

DOI: 10.1017/psa.2022.63

This is a manuscript accepted for publication in *Philosophy of Science*.

This version may be subject to change during the production process.

Novel Tool Development and the Dynamics of Control: The Rodent Touchscreen Operant Chamber as a Case Study

Jacqueline A. Sullivan

Rotman Institute of Philosophy

Western University

7170 Western Interdisciplinary Research Building

London, Ontario CANADA N6A 5B8

The author would like to thank two anonymous referees for very helpful comments on an earlier draft of this paper, Dan Weiskopf, who organized the “Beyond Explanation: Exploring the Diversity of Neuroscientific Practice” symposium, and Felipe De Brigard, Philipp Haueis, and Pamela Reinagel for helpful comments at this session. The author is grateful to Lisa Saksida, Tim Bussey, and the Translational Cognitive Neuroscience Laboratory at UWO for informing her philosophical thinking about rodent touchscreens. This paper is based on the author’s interpretation of the published literature on rodent touchscreens. This work was supported by a grant from the Social Sciences and Humanities Research Council (SSHRC) of Canada.

In the quest to discover the neural bases of cognition, rigorous behavioral tools are equally as important as sophisticated tools for neural intervention. This paper evaluates several episodes in the development of a novel behavioral tool for rodent cognitive testing, the *rodent touchscreen operant chamber* (Bussey et al., 1994). Using conceptual tools on offer in the philosophical literature on exploratory experimentation and control, I illuminate how optimization of this behavioral tool and an understanding of the causal knowledge it may be used to generate historically has involved dynamic interplay between community-driven exploratory practices operating in parallel with causal hypothesis-driven research.

1. Introduction. A primary aim of cognitive neuroscience is to identify the neural bases of cognition. Behavioral tools used to produce, detect and measure learning and memory play a crucial role in the achievement of this aim, but to date scant philosophical attention has been directed at understanding what factors inform and shape their development. My aim in this paper is to contribute to filling this gap in the philosophical literature via a critical examination of several key historical episodes in the development and refinement of a behavioral apparatus for rodent cognitive testing, the *rodent touchscreen operant chamber*, first described by Timothy Bussey and colleagues in 1994. Using conceptual tools on offer in the philosophical literature on exploratory experimentation and control, I show how a community-driven program of exploratory inquiry to increase the appropriateness of this tool for use in causal hypothesis testing has operated in parallel with and in the context of hypotheses-driven research and has historically played a key constitutive role in shaping the knowledge-producing capacity of this tool.

I begin, in Section 2, by describing a set of conceptual distinctions on offer in the philosophical literature on exploratory experimentation and control. In Section 3, I appeal to these conceptual distinctions to evaluate several historical episodes in the development of the first rodent touchscreen. As I aim to show, exploratory experiments were initially instrumental in informing and shaping the development of a touchscreen apparatus that could be used to test causal hypotheses about the role of specific brain areas and neurotransmitters in visual discrimination learning tasks. However, there were aspects of the experimental arrangement that the developers of the apparatus had not considered and thus not examined that could potentially confound experimental tests of causal hypotheses it was used to run. As I explain in Section 4, potential confounds later came to light, and new questions about

the apparatus continue to emerge and shape its use, but crucially due to a program of community-driven exploratory inquiry to probe for confounds, increase experimental control, and optimize the apparatus and associated cognitive testing methods for diverse lines of hypothesis-driven research. I use my analysis of the dynamic evolution of rodent touchscreens to make the case that the process of identifying confounds and gaining experimental control over behavioral tools and understanding the kind of causal knowledge they may be used to generate requires on-going dynamic interplay between community-driven exploratory practices and causal hypothesis-driven research.

2. Exploratory Experiments, Confounds, and Control. Philosophical work on whether experiments in neuroscience generate knowledge has focused predominantly on experimental tests of causal hypotheses linking neural activity to behavior. Although this makes good sense given that most experiments in neuroscience aim to discriminate among competing causal hypotheses about phenomena of interest, the emphasis has served to obscure how knowledge production works in neuroscience and the fundamental role that exploratory practices (e.g., Steinle 1997) play in this endeavor. Relevant to my aims in this paper are the confound probing functions that Jutta Schickore (2019) attributes to exploratory experiments (EEs) at the frontiers of science.

Schickore (2019) evaluates Charles Darwin's research on digestion in insectivorous plants (e.g., bladderwort). As she explains, Darwin was interested in determining whether the presence of animal matter was necessary for protoplasm production inside plant bladders. Darwin used a solution to remove animal residue from a group of bladders and then ran a series of experiments exposing them to different substances (e.g., ammonia, urea). According to Schickore, these EEs served as *diagnostic probes* that enabled Darwin to gain an understanding

of “factors that may have an impact on (disrupt, change)” or *confound* “the relationship between the independent and the dependent variable in an experimental situation” (Schickore 2019, 210). They allowed Darwin to construct what Schickore (2019, 211) dubs a “confounder repertoire”, which she claims differs from Deborah Mayo’s (1996) “error repertoires” because rather than unearthing “canonical” kinds of experimental errors like mistaking correlations for causal relations or misidentifying artifacts as real effects, “diagnostic probes” are devised to detect confounders idiosyncratic to specific experimental contexts. Investigators may struggle to determine what they are and how and where to probe for them. For example, Darwin had to hypothesize what kinds of factors may impede his experiments, what kinds of substances to test to determine their impact on digestion and how best to clean bladders so no foreign substances remained.

Such exploratory work, however, does not end once potential confounders have been identified. Another type of EE, which Schickore dubs a “determinative probe” (Schickore 2019, 215), also must precede experimental tests of causal hypotheses. Investigators must determine whether potential confounders they have identified in the context of diagnostic probes have been adequately controlled in their experimental set-ups prior to running experimental tests of causal hypotheses. Darwin, for example, had to ensure in his experiments on protoplasm production that the experimental and control groups differed only with respect to the independent variable. The relevant notion of *control*, then, is the practice of ensuring that the effects of variables other than the independent variable on the variable of interest are minimized and, in the ideal case, completely absent from the experimental arrangement.

Importantly, Schickore emphasizes that diagnostic probes are rarely if ever exhaustive in unearthing potential confounds, leaving open the possibility that experimental tests of causal

hypotheses occurring on their heels will fail to be appropriately “controlled” or sufficiently “severe”. Yet this means correctives must be in place to ensure that experimental tests of causal hypotheses eventually evolve in ways that increase control or “severity” (Mayo 1991). Schickore proposes her concepts of diagnostic and determinative probes could inform peer review in scientific journals by requiring manuscript authors to identify explicitly the types of probes used in their research studies. Schickore’s conceptual distinction also may be used as a guide for determining where correctives are and/or should be in place in the scientific process to improve control over scientific tools. In what follows, I use a historical case study from rodent behavioral neuroscience to demonstrate that in some areas of neuroscience, diagnostic probing is a fundamental and routine part of the research landscape and culture when novel tools for testing causal hypotheses are first developed and as they come to be used more widely within a research community.

3. Rodent touchscreens and early diagnostic probes. Rodents are widely used in neuroscience because they are inexpensive to procure and maintain, are regarded as representative models for the mammalian brain and can be readily used in intervention experiments. Research in rodent behavioral neuroscience has historically involved the use of transient or permanent intervention techniques to alter brain activity and behavioral assessment tools to determine the effects of such interventions on behavior. In a typical experiment, an investigator aims to manipulate a single variable, such as neuronal activity in a single brain region, to determine the impact on behavioral performance in a task designed to assess a cognitive function. If differences in behavioral performance between the experimental and control groups are observed, the activity of neurons in the region that was the target of the intervention are interpreted as playing a causal role in behavioral performance on the task and in the psychological function the task is designed to

measure. To establish such causal relationships, intervention techniques must act upon a single causal variable and behavioral assessment tools must individuate a single psychological function.

Prior to the early 1990s, behavioral tools for assessing cognition in rodents were limited to classical and operant conditioning chambers and different types of mazes. Given that rodents were assumed to be capable of more complex and human relevant forms of cognition than could be assessed using these tools, researchers were in pursuit of novel behavioral techniques. At that time, a common approach for studying the neural correlates of cognitive functions in brain-damaged human patients and brain lesioned non-human primates was to use touchscreen-based tasks (e.g., Sahakian and Owen 1992) with 2-dimensional visual images presented on a computer screen. Subjects responded to visual images with a finger press to the screen and were provided feedback for correct and incorrect responses. Because rodents were thought to lack visual acuity to see and respond to computer-based stimuli, however, skepticism about the feasibility of developing a comparable touchscreen-based tool for assessing cognition in rodents was widespread.

Undeterred by such skepticism, in the early 1990s, Timothy Bussey and colleagues decided to place rats in a touchscreen operant chamber designed to investigate visual discrimination learning in pigeons to see what would happen and found that rats “could attend and respond to [the] computer graphic visual stimuli (Bussey, Saksida, & Wilkie, unpublished observations)” (Bussey et al. 1994). Bussey and his collaborators (Bussey et al. 1994) thus set out to build a species-specific apparatus for rats. In designing and constructing the apparatus (Figure 1), they needed to anticipate, explore, and test potential variables that might compromise its appropriateness for testing causal hypotheses about the role of specific brain areas in reward-based forms of visual discrimination learning. They engaged in “parametric analyses”—a type of

diagnostic probing in which different mutable components of an experimental set-up are systematically varied to determine their impact on the dependent variable (e.g., behavior). Mutable features of the original rodent touchscreen included the video-display unit (VDU) on which 2-D visual stimuli were presented and its features (e.g., touch or pressure sensitive surround to detect nose pokes), features of the visual stimuli (e.g., size, shape, color, luminance), type of rewards (e.g., sucrose pellets or chow), location of the reward-delivering food magazine, spatial arrangement of the apparatus and types of materials used for constructing it (e.g., wood, plexiglass) and features of those materials (e.g., sound attenuating, wipeable).

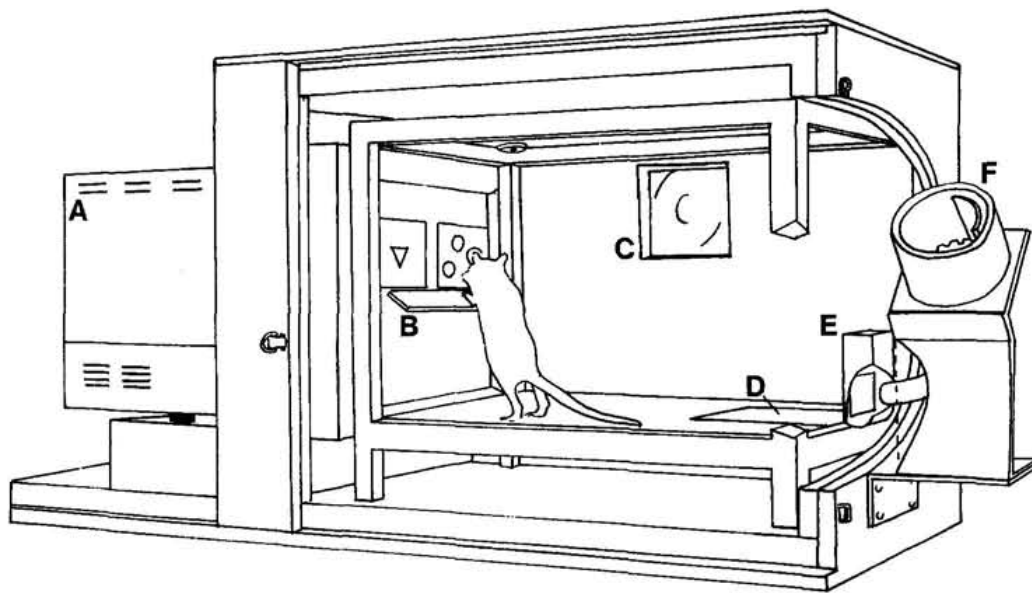


Figure 1. The First Rat Touchscreen Operant Apparatus. Lister-hooded rat rearing and placing paws on plexiglass shelf (B) facing stimuli presented on video display unit (A) where computer graphic stimuli are presented in windows created by black plexiglass mask. Reward magazine at back of chamber (E), pressure-sensitive floor panel (D), and wooden sound attenuating box (F). (Bussey et al. 1997). Reprinted with permission.

Given that the apparatus was developed for use with different computer automated learning tasks, it is relevant to note that each task and task protocol also had mutable features. For example, consider the delayed non-matching to sample (DNMS) task, which was used in the original study to test working memory in the touchscreen. In the original touchscreen version of this task, a rodent learned to select a different stimulus from one it has just seen. During the training condition, an image appeared in one of three locations on the touchscreen. When the rat nose-poked the image, it disappeared, and two images appeared in the other two windows of the touchscreen—the image just seen and a different image. If the rat selected the correct image, the magazine light went on, the tone sounded, and it received a food reward; incorrect responses received a 5-second time-out. The DNMS task had variable features in addition to those that were part of the touchscreen apparatus including, the location where the sample stimulus appeared on the screen, the number of training and testing trials and duration of training and testing sessions, the duration of intertrial intervals and the time between training and testing sessions.

Bussey and colleagues knew to build standard “canonical” *controls* into the apparatus (Figure 1) and cognitive testing method in response to an established repertoire of potential confounds arising from historical work on operant conditioning in rodents. For example, to ensure rats were not distracted by extraneous stimuli (e.g., sounds) during training and testing, the testing chamber, which was plexiglass, was placed in a “sound attenuating” wooden box and a ventilation fan served to “mask extraneous noise” (Bussey et al. 1994, 105). Rats were placed on a food restrictive diet prior to training and testing, a common strategy to ensure motivation to learn in reward-based learning experiments. Computerized automatization of touchscreen-based

tasks like DNMS served to eliminate investigator errors and biases and reduced investigator-animal interactions, which could have unintended impacts on the animal's behavior.

Other features of the apparatus and associated testing method resulted from parametric analyses testing different configurations of the apparatus while training rats in visual discrimination learning tasks. These “diagnostic probes” yielded novel information instrumental for improving the design of the apparatus for testing causal hypotheses. For example, sucrose reward pellets originally were delivered in food cups just below visual stimuli on the touchscreen. When Bussey and colleagues “observed that rats appear[ed] to orient toward the food cup rather than the stimuli”, making it difficult to determine “what”, if anything, was being learned, “[...] subsequent configurations of the apparatus delivered food reward” via a food magazine in “the rear of the chamber” (Bussey et al. 2008, 520) directly opposite the VDU.

Other parametric analyses were aimed at ensuring rats had full view of and were responding to all visual stimuli presented on the touchscreen and learning the intended stimulus relationships. These analyses prompted modifications to the design of the apparatus. For example, a pressure sensitive floor panel was installed in front of the food magazine directly opposite the VDU so that only when a rat was on the panel would the computer program initiate the next behavioral trial. The distance from the panel to the VDU was selected to ensure a rat would approach the screen “head on”, with all visual stimuli in full view. A plexiglass mask placed over the VDU, “allowing access to the stimuli only through windows just larger than the stimuli” (Bussey et al. 1994, 109) served to focus a rat's attention on those parts of the screen on which visual stimuli appeared. To ensure that a rodent's nose-poke response to a stimulus was neither impulsive nor inadvertent, a plexiglass shelf was placed directly underneath the stimulus windows, “forc[ing] the animal to stop in front of the stimuli, rear [on its hind legs], and stretch

toward a stimulus in order to make its response (Bussey et al. 1994, 109).” The infrared touchscreen was designed to detect nose-pokes to an image without a rat having to physically nose-press the screen—a response that when repeated could cause discomfort that may impede learning. As rats were known to exhibit spatial response biases, if a rat subject “developed a spatial response bias (defined as >70% responses to the left or right window within a session)” during training, the investigators (1994, 106) “[implemented] a correction procedure.”

Bussey and colleagues emerged triumphantly from these extensive parametric analyses with an innovative apparatus for testing causal hypotheses linking neural activity to behavior in rodents. The tool was also revolutionary insofar as it could be used to compare the performance of rodents with brain lesions and their brain-damaged human counterparts on similar kinds of psychological tasks, enabling cross-species comparisons linking neural activity to cognition and behavior. Yet, as I explain in the next section, during the next two decades after its introduction in the literature and the development of a similar apparatus for mice, concerns related to whether rodent touchscreens could be used to run adequately controlled tests of causal hypotheses linking neural activity to behavior began to emerge. As the community of rodent touchscreen researchers began to grow (Dumont et al. 2021), hypothesis-driven experiments aimed at improving the knowledge-producing capacity of rodent touchscreens began to be undertaken in parallel with experiments using the tool to test causal hypotheses linking neural activity to behavior.

4. Evolution of Rodent Touchscreens. During the next decade, Bussey and colleagues began using the rat touchscreen operant chamber for hypothesis-driven experiments designed to determine the role of specific brain areas in different types of visual-based learning. One set of experiments compared the performance of rats with quinolinic acid-induced lesions in anterior and posterior cingulate and medial frontal cortices and “sham” (operated on surgically but not

lesioned) rats on visual discrimination learning tasks (e.g., Bussey et al. 1997). In another study, Arjun Sahgal and Thomas Steckler (1995) used a similar experimental arrangement (“TouchWindows”) to determine the impact of intraperitoneal amphetamine and serotonin agonist injections on performance in a task designed to assess visual attention.

By 2008, several other research groups (34 studies in total) had used rodent touchscreen operant chambers in some form to study behavior and/or test causal hypotheses linking brain lesions or pharmacological manipulations to behavior (See Bussey et al., 2008; Dumont et al. 2021)). In part because the overall configuration of the rat touchscreen apparatus was not standardized across laboratories, with different researchers using different (a) types of touchscreens (b) types of food reward, (c) placements of reward magazines in the chamber, (d) types of visual images (e) types of learning tasks, (f) types of training and testing protocols, and (g) types of rat strains, novel questions began to emerge about additional variables that had not been previously considered that may confound tests of causal hypotheses the apparatus was used to run. In the context of causal hypothesis-driven experiments, researchers engaged in probative work, analyzing the behavior of their rodent subjects during training and testing and entertaining the possibility that observed changes in behavior that were taken as indicative of forms of visual discrimination learning may be due instead to other factors including: biases for certain visual stimuli, poor visual acuity (making it difficult to assess the kind of stimulus relationships being learned) and differences in intensity between stimuli used in the context of the same experiment in which a choice between stimuli had to be made.

Bussey and colleagues (2008) engaged in a set of hypothesis-driven experiments to probe for these potential confounders, to systematically investigate those parameters that had come to be recognized as potentially impacting learning in the apparatus with an eye towards

“optimizing” the touchscreen cognitive testing method for causal hypothesis-testing (Bussey et al., 2008). To take one representative example from this study, although it was known that rats exhibited stimulus preferences in experimental contexts in which 3-dimensional objects were used, whether they preferred some 2-dimensional touchscreen computerized visual stimuli over others had not been previously explored. If rats did have stimulus biases, an investigator would be unable to determine if changes in behavior following training were due to bona fide changes in cognition or merely preferences for certain stimuli. In intervention experiments using drugs or lesions to manipulate brain activity, then, she also would be unable to determine if the observed effects on behavioral performance in the apparatus were due to the intervention impacting cognition or instead “affecting [...] the expression of a stimulus bias” (Bussey et al. 2008, 517).

In one experiment to probe for stimulus biases, Bussey and colleagues (2008) selected six pairs of then standard touchscreen-based visual stimuli (e.g., plane-flower; spider-scissors) and presented them side-by-side in pairs in a single learning session to Lister hooded male rats. Rats had to select one of two stimuli with a nose-poke; one stimulus was rewarded with a sucrose pellet; the other was unrewarded. To determine stimulus bias, they compared each rat’s performance on this first session to chance (50% correct). They determined that rats expressed a bias that was significantly different from chance with respect to one stimulus (the ribbon) in one pair of stimuli tested (ribbon-cross). This experiment served two functions. First, it identified a bias that could be eliminated by removing the bias-producing stimulus from the bank of touchscreen-based stimuli to be used in future experiments. Second, it specified a method for detecting spontaneous stimulus biases that might disrupt assessments of visual discrimination learning in the apparatus that could be used in future experiments aimed at establishing causal links between brain and behavior.

Yet, exploratory research to probe for previously unconceived potential confounds related to the touchscreen cognitive testing method for rodents did not end with Bussey and colleagues' 2008 paper. Rodent touchscreen technology and associated tasks and testing methods have continued to evolve and been improved over the past two decades. In 2001, Bussey, Lisa Saksida and Lawrence Rothblatt described a novel touchscreen operant apparatus to assess learning in mice using touchscreen-based tasks. With the introduction of cutting-edge technologies for creating genetically altered mice and for intervening in neural activity in awake behaving mice (e.g., chemogenetics, optogenetics), the past two decades have witnessed a steady increase in the use of rodent touchscreens and with it a growing community of researchers using them not only to test causal hypotheses, but also working collaboratively to probe for confounders, improve experimental control over the behavioral apparatuses and associated testing methods, and gain a better understanding of the kind of causal knowledge they can be used to generate. I have space to consider only a few examples of this type of work here (But see Dumont et al. 2021; Sullivan et al. 2021 bibliographies).

Some hypothesis-driven studies aimed at diagnostic probing have sought to determine whether mice exhibit preferences for specific stimuli or stimulus locations in touchscreen-based tasks (e.g., Delotterie et al. 2014) or whether they have biases for specific images or their aspects (e.g., Belarde et al. 2021). As new learning tasks have been proposed for use in touchscreens, probative work also has shaped their development and optimization (e.g., Delotterie et al. 2014). Another experimental study sought to assess whether different appetitive rewards used in mouse touchscreens (e.g., super-saccharin or strawberry milkshake) correlate with differences in task performance. That some rewards are more reinforcing than others may explain differences in behavioral results across labs using different reinforcers (e.g., Phillips et al. 2017).

In another example of a hypothesis-driven experiment aimed at diagnostically probing mouse touchscreens, Helene Richter and colleagues, who previously used touchscreens to assess cognitive flexibility in mice, sought to determine whether touchscreen training, which can “last several weeks until months” with “animals undergo[ing] daily training sessions of durations up to one hour”, causes a stress response detectable in the “physiology and behavior” of mice (Krakenberg et al. 2021). In their study, compared to control mice, they found that mice trained in a paradigmatic version of the touchscreen Visuomotor Conditional Learning task exhibited significant increases in anxiety-like behavior and had increased levels of corticosterone metabolites in their fecal matter. Given these findings, which they took to indicate “profound effects” of touchscreen training on the physiology and behavior of their mice, they suggest that future “hypothesis-driven research” should be directed at clarifying “the influence of touchscreen training on the affective state of mice” and “disentangl[ing] the effects of different aspects of touchscreen training procedures” on behavior (Krakenberg et al. 2021).

This brief analysis of some of some stages in the evolution and history of rodent touchscreen operant chambers reveals some important features of the dynamics of behavioral tool development. Exploratory experiments serve as *initial probes* to identify confounders and are fundamental for getting hypothesis-driven experiments linking the brain and behavior off the ground. However, these initial probes are rarely if ever exhaustive, nor can they be, because the experimental set-ups and the contexts in which they occur have many moving parts, and even when investigators due their due diligence to itemize and control for confounders, recognition of other possible confounders emerges only with time as the behavioral tool is used more regularly and more widely, as variables not previously considered are investigated and as conflicting

results within or across research groups prompt further probative work to improve the tool for causal hypothesis-testing.

One implication of this, then, is that hypothesis-driven research using a behavioral tool serves as an important context for looking for and thinking about potential confounders that may later become the focus of hypothesis-driven experiments to probe for those very confounders, which may then yield findings fundamental for improving the behavioral tool and its appropriateness for testing causal hypotheses or shedding new light on the precise kind of causal knowledge the tool may provide.

5. Conclusion. Hypothesis-driven research in science is traditionally regarded as epistemically superior to exploratory research (e.g., Aragon 2011; Haufe 2013). Yet the analysis of several episodes in the history of rodent touchscreens is suggestive that in some areas of science, exploratory research, particularly that which involves hypothesis testing to probe for confounders and improve experimental tools—is regarded as being on an epistemic par with hypothesis-driven research. Indeed, rodent behavioral researchers are somewhat unique with respect to the critical scrutiny to which they subject their behavioral tools while simultaneously engaged in hypothesis-testing using them (e.g., See Nelson 2018 on the elevated plus maze; Sullivan 2020). The exploratory hypothesis-driven work prompted by such critical scrutiny is important not only for improving available behavioral assessment tools for hypothesis-driven research but also for the fundamental role it plays in putting causal knowledge arrived at in causal hypothesis-driven experiments into a broader epistemic context.

References

- Aragon, Richard. 2013. "Thinking Outside the Box: Fostering Innovation and Non-Hypothesis-Driven Research at NIH". *Science Translational Medicine* 3(70):cn5.
- Belarde, James, Chen, Claire, Rafikian, Elizabeth, Yang, Mu, Troy, Carol. "Optimizing Touchscreen Measures of Rodent Cognition By Eliminating Image Bias".
bioRxiv 2021.04.05.438342; doi: <https://doi.org/10.1101/2021.04.05.438342>
- Bussey, Timothy, Muir, Janice and Trevor Robbins. 1994. "A Novel Automated Touchscreen Procedure for Assessing Learning in the Rat Using Computer Graphic Stimuli",
Neuroscience Research Communications 15(2): 103-110.
- Bussey, Timothy, Muir, Janice, Everitt, Barry, and Trevor Robbins. 1997. "Triple Dissociation of Anterior Cingulate, Posterior Cingulate, and Medial Frontal Cortices on Visual Discrimination Tasks Using a Touchscreen Testing Procedure for the Rat", *Behavioral Neuroscience* 111(5): 920-936.
- Bussey, Timothy, Padain, Tina, Skillings, Elizabeth, Winters, Boyer, Morton, A. Jennifer, and Lisa Saksida. 2008. "The Touchscreen Cognitive Testing Methods for Rodents: How to Get the Best Out of Your Rat." *Learning and Memory*, 516-523.
- Bussey, Timothy, Saksida, Lisa, & Rothblat, L. 2001. Discrimination of computer-graphic stimuli by mice: A method for the behavioral characterization of transgenic and gene knockout models. *Behavioral Neuroscience*, 115(4), 957–960.
- Delotterie, David, Mathis, Chantal, Cassel, Jean-Christophe, Dorner-Ciossek, Cornelia, Marti Anelise (2014). Optimization of Touchscreen-Based Behavioral Paradigms in Mice: Implications for Building a Battery of Tasks Taxing Learning and Memory Functions. *PLoS ONE* 9(6): e100817.

- Dumont, Julie, Salewski, Ryan & Flavio Beraldo. 2021. "Critical Mass: The Rise of a Touchscreen Technology Community for Rodent Cognitive Testing". *Genes, Brain and Behavior* 20 (1): e12650.
- Haufe, Chris. 2013. Why do Funding Agencies Favor Hypothesis Testing? *Studies HPS A 4* (3), 363-374
- Krakenberg, Viktoria, Wewer, Maximilian, Plame, Rupert, Kaiser, Sylvia, Sachser, Norbert, Richter, S. Helene (2021). Regular Touchscreen Training Affects Faecal Corticosterone Metabolites and Anxiety-like Behavior in Mice." *Behavioral Brain Research* 401. 113080.
- Mayo, Deborah. 1996. *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.
- Nelson, Nicole. 2018. *Model Behavior: Animal Experiments, Complexity and the Genetics of Psychiatric Disorders*. Chicago: University of Chicago Press.
- Phillips, Benjamin, Heath, Christopher, Ossowska, Zofia, Bussey, Timothy, Saksida, Lisa. 2017. "Optimisation of Cognitive Performance in Rodent Operant (Touchscreen) Testing: Evaluation and Effects of Reinforcer Strength." *Learning & Behavior*, 45(3):252-262.
- Sahgal, Arjun and Steckler, Thomas. 1994. TouchWindows and operant behaviour in rats. *Journal of Neuroscience Methods* **55**: 59–64.
- Sahakian, Barbara and Owen, Adrian. 1992. "Computerized Assessment in Neuropsychiatry Using CANTAB: Discussion Paper." *Journal of Research in Social Medicine*, 85(7): 399-402.
- Schickore, Jutta. 2019. "The structure and function of experimental control in the life sciences." *Philosophy of Science*, 86(2), 203–218.
- Steinle, Frederich. (1997). "Entering new fields: Exploratory uses of experimentation." *Philosophy of Science*, 64, S65–S74.

Sullivan, Jacqueline. 2020. “Are There Model Behaviors for Model Organism Research?” *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 82

Sullivan, Jacqueline, Julie Dumont, Sara Memar, Miguel Skirzewski, Jinxia Wan, Maryam Mofrad, Hassam Zafar Ansari, Yulong Li, Lyle Muller, Vania Prado, Marco Prado, Lisa Saksida, Timothy Bussey. 2021. “New Frontiers in Translational Research: Touchscreens, Open Science, and The Mouse Translational Research Accelerator Platform (MouseTRAP).” *Genes, Brain and Behavior*.