” *World Journal of Biological Psychiatry* 10 (4 PART 2): 442–51. https://doi.org/10.1080/15622970701849986.
Wright, Anthony A, Robert G Cook, Jacquelyne J Rivera, Stephen F Sanns, and Juan D Delius. 1988. “Concept Learning By pigeons: Matching-to-Sample with Trial-Unique Video Picture Stimuli.” *Animal Learning & Behavior*. Vol. 16.

Figures and Legends

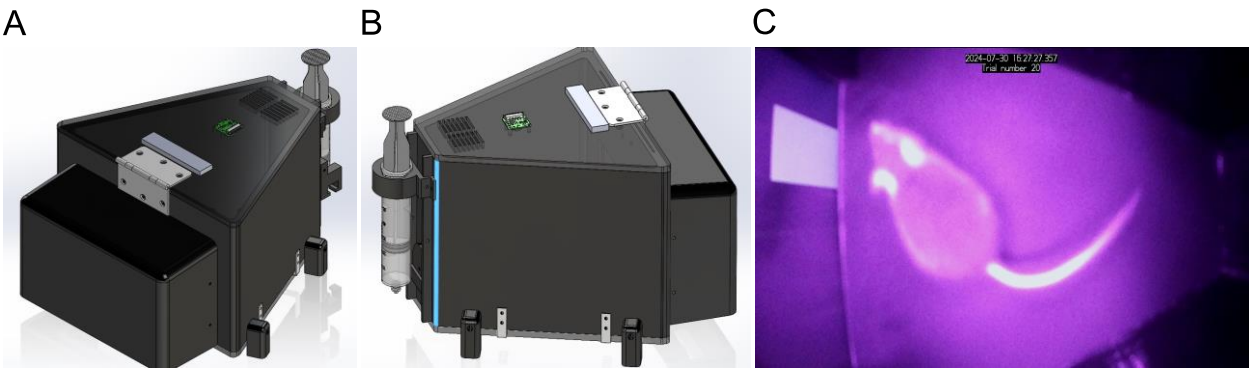

Figure 1: **Custom-built touchscreen chamber for mouse behavioral studies**. The tree-quarter (A) and side profile (B) views of the touchscreen chamber, highlighting the integrated design features and structural components. C) Interior view of the chamber during a pre-training session of the location discrimination serial reversal task, with a blinking cue on the right side of the screen.

```
<Experiment>
    <function id="prepare">
        <!-- General settings -->
    </function>
    <function id="main">
        <!-- Experiment setup -->
    </function>
    <function id="reward">
        <!-- Reward setup -->
    </function>
    <function id="time-out">
        <!-- Time-Out setup -->
    </function>
    <function id="trial">
        <!-- Trial Setup -->
    </function>
</Experiment>
```

Figure 2: **General outline of an XML configuration file.** Each section is enclosed in a "function" tag.

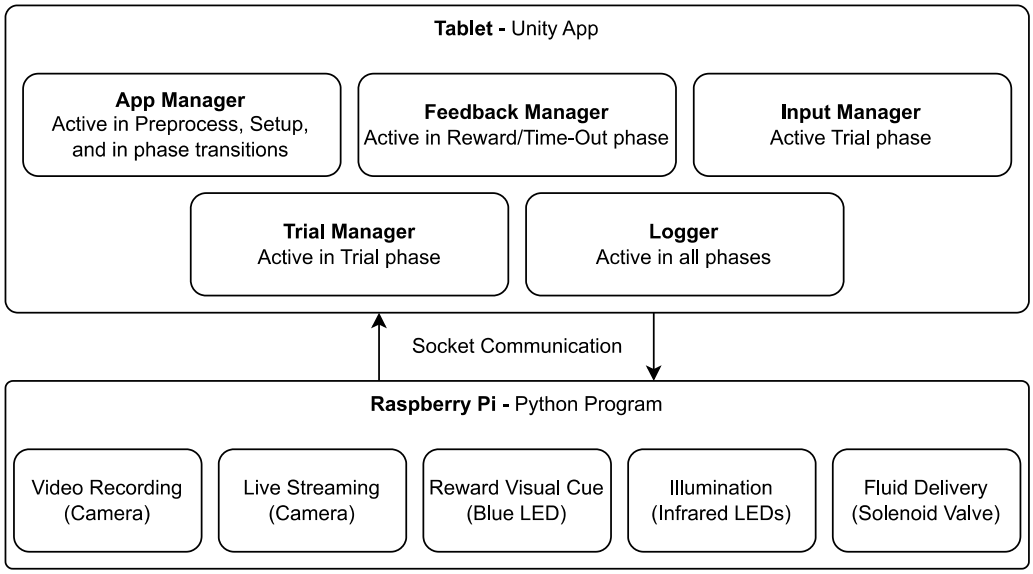

Figure 3: **The main software and hardware components.** Components in the Unity application run in their own event loops, which are only active in specific conditions based on the phase of the trial.

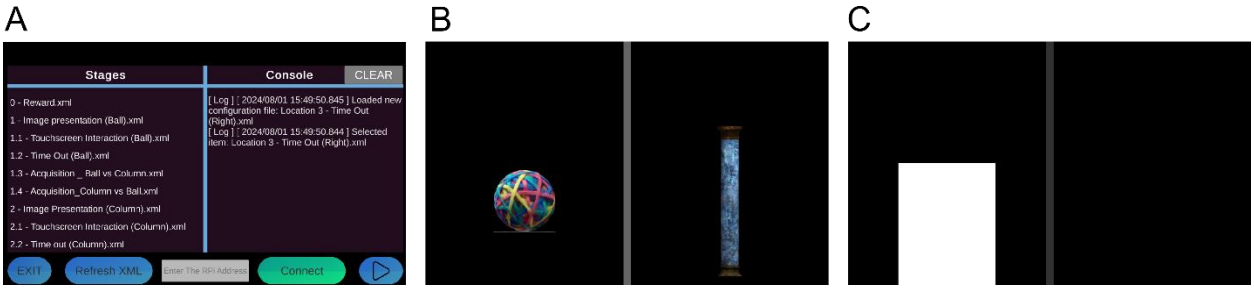

Figure 4: **Screenshots of the software. A)** Main menu of the program, where the user gets to choose the configuration file (learning stage) for the experiment. Configuring a connection to the RPi controller is also established in this page; the user inputs the network address of the RPi and initiates the connection. **B)** Screenshot of a 2-section visual discrimination task. **C)** Screenshot of a 2-section location task during pre-training, where a blinking cue appears on the screen to signal the S+ location.

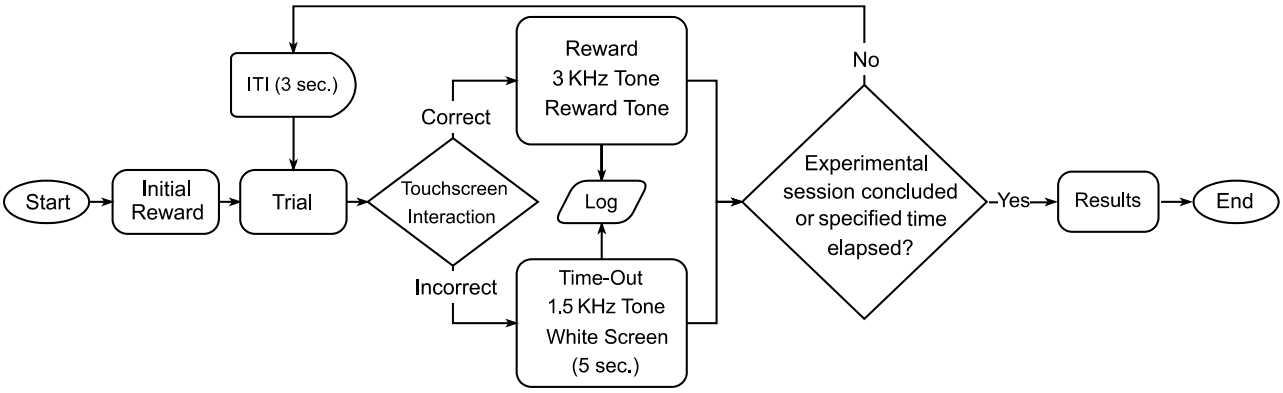

Figure 5: **Overview of a typical experimental session.** The flowchart represents the basic setup during every experimental session after the time-out is introduced during pre-training in boths tasks.

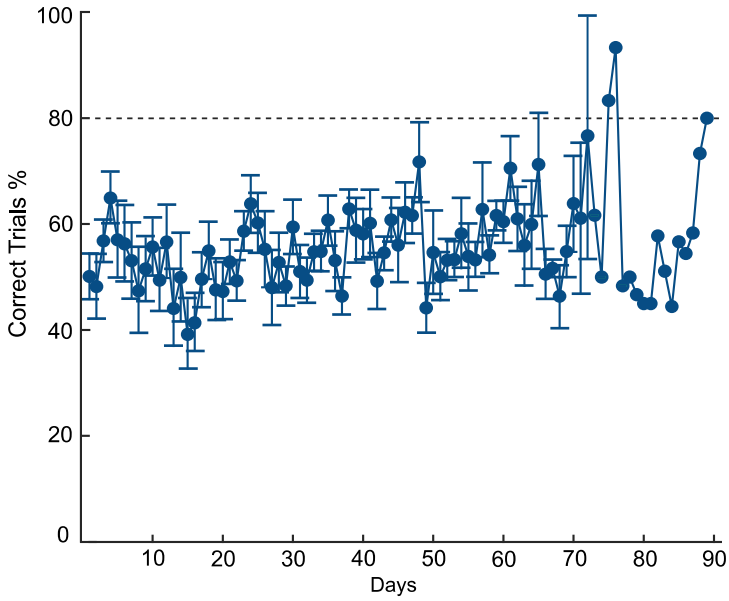

Figure 6: **Average learning curve of all animals during the acquisition stage and subsequent reversals (Acq. to R5) in a serial reversal visual discrimination task.** The dashed line represents the passing criterion of at least 80% correct trials for two consecutive sessions. Error bars indicate the standard error (SE) across days. As the task progresses, fewer animals remain in the behavioral testing, leading to the disappearance of error bars towards the end of the plot, as only one animal continued the task beyond this point.

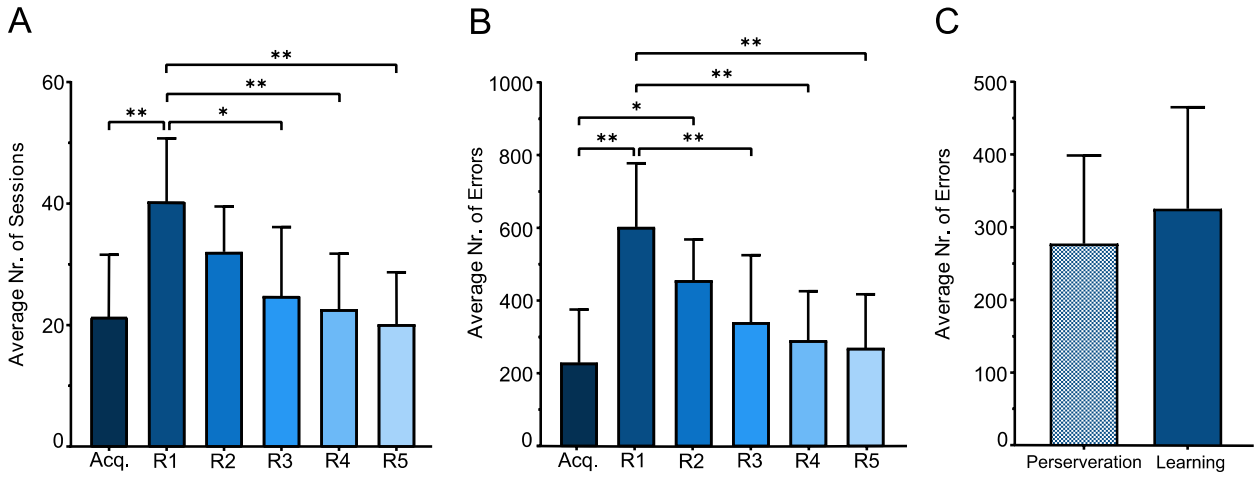

Figure 7: **Performance in the serial reversal visual discrimination task. A)** Average number of sessions across all learning stages. **B)** Average number of errors across all learning stages. **C)** Comparison between perseverance errors (sessions with ≤ 45% correct responses) and learning errors (errors in sessions with performance above 45%). Mean (M) ± SD in each learning stage. Statistical significance indicated as * $p < 0.05$, ** $p < .001$.

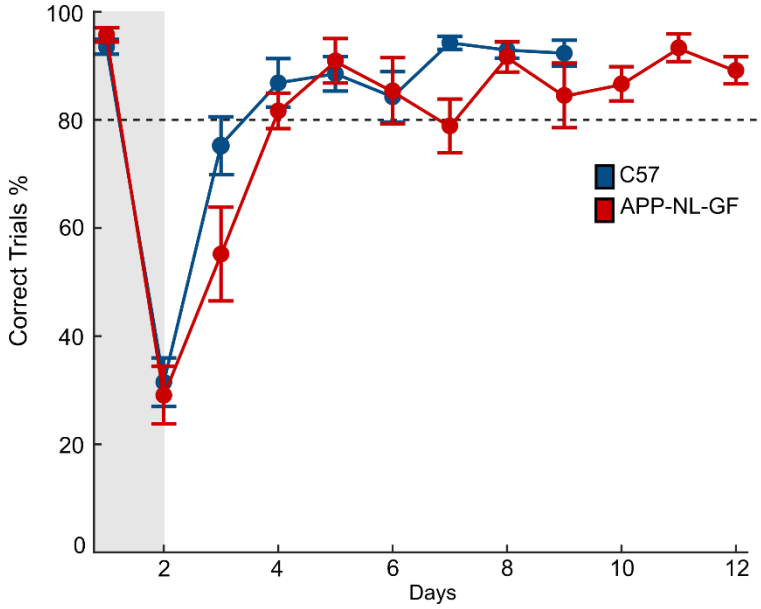

Figure 8: **Average learning curve of C57 (blue) and APP-NL-GF (red) mice during acquisition stage and subsequent reversals (Acq. to R5) in a serial reversal location task.** The dashed line represents the passing criterion of at least 80% correct trials for two consecutive sessions. Error bars indicate the standard error (SE) across days.

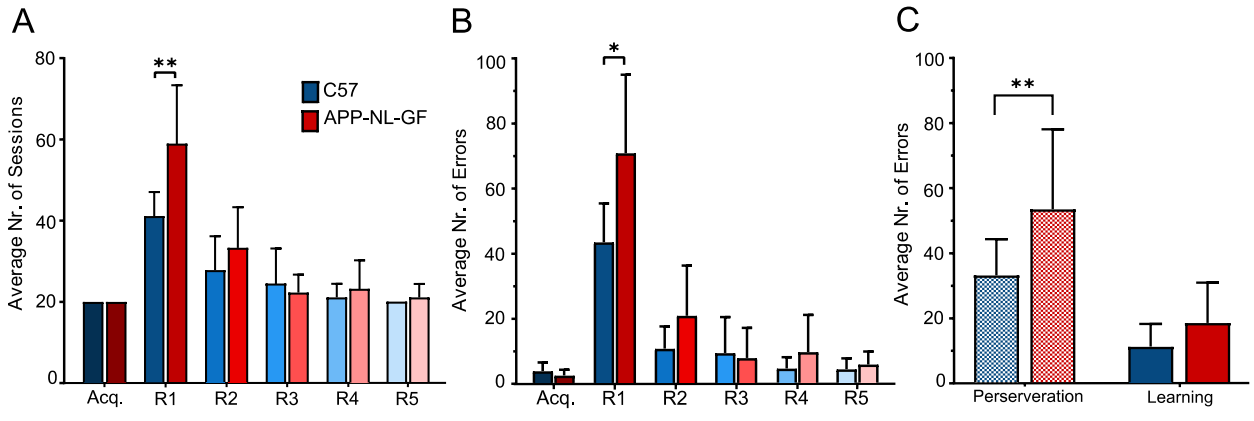

Figure 9: **Performance in the serial reversal location task**. **A)** Average number of sessions across all learning stages. **B)** Average number of errors across all learning stages. **C)** Comparison of perseverance errors and learning errors. Mean (M) ± SD in each learning stage. Statistical significance indicated as * $p < 0.05$, ** $p < .001$.