Optimization of experimental procedures for the assessment of auditory cognition in non-human primates

Dissertation

To obtain the doctoral degree in Sensory-Motor Neuroscience

"Doctor rerum naturalium"

Of the Georg-August Universität Göttingen

In the doctoral program Göttingen Graduate School for Neuroscience, Biophysics and Molecular Biosciences (GGNB)

Of the Georg-August University School of Science (GAUSS)

Submitted by

Jorge Cabrera-Moreno

From Puebla, Mexico

Göttingen, 2022

Thesis Committee

Dr. Marcus Jeschke & Prof. Dr. Tobias Moser
Cognitive Hearing in Primates Laboratory, German Primate Center
Auditory Neuroscience and Optogenetics Laboratory, German Primate Center
Institute for Auditory Neuroscience and InnerEarLab, University Medical Center

Prof. Dr. Julia Fischer
Cognitive Ethology Laboratory, German Primate Center

Prof. Dr. Hansjörg Scherberger
Neurobiology Laboratory, German Primate Center

Members of the Examination Board

Referee: **Dr. Marcus Jeschke & Prof. Dr. Tobias Moser**
Cognitive Hearing in Primates Laboratory, German Primate Center
Auditory Neuroscience and Optogenetics Laboratory, German Primate Center
Institute for Auditory Neuroscience and InnerEarLab, University Medical Center

Co-referee: **Prof. Dr. Julia Fischer**
Cognitive Ethology Laboratory, German Primate Center

Further Members of the Examination Board

Prof. Dr. Hansjörg Scherberger
Neurobiology Laboratory, German Primate Center

Prof. Dr. Alexander Gail
Sensorimotor Laboratory, German Primate Center

Prof. Dr. Melanie Wilke
Cognitive Neurology Unit, University Medical Center Göttingen

Prof. Dr. Siegrid Löwel
Department of Systems Neuroscience, Georg-August Universität Göttingen

Date of examination: 28th of February 2023

Declaration

Hereby I declare that I have written this thesis independently and with no other aids and sources than quoted.

Göttingen, January 13, 2023

Jorge Cabrera-Moreno

Acknowledgments

It is clear to me that the work described in this dissertation is not the result of a single person effort but rather a collective labor of several people. The specific scientific contribution to each of the described projects is acknowledged in each chapter, yet, I would like to use this space to thank all those people whose help went beyond data collection.

I would like to express my profound gratitude to Marcus Jeschke and Tobias Moser for giving me the great opportunity to pursue my PhD in their lab. I especially would like to thank Marcus Jeschke for his guidance, enthusiasm, supervision, and support throughout these four years.

Sincere gratitude goes to my committee members Julia Fischer and Hansjörg Scherberger for providing essential advice across these four years. I also would like to thank Alexander Gail, Melanie Wilke, and Siegrid Löwel for serving as examination board in my thesis defense, listening to my presentation, and providing thoughtful criticism.

I also would like to thank my laboratory colleagues, without whom several aspects of laboratory work would have been definitely more challenging. Antonino Calapai for his mentorship and full-time willingness to provide support. Lukasz Jablonski, Lakshay Khurana, and Josep Cardona Audi, for their patience in providing friendly and enthusiastic support on everything that had to do with cables, which biologists certainly never learned during college time. Burak Bali, Fadhel El May, Sabina Nowakowska, Lea Schott, and all previously mentioned for making life in the laboratory more fun and enjoyable. Great thanks for the laughs and the good time spent as a team.

Deep appreciation goes to the three exceptional students I had the opportunity to supervise during these years: Jonas Grunenberg, Lena Jeanson, and Christin Korb. Thanks for the time dedicated to the projects and for the friendly collaboration.

A big part of the projects presented owes a deep gratitude to Victoria Müller and Tamara Meyer-Burhenne for handling and providing care to the animals.

Similarly to the veterinarians Tamara Becker, Annette Schrod, Karen Lampe, and Birgit Kamp for ensuring animals' health.

Administrative and technical assistance staff whose role certainly is the backbone of every laboratory on earth. Thanks to Nadine Dietrich and Patricia Räke-Kügler for all administrative and technical support. Likewise, thanks to Ludwig Ehrenreich and Gerhard Hoch for building all the electronic gadgets required for the experiments.

A heartfelt thank goes to Kristin Kaduk, Antonino Calapai, Hayley Silver, Riyoko Shibe, and Iona McEwan for proofreading and providing feedback on parts of this dissertation.

Growth and development during a PhD surpass the professional component. A big part of my personal growth owns immense gratitude to the people I got to know during my time in Göttingen who valued, encouraged, and supported me throughout this journey. Thanks to Natalia Gritiuc and Kristin Kaduk for the time together, the fantastic and endless conversations, and the emotional support. Special thanks to Martin Garro, Daniel Ramos, Stefani Diaz, for being my little Latin-American refuge in Göttingen. Thanks to Stephanie Hock, Inga von Freitag Lörinhoff, and Alice Lippi for all the adventures and the fun side of my PhD odyssey.

Deep thanks to Ivan Serapio Ramos, Arturo Gutierrez Fuentes, and Ximena Torres Hernandez for almost two decades of friendship, from which, sadly, the last five years had to be long-distance. Thanks for being there for me in difficult times and demonstrating that love and family come in different shapes.

Finally, my deepest gratitude goes to my family, who has never let me down on any decision I have made in my life. Thanks for all the support for my education, endless love, and unconditional understanding.

Table of Contents

List of abreviations

Abbreviation	Definition
2AC	Two-Alternative Choice Task
3AC	Three-Alternative Choice Task
AAV	Adeno-Associated Virus
ABR	Auditory Brain Response
AMC	Amplitude Modulated Contour
AUT	Automated Unsupervised Training
BW	Frequency Bandwith
CF	Center Frequency
Chn	Channel
ChR2	Channelrhodopsin-2
cTr	Complex Train
d’	d-Prime
dB	Decibels
DPOAE	Distortion Product Otoacoustic Emissions
DPZ	German Primate Center
ESB	Enhanced Shockburst
EYFP	Enhanced Yellow Flourecent Protein
hSyn	Human Synapsin Promoter
Hz	Herz
I2C	Inter-Integrated Circuit
ILD	Interaural Level Difference
IPI	Inter-Phrase Interval
ITD	Interaural Time Difference
ITI	Inter-Trial Inverval
LXBI	Long-Tailed Macaque Experimental Behavioral Instrument
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
MXBI	Marmoset Experimental Behavioral Instrument
NHP	Non-human Primate
NHR	Number of Harmonics
NPHR	Number of Phrases
PCB	Printed Circuit Board
PRT	Positive Reinforcement Training

RFID	Radio Frequency Identification
RPA	Relative-Phrase Amplitude
SPI	Serial Peripheral Interface
SPL	Sound Pressure Level
sTr	Simple Train
THR	Threshold
UART	Universal Asynchronous Receiver Transmitter
USB	Universal Serial Bus
voc	Marmoset Vocalization
WGTA	Wisconsin General Appartus

Chapter 1

Introduction

Abstract

Non-human primates rely on acoustic signals to navigate their physical and social environment, the processes of detecting, locating, identifying, and interpreting auditory information are therefore crucial to making sense of the sound world. To do so, peripheral structures and specialized brain regions engage in complex coordination to perform low-level perceptual processes (sensory information processing) and higher-level cognitive processes (symbolic information processing). Primate audition research has thus devoted much effort to addressing how individuals acquire, process, and use acoustic information to interact with their environment.

Historically, studying non-human primate cognition has been based on observing animals' adaptive actions toward given situations, with fieldwork evaluating animals' behavioral responses in their native environment, and laboratory research assessing animals' behavior in artificially created setups. Research has primarily been concentrated in laboratory experiments due to better control of the experimental conditions. Laboratory setups are also beneficial as they facilitate the implementation of neurophysiological recordings to investigate the physiological mechanisms underlying cognitive processes. Despite the benefits of laboratory research, it relies heavily on animal training to habituate animals to the experimental setups and to instruct them to reliably interpret the principles of the experimental task. Training non-human primates thus demand a high amount of human involvement, and is susceptible to biases introduced by trainers. Additionally, in cognitive neuroscience, animals are often physically restrained to facilitate the recording of neural activity, drastically reducing the variety of behaviors

depicted by the animals, limiting research to basic processes of perception and memory, and raising animal welfare concerns.

Technological development in the last decade has made it increasingly possible to automate the behavioral assay and thereby reduce human manual labor, improve reproducibility, broaden the variety of behaviors assessed, and scale up data collection. However, auditory cognition research, specifically in non-human primates, lags in comparison to the sophisticated protocols available for visual cognition or even other species like rodents where automation has already proven a substantial contribution toward optimization.

This dissertation discloses in the following chapter a series of experiments aming to overcome the aforementioned methodological limitations in auditory cognition in non-human primates through automating training and testing protocols of a wide variety of perceptual and cognitive processes relying on a touchscreen-base experimental device directly attached to the home enclosure of socially housed common marmosets (*Callithrix jacchus*) and long-tailed macaques (*Macaca fascicularis*).

The chapter outline' favors topic coherence over the chronological order of experimentation. Chapter 1 aims to serve as a general introduction to provide a background overview as well as a description of the current status in non-human primate auditory cognition research.

1.1 Animal behavior

One of the most remarkable phenomena within animal life that distinguishes it from other living organisms is the wide variety of responses they depict as a consequence of internal and/or external stimuli (own state, conspecifics, members of different species, and environment). Animal behavior can be defined as everything animals do, or more accurately, what is observed by another organism to be doing (Skinner, 1938). This includes movements or the absence of them and underlying mental processes. Human (*Homo sapiens*) interest in animal behavior can probably be extended hundreds of thousands of years back when our ancestors needed to understand the behavior of animals for survival. Hunting, domestication, or escaping from predators required a great insight into animal behavior. Even today, the subject of animal behavior still has great practical significance. Crop production, for example, relies heavily on understanding pollination foraging behavior, where understanding transition rates between flowers is key to maximizing food production (Gagic et al., 2021). Likewise, understanding the foraging and nesting behavior of some mammals (Red fox – *Vulpes vulpes*), which are disease vectors, is essential to develop proper strategies for wildlife management to prevent human infections (Hodžić et al., 2015). Ultimately, studying the similarities and differences across human and non-human animal behaviors can give us insights into human cognition's developmental and evolutionary processes (Beran et al., 2014; Chittka et al., 2012; Roitblat, 1987; Wasserman et al., 2006; Zentall and Wasserman, 2012).

History and basic concepts

Like many research fields in biology, animal behavior has its foundations in the evolution theory by natural selection of Charles Darwin, presented in "The Origin of Species" in 1859 (Darwin, 2004). In his theory, Darwin proposed that the process of natural selection is an inevitable result of the variations among individuals of the same species, the inheritance of these variations, and the differences in reproduction success among individuals. As such, the specific set of characteristics that constitute an organism result from its ancestors' attributes that helped them to better survive and reproduce in their current environment more effectively. Since then, behavioral biologists have recognized that the behavior of animals and their anatomical traits are adaptations shaped by their natural history that have helped the animal acquire energy to survive and ultimately produce offspring (Pontzer, 2015).

After Darwin's natural selection theory, the study of animal behavior evolved into a new science, ethology, which sought to understand behavior from a physiological and psychological perspective rather than behavior's biological and evolutive relevance. In the 19th century, Ivan Pavlov, John B. Watson, Edward Tolman, Karl Lashley, Burrhus .F. Skinner, Wolfgang Köhler, and Robert Yerkes, among others, studied the behavior of laboratory dogs, rats, birds, and NHP intending to understand processes such as learning (Kohler, 2018; Lashley, 1951; Pavlov, 1928; Skinner, 1965, 1938; Tolman, 1973; Watson, 1913; Watson and Rayner, 1920; Yerkes and Dodson, 1908).

In the 1930s, field biologist Konrad Lorenz, together with Nikolaas Tinbergen, Karl von Frisch, and William Morton Wheeler stressed the importance of direct observation of animals in their natural habitat to understand behavior from a biological perspective (Lorenz, 1937, 1950; Tinbergen, 1963; Von Frisch, 2013; Wheeler, 2015). Nikolaas Tinbergen argued that to have a comprehensive understanding of any behavior, four "levels of analysis" must be applied: Causation – what causes the behavior?; Ontogeny – how does the behavior develop?; Function – how does the behavior affects fitness?; and Phylogeny – how does the behavior evolve? Since each of these levels of analysis focuses on different characteristics of behavior, several scientific fields emerged which contribute to understanding the behavior of animals (including humans).

What causes behavior?

Causation is a central concept of the behavioral assessment paradigm. For Tinbergen, causation referred to the immediate effects that, both external or internal, factors have on the occurrence of behavior (Lorenz, 1937; Tinbergen, 1963). To approach this question, it is ubiquitous to identify the physiological and cognitive processes that underlie a given behavior. Since the nervous system is taken as the foundation of animal behavior, at this level, behavior is understood as an output of neuronal processing given internal or external causes.

Motor-sensory-motor mechanisms

Living creatures need to solve the problem of staying alive and eventually reproduce. For this, they need to obtain energy, seek shelter, and avoid being the prey. In order to approach these challenges, animals must gather information, process it, and act upon it.

Acquiring information

Animals gather information from the world relying on a specific set of organs that allow them to transduce a particular type of energy into electrical signals that the nervous system can process. Generally speaking, there are three modalities of sensors that can be classified depending on the type of stimulus they transduce: mechanoreceptors (mechanical stress), chemoreceptors (chemical compounds), and photoreceptors (light) (Purves and Williams, 2001). Even though animals of different species might possess sensors from the same modality, the capabilities of these sensors to obtain information from the environment are specialized to the animals' expected ecological requirements. Consequently, animals that share the same environment do not necessarily sense the world in the same way. In 1957 the German biologist Jakob von Uexküll coined the word *Umwelt* (Environment, in the German language) to describe an organism's unique sensory world (meaningful environment) and differentiated it from the *Umgebung* (physical surroundings) (Canguilhem, 2022; von Uexküll and O'Neil, 2013). For example, humans (*Homo sapiens*) and mosquitoes from tropical regions (*Aedes aegypti*) share a similar *Umgebung* but different *Umwelts* (or Umwelten in proper German). While humans' hearing range goes from 20 to 20 000 Hz (Fay, 2012), the male mosquito's hearing organ detects sounds from 150 to 400 Hz. This apparent narrower sensitivity to sound serves the mosquito sufficiently in order to detect the presence of a female up to 10 meters away (Menda et al., 2019).

Processing information

As with the sensory system, the neural mechanisms that process information are a product of evolution and therefore limited to processing a narrow part of the information received from the sensory organs, usually associated with the ecological challenges that animals face in their environments. The processing of information may happen at the peripheral level of the nervous system, like in reflexes, where a nearly immediate movement is coordinated by a reflex arc pathway (Purves and Williams, 2001) or involve a more complex communication between different areas of the central nervous system (not necessarily relying on sensory input to take place but instead on the animal's internal physiological state) to coordinate behavioral states such as mating or foraging (Ji et al., 2021; Jung et al., 2020).

Responding

Once the information is processed, a specific response to solve a particular problem is deployed, called behavior. Ultimately, controlling an animal's movements involves a complex coordination of neural control to either respond with a simple sensory reflex such as the startle reflex, a rhythmic pattern like walking, or a whole state like foraging. Reacting appropriately, accurately, and flexibly to the different challenges along the lifespan of an animal warrants its survival and reproduction (Hunt, 2018).

Cognitive mechanisms

The study of mental organization for information processing can also explain behavior. For example, how animals, which are seen as goal-seeking agents, acquire, store, retrieve, and internally process information at different levels of cognitive complexity (Hebb and Donderi, 2013; Menzel and Fischer, 2011). In this framework, the animals' internal representation of the world is a crucial part of cognition because behavior is understood as the output of the information processing that takes place based on these representations. Consider, for example, the western scrub jays (*Aphelocoma californica*) that store different types of food (seeds, insects, or fruits) in dispersed caches to then retrieve it in the future in order to consume it (Clayton, 2006; Clayton et al., 2007, 2003; Dally et al., 2006; Raby et al., 2007). How do they decide what and where a piece of food should be cached? How do they judge the right time for retrieving a particular food? How do they face conflicting conditions, such as being watched by a conspecific during the caching episode? How do they remember what was cached and where it was cached? Thus, numerous processes of perception, attention, learning, memory, and decision-making activity underlie the scrub jay food caching and recovery behavior. The study of these perceptual and central mechanisms in animals begins by asking what kinds of cognitive structures each species or individual possesses, how they are activated, what their function is, and how they have evolved (Hogan, 2005).

Several of these questions have been addressed with respect to NHPs (in the wild and in the laboratory) due to the common interest in comparing the cognitive capabilities of humans with those species considered the most closely related to us to understand the evolution of the human mind (Cheney and Seyfarth, 1990; Povinelli et al., 2000; Tomasello, 2000).

1.2 Cognitive mechanisms in NHPs

Humans are primates, and together with bonobos (*Pan paniscus*), chimpanzees (*Pan troglodytes*), gorillas (*Gorilla beringei* and *Gorilla gorilla*), and orangutans (*Pongo pygmaeus*, *Pongo abelii*, and *Pongo tapanuliensis*) belong to the great ape or Hominidae family. Great apes are closely related to the Old-world monkeys (Family Cercopithecidae – common ancestor ~25 Ma), followed by New World monkeys (Families Callitrichidae, Cebidae, Aotidae, Pitheciidae, and Atelidae – common ancestor ~40 Ma), and finally, the Prosimians (Family Tarsiidae and Suborder Strepsiirhini – common ancestor ~80 Ma) (Brandon-Jones et al., 2004; Glazko, 2003; Perez et al., 2013; Rose, 2006; Shumaker and Beck, 2003; Wakeley, 2008; Wilson and Reeder, 2005). Although the study of the cognitive mechanisms that underlie behaviors for each of these species has its value on its own, those phylogenetically closer to humans are often studied under the assumption that the knowledge regarding their cognitive mechanisms would serve as a better comparison to understanding human cognition and behavior evolution.

The study of primate cognition is usually divided into physical and social cognition. On the one hand, studies on physical cognition focus on the mechanisms related to the interaction with the physical world, including foraging and defense skills, learning, and problem-solving. On one hand, object manipulation, tool use, limits of perception, categorization, numerosity, gratification delay, planning, memory, and metacognition are some of the mechanisms studied by physical cognition. On the other hand, social cognition investigates the mechanisms that allow primates to recognize themselves in relationship to others, as well as others' relationships, make judgments based on that knowledge, and use past experiences to anticipate the future (Byrne, 1995; Parrish and Brosnan, 2012; Seed and Tomasello, 2010; Shettleworth, 2009). Cooperation, decision-making, social learning, communication, deception, and theory of mind, are some of the mechanisms studied by social cognition.

Methodological approaches

Studies in primate cognition tend to be either conducted in the laboratory with captive colonies or with wild animals in the field. Fieldwork usually relies on systematic observations of colonies for long periods of time, where animals are studied facing challenges they naturally encounter in nature or by challenges designed and presented by the experimenter (Goodall, 1964; Menzel, 1991; Visalberghi et al., 2009). Studying animal cognition in their natural habitats is likely more productive when the questions addressed correspond to the species' ecology. However, challenges such as the experiments' replicability due to the impossibility of controlling relevant factors, the significant time needed to make those observations (exceptional expertise and experience are required), or the limited variety of methods that have been used have been highly questioned over the past decades (Pritchard et al., 2016; Ryan, 2011; Tomasello and Call, 2011).

Laboratory experiments rely on controlling the environment (e.g., food and water availability, day cycle, social contact) and internal factors (e.g., thirstiness, hunger, or social drive) of the animal to influence the motivation and performance at a given time. Therefore, experimenters in the laboratory usually bring the animals into carefully designed setups where sensory (e.g., luminescence, sound loudness) and social (e.g., number of conspecifics in the room) disturbances are controlled. Here the animals are presented with problems that have to be solved by relying on specific tools provided by the experimenter in order for the experimenter to easily observe the desired behavior then replicate it even with naive experimenters. Additionally, due to the well-designed setups, cognitive research in the laboratory also offers the possibility to investigate the physiological mechanisms that underly specific cognitive processes by registering the neural activity of the animal while performing a given task. However, a big concern that researchers usually raise in experiments in the laboratory is the low or sometimes null similarity that these setups have with the animal's natural environment and, therefore, whether the measured cognitive capabilities have any ecological validity. Additionally, variability across colonies, constraints in the low number of available individuals per colony, and the time required to train an animal to perform the desired task make the collection challenging and the findings difficult to generalize, even across captive colonies (Tomasello and Call, 2011).

Due to the influx of technology, the trade-off problem between tight control of individual conditions and the more natural settings has been slowly mitigated in the last decade. Current efforts are being made to improve the accuracy of measurements in the field (e.g., remote sensing devices, satellite tracking, machine learning algorithms for animal identification) and to develop tasks and settings in the laboratory that are ecologically valid for the species (e.g., group testing, and more naturalistic setups) (Bala et al., 2020; Hayden et al., 2021; Jacob et al., 2021; Nourizonoz et al., 2020; Schofield et al., 2019). Cognitive research in NHPs is a collaborative endeavor that requires collaboration between the two approaches.

Consequently, part of the challenge to better understand the processes of primate cognition lies in the good practice of designing and performing more standardized experiments that facilitate replicability while stressing the importance of sharing data among the community, given that research groups will always be constrained to a few species with a limited number of animals available for testing over time (Cauchoix et al., 2017; Hurlbert, 1984; Palmer, 2000; Primates et al., 2019; Ryan, 2011; Tomasello and Call, 2011).

Challenges in the laboratory

Keeping NHPs in captivity comes with a significant number of challenges requiring a synergic collaborative effort between several experts to overcome (e.g., veterinarians, caretakers, and scientists). The first line of challenges relates to housing and husbandry. Because of the close phylogenetic relationship to humans, NHPs hold a special status compared to other non-primate animals and, therefore, need to be provided with additional physical and psychological care to warranty their well-being (Jennings et al., 2009; Kirkwood and Hubrecht, 2010; Rennie and Buchanan-Smith, 2006a, 2006b; Röder and Timmermans, 2002; Wolfensohn and Honess, 2008). Since the welfare of animals kept in captivity is of high importance for humanitarian, scientific, and economic reasons, strict regulation of the housing conditions and experiments with NHPs come under legal regulation for every country (e.g., in Europe – European Directive 2010/63/EU, in the United States of America – Animal welfare Act 1996 S.1378).

The second line of challenges refers to the experimental procedures. Despite the wide variety of experiments that can be performed in NHPs in the laboratory, most of them require the animals to tolerate specific circumstances arising from the experimental needs (e.g., human interaction, handling, transportation, social separation, use of experimental tools, remain still in a chair, etc.). To reduce stress during experimental procedures, researchers rely on training techniques to habituate the animals to cooperate during experimentation. The training of NHPs is essential not just for scientific purposes but also for husbandry and veterinary procedures. For these reasons, training in NHPs is considered good practice and recommended in several legislations and professional guidelines (Animal Welfare Institute (Reinhardt, 1997), National Centre for the Replacement Refinement & Reduction of Animals in Research (NC3Rs) (Prescott and Buchanan-Smith, 2003) Directive of the European Union – 2008/0211/COD).

Training NHPs

Positive reinforcement training (PRT) has been one of the most frequent techniques to train NHPs to cooperate voluntarily in scientific, veterinarian, and husbandry procedures. PRT is a type of operant conditioning where the animals are rewarded for performing the desired behavior (Skinner, 2005, 1938). In this type of training, the trainer ignores unwanted behaviors while focusing on the correct ones. The behaviors of interest are rewarded every time they appear and thereby reinforced. As a result, the animal is prone to show the desired behaviors in exchange for a reward over time. However, while relatively simple behaviors such as those needed for husbandry and veterinary procedures (e.g., blood sampling, routine physical check-ups, transportation) can be relatively easy to accomplish by expert staff, more complex behaviors like those needed for assessing cognitive capabilities, drastically increase human time investment (Laule et al., 2003; Schapiro et al., 2005). Additionally, manual training is susceptible to biases introduced by the trainers. This source of variability in training histories is of particular importance in experimental procedures where comparisons across animals' cognitive performance might lead to limited or misleading interpretations of the behavior and neurophysiological data due to mismatching in the task-solving strategies of individual animals (Farrar et al., 2021; Tomasello and Call, 2011).

Finally, the training and experimental success heavily rely on matching the animal's internal state (motivation) with the time chosen by the trainer. The confined number of working hours per trainer and the number of animals that need to be trained per day make the ideal conditions challenging, which consequently reduces the efficacy of training (Balleine, 1992; Colgan, 2013; Dickinson and Balleine, 1994; Ward et al., 2015).

Testing batteries - A small history of optimization

For many decades the Wisconsin General Test Apparatus (WGTA), a manually operated device that presented experimental subjects with discriminative stimuli while the researcher observed the responses through a one-way screen, allowed the study NHP learning (Harlow and Bromer, 1938). With the advances in technological complexity and affordability, numerous modifications to the WGTA, such as the implementation of electrical sensors to register responses (Sidowski and Spears, 1970), motors to control the movement of panels and trays (Davenport et al., 1970; Schrier, 1961) and later computers for automatic data collection (Polidora and Main, 1963), were some of the first attempts to optimize the system. The time savings and the gains in precision and reliability soon made computer technology an essential part of NHP cognition assessment in the laboratory.

The improvement of display technology enabled scientists to implement video-tasks and benefited from the relatively unlimited variety of sophisticated stimuli improving the flexibility and complexity of the task design. Although several species of NHP learned to interact with computer-generated targets presented in the screen through the manipulation of joysticks, buttons, or levers, to obtain a food reward (Perdue et al., 2018; Richardson et al., 1990; Roberts et al., 1988; Rumbaugh et al., 1989a, 1989b; Savage-Rumbaugh, 1986; Vauclair and Fagot, 1993; Washburn et al., 1989; Washburn and Rumbaugh, 1992) the spatial separation between the controls, monitors and location of reward delivery represented a mayor challenge for other species such as squirrel monkeys (*Saimiri*) (Andrews, 1993). Subsequently, the development of touchscreen technology permitted design tasks where the target and the response could be placed in the same physical area for the first time, solving the problem of stimulus-response spatial discontinuity (Washburn et al., 1989). Video tasks, together with the new touchscreen technology, gained popularity among NHP research in the subsequent years (Roberts et al., 1988; Crofts et al., 1999; Dias et al., 1996; Pearce et al., 1998; Spinelli et al., 2004; Weed et al., 2008) and nowadays, test batteries are a common tool among research groups studying primate cognition and behavior, allowing the standardization of training and assessment of cognitive mechanisms in NHP, in the laboratory and the field (Berger et al., 2018; Calapai et al., 2022, 2017; Fichtel et al., 2020; Inoue and Matsuzawa, 2007; Joly et al., 2017; Leinwand et al., 2020; Schmitt et al., 2012).

1.3 Automation

Automation refers to technology applications where human intervention is reduced. This is achieved by employing programmed commands that ensure the proper execution of a process (Groover, 2020; Lyshevski, 2008). Control loops can be either open or closed. In an open-loop, the control of an action is independent of the process output. For example, the heating system of an indoor monkey colony can be automatically controlled by a timer that turns on/off the heater (the process output) for a fixed amount of time to heat the room, this being independent of the room temperature. In a closed-loop, the action's control is dependent on the process output (Mayr, 1970). Following the previous example, in a closed-loop setting, the turning on/off of the heating system depends on the room's temperature, which can be monitored by a temperature sensor that sends feedback to a controller.

In the past three decades, behavioral research has benefited from an exceptional technological influx facilitating the optimization of experimental procedures. Microcontrollers have played a key role in this regard by facilitating the modularization of complex tasks through the individual control of smaller processes. Microcontrollers appeared in 1960, with the invention of the MOSFET (metal-oxide-semiconductor field-effect transistor) (Atalla et al., 1959), and are essentially a simple miniature computer designed to perform a single task. Microcontrollers typically contain single or multiple processors, memory, and programmable input/output peripherals. These peripherals enable them to interpret external signals when programmed as input (such as acceleration, temperature, movement, or button presses) or to control external devices when programmed as output (such as pumps, lights, speakers, or motors) (Heath, 2003). Additionally, the capability to transfer and receive data from other devices via various serial communication protocols (e.g., I2C, SPI, USB, UART) facilitates synchronization across different modules and the implementation of more complex systems (Harris and Harris, 2013), allows for synchronization across different modules and implementation of more complex systems. Nowadays, microcontrollers are no longer programmed in assembly language but in high-level programming languages such as C, Python, and JavaScript, which consequently improves their accessibility to non-experts (Mazzei et al., 2015). For these reasons, microcontrollers are a common low-cost means of data collection, sensing, and actuating in the physical world. They provide access to novel spatial and temporal scales of measurement, with the great possibility to flexibly tailor experimental setups based on the research requirements.

An example that illustrates the advantages of modularity in the case of automated training and testing in NHPs is the use of two different approaches to identify individual animals of different species housed under different conditions, as described in Chapters 2, 3, and 4 of this dissertation. In Chapters 2 and 3, individual marmosets were identified employing radio frequency identification (RFID) technology which detects and identifies tags attached to the targets (implanted subcutaneously in the marmosets) using electromagnetic fields. However, in Chapter 4, the use of RFID technology was not feasible due to physical constraints (housing and animal husbandry regulations). Instead, a machine learning algorithm was implemented to identify individual long-tailed macaques (*Macaca fascicularis*) based on pictures taken on a trial-by-trial basis. The two independent modules provided feedback to the main computer to reliably identify individual animals, keeping the system's architecture, functionality, and performance across several experimental conditions.

The next phase

Automation allows for the reduction of human intervention, saving human labor while also improving the accuracy and reproducibility of the process. Autor and colleagues (Autor and Duggan, 2003) divided automation into a two-by-two matrix: in one axis the routine versus non-routine tasks and in the other axis the manual versus cognitive tasks. In summary, routine tasks can be understood as tasks that follow explicit rules (e.g., a calculator), while non-routine tasks are not necessarily explicitly specified in computer code (e.g., a chatbot based on a machine learning algorithm). Each of these categories can, in turn, be of either manual (e.g., 3D printing) or cognitive (e.g., face and speech recognition) nature. They both relate either to the physical or knowledge work. Laboratory labor is comprised of tasks that fall into these four categories. Yet most of the current automated tasks belong to the physical labor category (routine and non-routine manual tasks) since they are the easiest to automate. However, as computing power increases, the type of tasks that machines can perform also increases and will inevitably continue to do so. Advances in electronics, mobile robotics, and artificial intelligence technology allow automation to expand to non-routine cognitive tasks, which until now have largely remained under human control (Brynjolfsson and McAfee, 2011).

Computer vision is an excellent example of the gradual automation of cognitive tasks in research. It aims to understand and automate processes that the human visual system can perform (Ballard and Brown, 1982; Huang, 1996). It has contributed significantly to the sophistication and detailed analysis of the behavior (mainly in rodents, with a few examples in NHPs). It has been crucial for the automation of human observation, which has been the standard approach for identifying, counting, and annotating individual or group animal movements and interactions for centuries (Datta et al., 2019;

De Chaumont et al., 2019; Geuther et al., 2021; Krynitsky et al., 2020; Lauer et al., 2022; Marshall et al., 2021; Mathis et al., 2018; Nath et al., 2019; Noldus et al., 2001; Norouzzadeh et al., 2018; Schofield et al., 2019; Sun et al., 2021; Wiltschko et al., 2015). Additionally, combining video analysis with state-of-the-art sensors (that can be fixed to the animal and continuously transmit data wirelessly) opens up new possibilities for measuring behavior (Berman, 2018; Iriki and Tramacere, 2022; Nishinaka et al., 2021; Nourizonoz et al., 2020; Schwarz et al., 2014; Venkatraman et al., 2010).

Historically, the amount and replicability of behavioral data have been among the major constraints for NHPs research. Technological advances have made feasible the automation of behavioral assessment, gaining reproducibility and replicability with a considerable reduction of human intervention. It is clear that researchers will soon have the capacity to acquire unprecedented amounts of data. However, several challenges of similar magnitude lie ahead and call for the development of adequate conceptual frameworks on how to properly collect, analyze, and handle behavioral data (Gomez-Marin et al., 2014; Krakauer et al., 2017). One big methodological challenge that is likely to remain in experimental research in NHP cognition is the design of innovative experiments that can exploit the curiosity of the individuals encouraging them to use their cognitive capabilities flexibly and creatively to solve the problems presented and yet still possess ecological validity for the species (Tomasello and Call, 2011).

1.4 Auditory cognition

Synopsis of the auditory pathway

Ascending auditory pathway

Auditory information processing involves the complex coordination of several systems along the auditory pathway, including peripheral sensory structures and several nervous system regions, to transform sound waves into distinct patterns of neural activity and guide behavior (Purves et al., 2000). The first stage of sound processing occurs in the outer and middle ear, funneling the sound waves into the ear canal and amplifying their pressure so the energy can be transferred to the fluid-filled cochlea. In the cochlea, a series of biomechanical mechanisms segregate the signal into simpler components, which are transduced by the inner hair cells and encoded as action potentials by the spiral ganglion neurons. Sound decomposition in the spectral domain occurs due to the cochlea frequency sensitivity arranged in space (Tonotopy). This spectral decomposition is the central organizational principle in the auditory system and is kept along the auditory pathway up to the primary auditory cortex. After the sound has been transduced to electric signals, the information reaches the cochlear nucleus, where parallel projections travel upstream to several targets. The main targets are the superior olivary complex and the lateral lemniscus, where binaural cues for sound localization are integrated. The inferior colliculus of the midbrain is the target of most lower brainstem stages and processes information regarding frequency and integrative cues for sound localization in space. Additionally, the inferior colliculus is the first stage, where auditory information can interact with motor information. From the inferior colliculus projections target the medial geniculate nucleus in the thalamus and finally to the primary auditory cortex, the first region in the cortex to receive auditory input, where integrative features of sound such as harmonic and temporal combinations are processed. Sound is thought to become the first consciously experienced at this level (Demanez and Demanez, 2003; Moore, 2000; Oliver, 2000; Purves et al., 2000; Ramirez-Moreno and Sejnowski, 2012; Shamma, 2001; Winer and Schreiner, 2005).

Descending auditory pathway

Anatomical studies showed that thalamo-cortical and cortico-thalamic connections are largely reciprocal in the auditory cortex (A1 and median belt) of NHPs, with additional projections bypassing the thalamus and directly targeting the inferior colliculus (Aitkin et al., 1988; De La Mothe et al., 2006; Luethke et al., 1989). However, the function of these pathways is unclear but usually associated with inhibitory function through negative feedback, facilitating and impeding the ascending of information to upstream regions in the brain. For example, descending input to the inferior colliculus has been shown to affect sensory tunning to match cortical activation (Suga and Ma, 2003).

Methods in NHP auditory research

The study of NHP sound perception and cognition has been covered across a substantial number of species (nearly 10% of the order), relying on methods that vary among species and the research question.

Behavioral methods

Psychophysics is an experimental approach employed by cognitive neuroscience to study the biological processes that underlie perception (Albright et al., 2000). Psychophysics studies the relationship between physical stimuli and the perception they produce (Gescheider, 2013). Psychoacoustics specifically focuses on sound perception or hearing, which might include physiological aspects (Fechner, 1948; Moore, 2014, 2013; Zwicker and Fastl, 2013).

Psychoacoustic methods require the cooperation of the animals to engage in experimental tasks, such as multiple alternative forced-choice tasks, where individuals are presented with multiple acoustic stimuli, from which only one contains the target sound. These methods are usually accompanied by visual stimuli and feedback, improving accuracy and reliability (Fechner, 1860; Levitt, 1971; Lopez-Poveda, 2014). A similar method is the Go-NoGo tests, where individuals are required to perform an action (e.g., pressing a button) when the "Go" stimulus is presented and withhold this action when the "NoGo" signal is presented (Georgiou and Essau, 2011; Osmanski et al., 2016). Another method that requires less cooperation from the subject is the "observer-based method," which consists of observing the subject's reaction to stimuli presentation (Olsho et al., 1987). The aforementioned tasks' difficulty can be adjusted using a staircase or adaptative procedures.

In behavioral methods, the animal response is usually stored in single units called trials, collected along a series of experimental sessions. The relationship between the signal and the observed behavior defines the sensory detection, discrimination, or perceptual judgment. For example, if the response to a stimulus appears only 50% of the time, it is usually interpreted as an estimated threshold of the animal. In the same way, the minimum detectable difference between two stimuli can be estimated by the times the animal chooses both stimuli 50% of the time (McNicol, 2005).

Physiological methods

Physiological approachers allow for hearing assessment without the attention or cooperation of the subjects. Some of these methods are especially relevant when behavioral techniques are challenging to use (e.g., assessing hearing in infants). Auditory brainstem responses (ABRs) are auditory evoked potentials that evaluate the functionality of the cochlear and lower auditory pathway structures by registering electrical activity via electrodes placed in the scalp. ABRs use clicks or tone stimuli to measure a series of electric potentials known as *waves* that correspond to the activation of several structures along the lower auditory pathway, including the cochlea (Burkard et al., 2007; Hall, 2007; Kraus et al., 1985; Moore, 1983). Distortion product otoacoustic emissions (DPOAEs) is generated by simultaneously presenting a pair of primary tones (f1 and f2) and recording the resulting sound produced by the cochlear through a microphone placed inside the ear canal (Brown and Gaskill, 1990; Lonsbury-Martin et al., 1988; Martin et al., 1988). Single-neuron and multielectrode array recordings enable a more detail measurement of neuronal activation during sound processing. However, these techniques are more invasive, and in most cases, animals are required to remain head-fixed in a primate chair during the recording sessions, which limits the variety of cognitive processes that can be evaluated. Recently, improvements in wireless radiotelemetry systems have enabled scientists to perform brain recordings in freely moving animals while transmitting the data wirelessly (Berger et al., 2020; Mohseni et al., 2005; Roy and Wang, 2012).

Fundamental aspects of auditory perception

Limits of perception

NHPs make use of different frequencies highly linked to body size and ecological needs (Brown and Waser, 2017; Ramsier and Rauschecker, 2017). The ability to detect the presence of a sound in the environment is a fundamental function of the auditory system. This ability is measured by a psychoacoustic threshold, which is the minimum sound intensity required to elicit perception in at least half of the trials. Because hearing thresholds vary across frequencies, audiograms additionally provide information about the hearing range of the subjects. A wide variety of audiograms (a plot of intensity expressed in decibels (dB) sound pressure level (SPL) as a function of frequency) in NHPs have been collected over the last 80 years using different techniques (speakers or headphones) (Coleman, 2009). This data shows that strepsirrhines and tarsiers possess the best sensitivity in high frequencies (around 16 kHz). In contrast, haplorhines have better sensitivity at lower frequencies (1 kHz to 10 kHz). Generally, the frequency sensitivity area overlaps with the spectra of the species' vocalizations (Coleman, 2009; Heffner, 2004). Additionally, whereas monkeys and apes (except humans) show a sensitivity curve in a W shape with two sensitive peaks, loris lemurs and human sensitivity curves have a V shape with a single sensitive peak (Coleman, 2009; Heffner, 2004; Ramsier and Rauschecker, 2017). Frequency hearing ranges in octaves for most NHPs vary from 8 to 10, with the narrowest in *Saimiri sciureus* at 8.75 and the broadest in *Macaca fuscata* at 10.29 (Heffner, 2004).

Frequency discrimination

Frequency discrimination refers to the ability of the subject to detect a difference in frequency between two pure tones. It is usually expressed as a percentage of change with respect to the reference frequency.

In old-world monkeys (Prosen et al., 1990; Sinnott et al., 1987), the minimum reported frequency change that animals can detect was found at 0.27%, around 4kHz. While in new-world monkeys, the minimum detectable frequency change has been reported to be around 2.2%-3.4% from frequencies between ~3.5–14 kHz (Osmanski et al., 2016; Recanzone et al., 1991; Wienicke et al., 2001). Finally, a study with chimpanzees reported a higher sensitivity, around 8 kHz (Kojima, 1990). The maximal discrimination sensitivity for the reported NHPs tends to overlap with the frequency content of their vocalizations, potentially representing an advantage for vocal processing.

Sound Localization

The ability of the subject to identify the location of a sound source in space is known as sound localization. Absolute sound localization refers to the localization of a sound source in three-dimensional space. In contrast, relative sound localization refers to the ability to detect a shift in orientation from the sound source and can be quantified by measuring the minimum shift audible angle (Mills, 1958). Sound localization relies on monoaural and binaural cues. For example, localization in the horizontal axis relies on the detection of interaural time (ITD) and level differences (ILD) (Middlebrooks and Green, 1991). In contrast, the sound localization in the vertical axis depends on the pinna structures that act as a direction-selective filter (Batteau, 1967).

In old-world monkeys (*Macaca*), the minimum audible angles for localization in the vertical and the horizontal axis ranged from 3° to 20° and from 4° to 20°, respectively, using conspecific vocalizations and bands of noise for the vertical axis and pure tones for the horizontal axis (Brown et al., 1982). For new world monkeys (*Callithrix*), the minimum audible angles for localization in the vertical and horizontal axis ranged from 17° to 22° and at 15°, respectively, using bandpass Gaussian noise (Remington et al., 2022).

Masking

Auditory masking occurs when the perception of a sound is affected by the presence of another sound (Greenwood, 1961). Masking can occur in the temporal or spectral domain and can be simultaneous or non-simultaneous. It can be measured as an increase in the detection threshold caused by the masker, where the amount of masking is presented in decibels. Simultaneous masking occurs when two sounds of equal duration are presented simultaneously, and just one can be perceived. The ability to perceive frequencies separately played simultaneously is known as frequency resolution. This phenomenon is thought to occur at the level of the basilar membrane in the cochlea due to its physical properties and tonotopical arrangement (Florentine et al., 1980; Moore, 1985).

Temporal masking can be forward or backward masking. In these two arrangements, the masker precedes or follows the target sound. Forward masking is believed to be caused by the adaptation of the neural responses to the masker when the target occurs and/or because the neural responses of the masker persist when the target occurs (Meddis and O'Mard, 2005; Oxenham, 2001). Less information is available explaining the physiological mechanisms underlying backward masking.

In macaque monkeys (*Macaca*), it has been reported that pure tone thresholds can be higher when the masker (Noise) is continuous or gated on and off simultaneously during target presentation and lower when the masker is switched off during target presentation. In addition, maskers were reported to have higher effects on hearing thresholds when they temporally overlap with the target. Furthermore, noise durations longer than 100 ms were necessary to observe masking effects. Finally, hearing thresholds decreased when the target and noise were separated in space compared to when both were co-located (Rocchi et al., 2017).

Temporal resolution

Temporal resolution refers to the minimum change in time that can be perceived. It is usually measured in humans using a gap-detection or amplitude modulation detection task (Moore, 2013). In the gap-detection task, individuals are asked to detect the presence of silence in a continuous sound (usually a bandwidth noise). While in the amplitude modulation task, subjects are asked to detect amplitude modulation changes (pure tones and bandwidth noise).

In macaque monkeys (*Macaca*), a threshold of 6.6 ms was found using a 2 kHz single-frequency tone (Izumi, 1999; Petkov et al., 2003).

Auditory scene analysis

Auditory scene analysis refers to the process of parsing complex auditory input into meaningful perceptual objects (Bregman, 1990). Grouping principles that underlie auditory scene analysis are related to those discovered by the school of Gestalt psychology (Mather, 2006). This categorization can be temporal (sequential grouping mechanisms) or spectral (simultaneous grouping mechanisms) (Bregman, 1990; Noorden, 1975). For example, a sequence of sounds might be perceived as coming from the same source or from two or more sources. These phenomena are known as coherence (perceived as one source) and stream segregation (perceived as two or more sources). Each sound source perceived is called a stream. Stream segregation always occurs when large spectro-temporal differences (e.g., in the temporal envelope, fundamental frequency, lateralization, phase spectrum, and rate of presentation) occur in successive sounds. When the differences are of intermediate size, the percept often switches from one stream to two or more. Coherence and stream segregation are suggested to be highly linked to attention and time exposure to acoustic stimuli (Moore and Gockel, 2012).

Macaques monkeys (*Macaca*) can abstract the relative relationship between tone sequences when the contours are judged in the direction of the frequency change (ascending or descending), yet when a brief silence gap (200 ms) is inserted between the tones the grouping effect is affected (Brosch et al., 2004; Izumi, 2001, 1999). Common marmosets were found to be able to discriminate trains of either 0.5- or 2-kHz tones repeated in either 50- or 200-ms intervals but failed to discriminate ABAB from AABB patterns consisting of A (0.5-kHz/50-ms pulse) and B (2-kHz/200-ms pulse) (Wakita, 2020, 2019).

Vocal perception

In contrast to simple acoustic stimuli that can be determined by their amplitude, frequency, and phase, complex acoustic stimuli, such as vocalizations, are more challenging to define because their spectral components may vary independently over time across amplitude and phase dimensions. Field and laboratory studies have sought to understand NHP communicative vocal repertoire relying on different approaches. Playback studies in the field found that when playing back various acoustic signals, different behavioral responses were reliably elicited, suggesting the existence of meaning signal classes (Cheney and Seyfarth, 1999, 1980; Fischer, 1998; Fischer et al., 1998; Gouzoules et al., 1984; Green, 1975; Seyfarth et al., 1980). In laboratory experiments, animals can be individually tested for particular variations in the stimulus, assessing for the minimum change necessary for discriminable differences (Beecher et al., 1979; Brown et al., 1979; Hopp et al., 1992; May et al., 1989, 1988; Owren, 1990; Zoloth et al., 1979). For example, rhesus monkeys (*Macaca mulatta*), the manipulation of spectral features of screams (vocalization used for specific aggressive interactions), such as the fundamental frequency, frequency bandwidth, and harmonic structure changes, were highly effective in altering stimulus classification responses (Le Prell et al., 2002).

However, a caveat when comparing these findings is the limited and non-homologous acoustic manipulation of the stimuli used for testing. It has been suggested that in order to demonstrate a detailed influence of specific acoustic features in vocal perception, precise access and flexible manipulation of individual acoustic elements of the tested stimuli are essential (DiMattina and Wang, 2006).

Cochlear implants

Cochlear implants are neuroprosthetics designed for patients with profound hearing loss. Cochlear implants restore hearing by directly stimulating the spiral ganglion neurons bypassing the commonly damaged inner hair cells in the cochlea (Lenarz, 2017; Zeng et al., 2008). Despite the hearing restoration success of cochlear implants, patients face several perceptual limitations, such as difficulties in understanding speech in noisy environments or music appreciation (Friesen et al., 2001; Zeng and Galvin III, 1999). Although it has been suggested that such perceptual limitations likely have peripheral causes, the understanding of how the auditory system makes sense of the limited input provided by cochlear implants is limited.

The common marmoset, a highly vocal primate species, has recently been seen as a model for cochlear implant research. Several studies have already established the feasibility of cochlear implant research by characterizing the temporal bone anatomy for implantation and the neural response in the auditory cortex in awake animals (Eliades and Tsunada, 2019; Johnson et al., 2017, 2016, 2012). However, these studies have only relied on relatively simple stimulation patterns in restrained marmosets with intermittent use of cochlear implants. Therefore, it remains unclear how the chronic use of cochlear implants could impact the plasticity along the auditory pathway and how marmosets would use the cochlear implant in ecologically meaningful behaviors such as vocal communication.

1.5 Aims of this dissertation

In the last pages, I have briefly described some of the methodological challenges that auditory cognition research in NHPs has faced over the last decades, along with the implementation of novel tools that have helped to mitigate these constraints.

In a loose sense, the exceptional development of technology in the past three decades has allowed researchers to drastically improve the quality and quantity of behavioral data. However, the methodological advances achieved by this new wave of technological influx are still under development thus, limited in application to a few fields. Experimental procedures in auditory cognition for NHPs lag behind the sophisticated methods used for visual cognition. These differences might result from historical preferences to study different models in auditory research (e.g., bats, ferrets, and birds) and the difficulties in reliably training NHPs in auditory tasks, which generally display a bias towards vision (Schmitt and Fischer, 2009).

As such, this dissertation intends to fill the gap by contributing to the optimization of experimental procedures for auditory cognition research in NHPs. Focusing in common marmosets and long-tailed macaques.

In the following chapters, I describe the development of a novel approach to automatically train and test common marmosets and long-tailed macaques for various aspects of auditory perception using audio-visual tasks directly in their homecages, without the need for social separation, neither water nor food control.

Chapter 2 – Describes the development of a novel experimental device named MXBI (marmoset experimental behavioral instrument) together with a series of automated unsupervised protocols that assess the auditory perception skills in socially housed common marmosets. This chapter has been published in *Nature communications 13, 1648 (2022).*

Chapter 3 – Describes a series of experiments that ultimately demonstrate the flexibility of use of the MXBI by automating commonly used psychoacoustic tasks, proving the feasibility of wireless control of cochlear implants, and testing for more complex cognitive mechanisms, such as vocal perception.

Chapter 4 – Describes the performance of four long-tailed macaques to a visuo-acoustic discrimination task in a social context. It also describes the performance of a novel animal identification protocol based on computer vision. This chapter has been published in *Frontiers in Psychology, 13, 10472242 (2022)*

Chapter 5 – Summarizes and discusses the findings of previous chapters and provides an outlook for further directions.

1.6 References

Aitkin, L.M., Kudo, M., Irvine, D.R., 1988. Connections of the primary auditory cortex in the common marmoset, Callithrix jacchus jacchus. J. Comp. Neurol. 269, 235–248.

Albright, T.D., Kandel, E.R., Posner, M.I., 2000. Cognitive neuroscience. Curr. Opin. Neurobiol. 10, 612–624. https://doi.org/10.1016/S0959-4388(00)00132-X

Andrews, M.W., 1993. Video-Task Paradigm Extended to *Saimiri*. Percept. Mot. Skills 76, 183–191. https://doi.org/10.2466/pms.1993.76.1.183

Atalla, M.M., Tannenbaum, E., Scheibner, E.J., 1959. Stabilization of Silicon Surfaces by Thermally Grown Oxides*. Bell Syst. Tech. J. 38, 749–783. https://doi.org/10.1002/j.1538-7305.1959.tb03907.x

Autor, D.H., Duggan, M.G., 2003. The Rise in the Disability Rolls and the Decline in Unemployment. Q. J. Econ. 118, 157–206. https://doi.org/10.1162/00335530360535171

Bala, P.C., Eisenreich, B.R., Yoo, S.B.M., Hayden, B.Y., Park, H.S., Zimmermann, J., 2020. Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio. Nat. Commun. 11, 4560. https://doi.org/10.1038/s41467-020-18441-5

Ballard, D.H., Brown, C.M., 1982. Computer vision. Prentice-Hall, Englewood Cliffs, N.J.

Balleine, B., 1992. Instrumental performance following a shift in primary motivation depends on incentive learning. J. Exp. Psychol. Anim. Behav. Process. 18, 236–250. https://doi.org/10.1037/0097-7403.18.3.236

Batteau, 1967. The role of the pinna in human localization. Proc. R. Soc. Lond. B Biol. Sci. 168, 158–180. https://doi.org/10.1098/rspb.1967.0058

Beecher, M.D., Petersen, M.R., Zoloth, S.R., Moody, D.B., Stebbins, W.C., 1979. Perception of conspecific vocalizations by Japanese macaques. Brain. Behav. Evol. 16, 443–460.

Beran, M.J., Parrish, A.E., Perdue, B.M., Washburn, D.A., 2014. Comparative cognition: Past, present, and future. Int. J. Comp. Psychol.

Berger, M., Agha, N.S., Gail, A., 2020. Wireless recording from unrestrained monkeys reveals motor goal encoding beyond immediate reach in frontoparietal cortex. Elife 9, e51322.

Berger, M., Calapai, A., Stephan, V., Niessing, M., Burchardt, L., Gail, A., Treue, S., 2018. Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. J. Neurophysiol. 119, 796–807. https://doi.org/10.1152/jn.00614.2017

Berman, G.J., 2018. Measuring behavior across scales. BMC Biol. 16, 23. https://doi.org/10.1186/s12915-018-0494-7

Brandon-Jones, D., Eudey, A.A., Geissmann, T., Groves, C.P., Melnick, D.J., Morales, J.C., Shekelle, M., Stewart, C.-B., 2004. Asian Primate

Classification. Int. J. Primatol. 25, 97–164. https://doi.org/10.1023/B:IJOP.0000014647.18720.32

Bregman, A.S., 1990. Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge, Mass.

Brosch, M., Selezneva, E., Bucks, C., Scheich, H., 2004. Macaque monkeys discriminate pitch relationships. Cognition 91, 259–272. https://doi.org/10.1016/j.cognition.2003.09.005

Brown, A.M., Gaskill, S.A., 1990. Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses. J. Acoust. Soc. Am. 88, 840–849.

Brown, C.H., Beecher, M.D., Moody, D.B., Stebbins, W.C., 1979. Locatability of vocal signals in Old World monkeys: Design features for the communication of position. J. Comp. Physiol. Psychol. 93, 806.

Brown, C.H., Schessler, T., Moody, D., Stebbins, W., 1982. Vertical and horizontal sound localization in primates. J. Acoust. Soc. Am. 72, 1804–1811. https://doi.org/10.1121/1.388653

Brown, C.H., Waser, P.M., 2017. Primate Habitat Acoustics, in: Quam, R.M., Ramsier, M.A., Fay, R.R., Popper, A.N. (Eds.), Primate Hearing and Communication, Springer Handbook of Auditory Research. Springer International Publishing, Cham, pp. 79–107. https://doi.org/10.1007/978-3-319-59478-1_4

Brynjolfsson, E., McAfee, A., 2011. Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy. Lexington, Massachusetts: Digital Frontier Press.

Burkard, R.F., Eggermont, J.J., Don, M. (Eds.), 2007. Auditory evoked potentials: basic principles and clinical application. Lippincott Williams & Wilkins, Philadelphia.

Byrne, R.W., 1995. Primate cognition: Comparing problems and skills. Am. J. Primatol. 37, 127–141. https://doi.org/10.1002/ajp.1350370206

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., Gail, A., 2017. A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. Behav. Res. Methods 49, 35–45. https://doi.org/10.3758/s13428-016-0707-3

Calapai, A., Cabrera-Moreno, J., Moser, T., Jeschke, M., 2022. Flexible auditory training, psychophysics, and enrichment of common marmosets with an automated, touchscreen-based system. Nat. Commun. 13, 1648. https://doi.org/10.1038/s41467-022-29185-9

Canguilhem, G., 2022. Knowledge of life, in: Knowledge of Life. Fordham University Press.

Cauchoix, M., Hermer, E., Chaine, A.S., Morand-Ferron, J., 2017. Cognition in the field: comparison of reversal learning performance in captive and wild passerines. Sci. Rep. 7, 12945. https://doi.org/10.1038/s41598-017-13179-5

Cheney, D.L., Seyfarth, R.M., 1999. Recognition of other individuals' social relationships by female baboons. Anim. Behav. 58, 67–75.

Cheney, D.L., Seyfarth, R.M., 1990. How monkeys see the world: Inside the mind of another species. University of Chicago Press.

Cheney, D.L., Seyfarth, R.M., 1980. Vocal recognition in free-ranging vervet monkeys. Anim. Behav. 28, 362–367. https://doi.org/10.1016/S0003-3472(80)80044-3

Chittka, L., Rossiter, S.J., Skorupski, P., Fernando, C., 2012. What is comparable in comparative cognition? Philos. Trans. R. Soc. B Biol. Sci. 367, 2677–2685.

Clayton, N.S., 2006. The Prospective Cognition of Food Caching and Recovery by Western Scrub-Jays (Alphelocoma californica). Comp. Cogn. Behav. Rev. 1. https://doi.org/10.3819/ccbr.2008.10001

Clayton, N.S., Bussey, T.J., Dickinson, A., 2003. Can animals recall the past and plan for the future? Nat. Rev. Neurosci. 4, 685–691. https://doi.org/10.1038/nrn1180

Clayton, N.S., Dally, J.M., Emery, N.J., 2007. Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. Philos. Trans. R. Soc. B Biol. Sci. 362, 507–522. https://doi.org/10.1098/rstb.2006.1992

Coleman, M.N., 2009. What Do Primates Hear? A Meta-analysis of All Known Nonhuman Primate Behavioral Audiograms. Int. J. Primatol. 30, 55–91. https://doi.org/10.1007/s10764-008-9330-1

Colgan, P., 2013. Animal motivation. Springer Science & Business Media.

Crofts, H.S., Muggleton, N.G., Bowditch, A.P., Pearce, P.C., Nutt, D.J., Scott, E.A.M., 1999. Home cage presentation of complex discrimination tasks to marmosets and rhesus monkeys. Lab. Anim. 33, 207–214. https://doi.org/10.1258/002367799780578174

Dally, J.M., Emery, N.J., Clayton, N.S., 2006. Food-Caching Western Scrub-Jays Keep Track of Who Was Watching When. Science 312, 1662–1665. https://doi.org/10.1126/science.1126539

Darwin, C., 2004. On the origin of species, 1859. Routledge.

Datta, S.R., Anderson, D.J., Branson, K., Perona, P., Leifer, A., 2019. Computational neuroethology: a call to action. Neuron 104, 11–24.

Davenport, J.W., Chamove, A.S., Harlow, H.F., 1970. The semiautomatic Wisconsin general test apparatus. Behav. Res. Methods Instrum. 2, 135–138. https://doi.org/10.3758/BF03211023

De Chaumont, F., Ey, E., Torquet, N., Lagache, T., Dallongeville, S., Imbert, A., Legou, T., Le Sourd, A.-M., Faure, P., Bourgeron, T., 2019. Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. Nat. Biomed. Eng. 3, 930–942.

De La Mothe, L.A., Blumell, S., Kajikawa, Y., Hackett, T.A., 2006. Thalamic connections of the auditory cortex in marmoset monkeys: core and medial belt regions. J. Comp. Neurol. 496, 72–96.

Demanez, J.P., Demanez, L., 2003. Anatomophysiology of the central auditory nervous system: basic concepts. Acta Otorhinolaryngol. Belg. 57, 227–236.

Dias, R., Robbins, T.W., Roberts, A.C., 1996. Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380, 69–72. https://doi.org/10.1038/380069a0

Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18.

DiMattina, C., Wang, X., 2006. Virtual Vocalization Stimuli for Investigating Neural Representations of Species-Specific Vocalizations. J. Neurophysiol. 95, 1244–1262. https://doi.org/10.1152/jn.00818.2005

Eliades, S.J., Tsunada, J., 2019. Marmosets in auditory research, in: The Common Marmoset in Captivity and Biomedical Research. Elsevier, pp. 451–475.

Farrar, B.G., Krupenye, C., Rodrigo, A.M., Tennie, C., Fischer, J., Altschul, D., Ostojic, L., 2021. Replication and Reproducibility in Primate Cognition Research (preprint). Open Science Framework. https://doi.org/10.31219/osf.io/z52er

Fay, R.R., 2012. Comparative hearing: mammals. Springer Science & Business Media.

Fechner, G.T., 1948. Elements of psychophysics, 1860. In: Dennis W, editor. Readings in the History of Psychology. East Norwalk: Appleton-Century-Crofts. pp. 206-213. doi:10.1037/11304-026

Fechner, G.T., 1860. Elemente der psychophysik. Breitkopf u. Härtel.

Fichtel, C., Dinter, K., Kappeler, P.M., 2020. The lemur baseline: how lemurs compare to monkeys and apes in the Primate Cognition Test Battery. PeerJ 8, e10025.

Fischer, J., 1998. Barbary macaques categorize shrill barks into two call types. Anim. Behav. 55, 799–807. https://doi.org/10.1006/anbe.1997.0663

Fischer, J., Hammerschmidt, K., Todt, D., 1998. Local variation in Barbary macaque shrill barks. Anim. Behav. 56, 623–629.

Florentine, M., Buus, S., Scharf, B., Zwicker, E., 1980. Frequency Selectivity in Normally-Hearing and Hearing-Impaired Observers. J. Speech Lang. Hear. Res. 23, 646–669. https://doi.org/10.1044/jshr.2303.646

Friesen, L.M., Shannon, R.V., Baskent, D., Wang, X., 2001. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. J. Acoust. Soc. Am. 110, 1150–1163.

Gagic, V., Kirkland, L., Kendall, L.K., Jones, J., Kirkland, J., Spurr, C., Rader, R., 2021. Understanding pollinator foraging behaviour and transition rates between flowers is important to maximize seed set in hybrid crops. Apidologie 52, 89–100. https://doi.org/10.1007/s13592-020-00800-2

Georgiou, G., Essau, C.A., 2011. Go/No-Go Task, in: Goldstein, S., Naglieri, J.A. (Eds.), Encyclopedia of Child Behavior and Development. Springer US, Boston, MA, pp. 705–706. https://doi.org/10.1007/978-0-387-79061-9_1267

Gescheider, G.A., 2013. Psychophysics, 0 ed. Psychology Press. https://doi.org/10.4324/9780203774458

Geuther, B.Q., Peer, A., He, H., Sabnis, G., Philip, V.M., Kumar, V., 2021. Action detection using a neural network elucidates the genetics of mouse grooming behavior. Elife 10, e63207.

Glazko, G.V., 2003. Estimation of Divergence Times for Major Lineages of Primate Species. Mol. Biol. Evol. 20, 424–434. https://doi.org/10.1093/molbev/msg050

Gomez-Marin, A., Paton, J.J., Kampff, A.R., Costa, R.M., Mainen, Z.F., 2014. Big behavioral data: psychology, ethology and the foundations of neuroscience. Nat. Neurosci. 17, 1455–1462. https://doi.org/10.1038/nn.3812

Goodall, J., 1964. Tool-using and aimed throwing in a community of free-living chimpanzees. Nature 201, 1264–1266.

Gouzoules, S., Gouzoules, H., Marler, P., 1984. Rhesus monkey (Macaca mulatta) screams: representational signalling in the recruitment of agonistic aid. Anim. Behav. 32, 182–193.

Green, S., 1975. Variation of vocal pattern with social situation in the Japanese monkey (Macaca fuscata): a field study.

Greenwood, D.D., 1961. Auditory Masking and the Critical Band. J. Acoust. Soc. Am. 33, 484–502. https://doi.org/10.1121/1.1908699

Groover, M.P., 2020. Fundamentals of modern manufacturing: materials, processes, and systems. John Wiley & Sons.

Hall, J.W., 2007. New handbook of auditory evoked responses. Pearson, Boston, Mass. London.

Harlow, H.F., Bromer, J.A., 1938. A test-apparatus for monkeys. Psychol. Rec. 2, 434–436. https://doi.org/10.1007/BF03393227

Harris, D.M., Harris, S.L., 2013. Digital design and computer architecture, Second edition. ed. Elsevier, Amsterdam.

Hayden, B., Park, H.S., Zimmermann, J., 2021. Automated Tracking of Primate Behavior. https://doi.org/10.48550/ARXIV.2108.13486

Heath, S., 2003. Embedded systems design, 2nd ed. ed. Newnes, Oxford ; Boston.

Hebb, D.O., Donderi, D.C., 2013. Textbook of psychology (psychology revivals). Psychology Press.

Heffner, R.S., 2004. Primate hearing from a mammalian perspective. Anat. Rec. 281A, 1111–1122. https://doi.org/10.1002/ar.a.20117

Hodžić, A., Alić, A., Fuehrer, H.-P., Harl, J., Wille-Piazzai, W., Duscher, G., 2015. A molecular survey of vector-borne pathogens in red foxes (Vulpes vulpes) from Bosnia and Herzegovina. Parasit. Vectors 8, 88. https://doi.org/10.1186/s13071-015-0692-x

Hogan, J., 2005. Causation: the study of behavioural mechanisms. Anim. Biol. 55, 323–341. https://doi.org/10.1163/157075605774840969

Hopp, S.L., Sinnott, J.M., Owren, M.J., Petersen, M.R., 1992. Differential sensitivity of Japanese macaques (Macaca fuscata) and humans (Homo sapiens) to peak position along a synthetic coo call continuum. J. Comp. Psychol. 106, 128.

Huang, T., 1996. Computer Vision: Evolution And Promise. https://doi.org/10.5170/CERN-1996-008.21

Hunt, J., 2018. Animal Behaviour. EDTECH.

Hurlbert, S.H., 1984. Pseudoreplication and the Design of Ecological Field Experiments. Ecol. Monogr. 54, 187–211. https://doi.org/10.2307/1942661

Inoue, S., Matsuzawa, T., 2007. Working memory of numerals in chimpanzees. Curr. Biol. 17, R1004–R1005.

Iriki, A., Tramacere, A., 2022. "Natural Laboratory Complex" for novel primate neuroscience. Front. Integr. Neurosci. 16, 927605. https://doi.org/10.3389/fnint.2022.927605

Izumi, A., 2001. Relative pitch perception in Japanese monkeys (Macaca fuscata). J. Comp. Psychol. 115, 127–131. https://doi.org/10.1037/0735-7036.115.2.127

Izumi, A., 1999. The Effect of Marker Frequency Disparity on the Discrimination of Gap Duration in Monkeys. Perception 28, 437–444. https://doi.org/10.1068/p2812

Jacob, G., Katti, H., Cherian, T., Das, J., Zhivago, K., Arun, S., 2021. A naturalistic environment to study visual cognition in unrestrained monkeys. eLife 10, e63816. https://doi.org/10.7554/eLife.63816

Jennings, M., Prescott, M.J., Refinement (Primates), J.W.G. on, 2009. Refinements in husbandry, care and common procedures for non-human primates: Ninth report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement. Lab. Anim. 43, 1–47.

Ji, N., Madan, G.K., Fabre, G.I., Dayan, A., Baker, C.M., Kramer, T.S., Nwabudike, I., Flavell, S.W., 2021. A neural circuit for flexible control of persistent behavioral states. eLife 10, e62889. https://doi.org/10.7554/eLife.62889

Johnson, L.A., Della Santina, C.C., Wang, X., 2017. Representations of time-varying cochlear implant stimulation in auditory cortex of awake marmosets (Callithrix jacchus). J. Neurosci. 37, 7008–7022.

Johnson, L.A., Della Santina, C.C., Wang, X., 2016. Selective Neuronal Activation by Cochlear Implant Stimulation in Auditory Cortex of Awake Primate. J.

Neurosci. 36, 12468–12484. https://doi.org/10.1523/JNEUROSCI.1699-16.2016

Johnson, L.A., Della Santina, C.C., Wang, X., 2012. Temporal bone characterization and cochlear implant feasibility in the common marmoset (Callithrix jacchus). Hear. Res. 290, 37–44. https://doi.org/10.1016/j.heares.2012.05.002

Joly, M., Micheletta, J., De Marco, A., Langermans, J.A., Sterck, E.H.M., Waller, B.M., 2017. Comparing physical and social cognitive skills in macaque species with different degrees of social tolerance. Proc. R. Soc. B Biol. Sci. 284, 20162738. https://doi.org/10.1098/rspb.2016.2738

Jung, Y., Kennedy, A., Chiu, H., Mohammad, F., Claridge-Chang, A., Anderson, D.J., 2020. Neurons that Function within an Integrator to Promote a Persistent Behavioral State in Drosophila. Neuron 105, 322-333.e5. https://doi.org/10.1016/j.neuron.2019.10.028

Kirkwood, J., Hubrecht, R.C., 2010. The UFAW handbook on the care and management of laboratory and other research animals. John Wiley & Sons.

Kohler, W., 2018. The mentality of apes. Routledge.

Kojima, S., 1990. Comparison of Auditory Functions in the Chimpanzee and Human. Folia Primatol. (Basel) 55, 62–72. https://doi.org/10.1159/000156501

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., Poeppel, D., 2017. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 93, 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kraus, N., Smith, D.I., Reed, N.L., Willott, J., Erwin, J., 1985. Auditory brainstem and middle latency responses in non-human primates. Hear. Res. 17, 219–226. https://doi.org/10.1016/0378-5955(85)90066-8

Krynitsky, J., Legaria, A.A., Pai, J.J., Garmendia-Cedillos, M., Salem, G., Pohida, T., Kravitz, A.V., 2020. Rodent arena tracker (RAT): A machine vision rodent tracking Camera and closed loop control system. Eneuro 7.

Lashley, K.S., 1951. The problem of serial order in behavior. Bobbs-Merrill Oxford.

Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M.M., Di Santo, V., Soberanes, D., Feng, G., 2022. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat. Methods 19, 496–504.

Laule, G.E., Bloomsmith, M.A., Schapiro, S.J., 2003. The Use of Positive Reinforcement Training Techniques to Enhance the Care, Management, and Welfare of Primates in the Laboratory. J. Appl. Anim. Welf. Sci. 6, 163–173. https://doi.org/10.1207/S15327604JAWS0603_02

Le Prell, C.G., Hauser, M.D., Moody, D.B., 2002. Discrete or graded variation within rhesus monkey screams? Psychophysical experiments on classification. Anim. Behav. 63, 47–62.

Leinwand, J.G., Huskisson, S.M., Egelkamp, C.L., Hopper, L.M., 2020. Within-and between-species variation in the responses of three primate species to a touchscreen gambling task. Learn. Motiv. 71, 101635.

Lenarz, T., 2017. Cochlear implant–state of the art. Laryngo-Rhino-Otol. 96, S123–S151.

Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.

Lonsbury-Martin, B.L., Martin, G.K., Probst, R., Coats, A.C., 1988. Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy. Hear. Res. 33, 69–93. https://doi.org/10.1016/0378-5955(88)90021-4

Lopez-Poveda, E.A., 2014. Development of fundamental aspects of human auditory perception. Dev. Audit. Vestib. Syst. 287–314.

Lorenz, K., 1937. Über die Bildung des Instinktbegriffes. Naturwissenschaften 25, 324–331. https://doi.org/10.1007/BF01774269

Lorenz, K.Z., 1950. The comparative method in studying innate behavior patterns.

Luethke, L.E., Krubitzer, L.A., Kaas, J.H., 1989. Connections of primary auditory cortex in the New World monkey, Saguinus. J. Comp. Neurol. 285, 487–513.

Lyshevski, S.E., 2008. Electromechanical systems and devices. CRC Press.

Marshall, J.D., Aldarondo, D.E., Dunn, T.W., Wang, W.L., Berman, G.J., Ölveczky, B.P., 2021. Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire. Neuron 109, 420-437. e8.

Martin, G.K., Lonsbury-Martin, B.L., Probst, R., Coats, A.C., 1988. Spontaneous otoacoustic emissions in a nonhuman primate. I. Basic features and relations to other emissions. Hear. Res. 33, 49–68. https://doi.org/10.1016/0378-5955(88)90020-2

Mather, G., 2006. Foundations of perception. Taylor & Francis.

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. 21, 1281–1289.

May, B., Moody, D.B., Stebbins, W.C., 1989. Categorical perception of conspecific communication sounds by Japanese macaques, *Macacafuscata*. J. Acoust. Soc. Am. 85, 837–847. https://doi.org/10.1121/1.397555

May, B., Moody, D.B., Stebbins, W.C., 1988. The significant features of Japanese macaque coo sounds: a psychophysical study. Anim. Behav. 36, 1432–1444.

Mayr, O., 1970. The origins of feedback control. Sci. Am. 223, 110–119.

Mazzei, D., Montelisciani, G., Baldi, G., Fantoni, G., 2015. Changing the programming paradigm for the embedded in the IoT domain, in: 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT). Presented at the 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), IEEE, Milan, Italy, pp. 239–244. https://doi.org/10.1109/WF-IoT.2015.7389059

McNicol, D., 2005. A Primer of Signal Detection Theory, 0 ed. Psychology Press. https://doi.org/10.4324/9781410611949

Meddis, R., O'Mard, L.P., 2005. A computer model of the auditory-nerve response to forward-masking stimuli. J. Acoust. Soc. Am. 117, 3787–3798. https://doi.org/10.1121/1.1893426

Menda, G., Nitzany, E.I., Shamble, P.S., Wells, A., Harrington, L.C., Miles, R.N., Hoy, R.R., 2019. The Long and Short of Hearing in the Mosquito Aedes aegypti. Curr. Biol. 29, 709-714.e4. https://doi.org/10.1016/j.cub.2019.01.026

Menzel, C.R., 1991. Cognitive aspects of foraging in Japanese monkeys. Anim. Behav. 41, 397–402. https://doi.org/10.1016/S0003-3472(05)80840-1

Menzel, R., Fischer, J., 2011. Animal thinking: contemporary issues in comparative cognition. MIT press.

Middlebrooks, J., Green, D., 1991. Sound Localization by Human Listeners. Annu. Rev. Psychol. 42, 135–59. https://doi.org/10.1146/annurev.ps.42.020191.001031

Mills, A.W., 1958. On the Minimum Audible Angle. J. Acoust. Soc. Am. 30, 237–246. https://doi.org/10.1121/1.1909553

Mohseni, P., Najafi, K., Eliades, S.J., Wang, X., 2005. Wireless multichannel biopotential recording using an integrated FM telemetry circuit. IEEE Trans. Neural Syst. Rehabil. Eng. 13, 263–271.

Moore, B.C., 2014. Psychoacoustics, in: Springer Handbook of Acoustics. Springer, pp. 475–517.

Moore, B.C.J., 2013. An introduction to the psychology of hearing, Sixth edition. ed. Brill, Leiden.

Moore, B.C.J., 1985. Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. Br. J. Audiol. 19, 189–201. https://doi.org/10.3109/03005368509078973

Moore, B.C.J., Gockel, H.E., 2012. Properties of auditory stream formation. Philos. Trans. R. Soc. B Biol. Sci. 367, 919–931. https://doi.org/10.1098/rstb.2011.0355

Moore, E.J. (Ed.), 1983. Bases of auditory brainstem evoked responses. Grune & Stratton, New York.

Moore, J.K., 2000. Organization of the human superior olivary complex. Microsc. Res. Tech. 51, 403–412. https://doi.org/10.1002/1097-0029(20001115)51:4<403::AID-JEMT8>3.0.CO;2-Q

Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., Mathis, M.W., 2019. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat. Protoc. 14, 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Nishinaka, M., Jinno, H., Jimbo, Y., Lee, S., Wang, J., Lee, W., Yokota, T., Someya, T., 2021. High-Transconductance Organic Electrochemical Transistor

Fabricated on Ultrathin Films Using Spray Coating. Small Struct. 2, 2000088. https://doi.org/10.1002/sstr.202000088

Noldus, L.P.J.J., Spink, A.J., Tegelenbosch, R.A.J., 2001. EthoVision: A versatile video tracking system for automation of behavioral experiments. Behav. Res. Methods Instrum. Comput. 33, 398–414. https://doi.org/10.3758/BF03195394

Noorden, van L.P.A.S., 1975. Temporal coherence and the perception of temporal position in tone sequences. IPO Annu. Prog. Rep.

Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc. Natl. Acad. Sci. 115, E5716–E5725.

Nourizonoz, A., Zimmermann, R., Ho, C.L.A., Pellat, S., Ormen, Y., Prévost-Solié, C., Reymond, G., Pifferi, F., Aujard, F., Herrel, A., Huber, D., 2020. EthoLoop: automated closed-loop neuroethology in naturalistic environments. Nat. Methods 17, 1052–1059. https://doi.org/10.1038/s41592-020-0961-2

Oliver, D.L., 2000. Ascending efferent projections of the superior olivary complex. Microsc. Res. Tech. 51, 355–363. https://doi.org/10.1002/1097-0029(20001115)51:4<355::AID-JEMT5>3.0.CO;2-J

Olsho, L.W., Koch, E.G., Halpin, C.F., Carter, E.A., 1987. An observer-based psychoacoustic procedure for use with young infants. Dev. Psychol. 23, 627.

Osmanski, M.S., Song, X., Guo, Y., Wang, X., 2016. Frequency discrimination in the common marmoset (Callithrix jacchus). Hear. Res. 341, 1–8. https://doi.org/10.1016/j.heares.2016.07.006

Owren, M.J., 1990. Acoustic classification of alarm calls by vervet monkeys (Cercopithecus aethiops) and humans (Homo sapiens): I. Natural calls. J. Comp. Psychol. 104, 20.

Oxenham, A.J., 2001. Forward masking: Adaptation or integration? J. Acoust. Soc. Am. 109, 732–741. https://doi.org/10.1121/1.1336501

Palmer, A.R., 2000. Quasireplication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. Annu. Rev. Ecol. Syst. 441–480.

Parrish, A.E., Brosnan, S.F., 2012. Primate Cognition. Encyclopedia of Human Behavior: (Vol. 3). Elsevier Inc.

Pavlov, I.P., 1928. Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals. Liverwright Publishing Corporation, New York. https://doi.org/10.1037/11081-000

Pearce, P.C., Crofts, H.S., Muggleton, N.G., Scott, E.A.M., 1998. Concurrent Monitoring of EEG and Performance in the Common Marmoset: A Methodological Approach. Physiol. Behav. 63, 591–599. https://doi.org/10.1016/S0031-9384(97)00494-0

Perdue, B.M., Beran, M.J., Washburn, D.A., 2018. A computerized testing system for primates: Cognition, welfare, and the Rumbaughx. Behav. Processes 156, 37–50. https://doi.org/10.1016/j.beproc.2017.12.019

Perez, S.I., Tejedor, M.F., Novo, N.M., Aristide, L., 2013. Divergence Times and the Evolutionary Radiation of New World Monkeys (Platyrrhini, Primates): An Analysis of Fossil and Molecular Data. PLoS ONE 8, e68029. https://doi.org/10.1371/journal.pone.0068029

Petkov, C.I., O'Connor, K.N., Sutter, M.L., 2003. Illusory Sound Perception in Macaque Monkeys. J. Neurosci. 23, 9155–9161. https://doi.org/10.1523/JNEUROSCI.23-27-09155.2003

Polidora, V.J., Main, W.T., 1963. PUNCHED CARD PROGRAMMING AND RECORDING TECHNIQUES EMPLOYED IN THE AUTOMATION OF THE WGTA [1]. J. Exp. Anal. Behav. 6, 599–603. https://doi.org/10.1901/jeab.1963.6-599

Pontzer, H., 2015. Energy Expenditure in Humans and Other Primates: A New Synthesis. Annu. Rev. Anthropol. 44, 169–187. https://doi.org/10.1146/annurev-anthro-102214-013925

Povinelli, D.J., Bering, J.M., Giambrone, S., 2000. Toward a science of other minds: Escaping the argument by analogy. Cogn. Sci. 24, 509–541.

Prescott, M.J., Buchanan-Smith, H.M., 2003. Training Nonhuman Primates Using Positive Reinforcement Techniques. J. Appl. Anim. Welf. Sci. 6, 157–161. https://doi.org/10.1207/S15327604JAWS0603_01

Primates, M., Altschul, D., Beran, M.J., Bohn, M., Caspar, K., Fichtel, C., Försterling, M., Grebe, N., Hernandez-Aguilar, R.A., Kwok, S.C., Rodrigo, A.M., Proctor, D., Sanchez-Amaro, A., Simpson, E.A., Szabelska, A., Taylor, D., van der Mescht, J., Völter, C., Watzek, J., 2019. Collaborative open science as a way to reproducibility and new insights in primate cognition research (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/8w7zd

Pritchard, D.J., Hurly, T.A., Tello-Ramos, M.C., Healy, S.D., 2016. Why study cognition in the wild (and how to test it)?: Cognition in the wild. J. Exp. Anal. Behav. 105, 41–55. https://doi.org/10.1002/jeab.195

Prosen, C.A., Moody, D.B., Sommers, M.S., Stebbins, W.C., 1990. Frequency discrimination in the monkey. J. Acoust. Soc. Am. 88, 2152–2158. https://doi.org/10.1121/1.400112

Purves, D., Fitzpatrick, D., Katz, L.C., Lamantia, A.S., McNamara, J.O., Williams, S.M., Augustine, G.J., 2000. Neuroscience. Sinauer Associates.

Purves, D., Williams, S.M. (Eds.), 2001. Neuroscience, 2nd ed. ed. Sinauer Associates, Sunderland, Mass.

Raby, C.R., Alexis, D.M., Dickinson, A., Clayton, N.S., 2007. Planning for the future by western scrub-jays. Nature 445, 919–921. https://doi.org/10.1038/nature05575

Ramirez-Moreno, D.F., Sejnowski, T.J., 2012. A computational model for the modulation of the prepulse inhibition of the acoustic startle reflex. Biol. Cybern. 106, 169–176.

Ramsier, M.A., Rauschecker, J.P., 2017. Primate audition: reception, perception, and ecology, in: Primate Hearing and Communication. Springer, pp. 47–77.

Recanzone, G.H., Jenkins, W.M., Hradek, G.T., Merzenich, M.M., 1991. A behavioral frequency discrimination paradigm for use in adult primates. Behav. Res. Methods Instrum. Amp Comput. 23, 357–369. https://doi.org/10.3758/BF03203397

Reinhardt, V., 1997. Training nonhuman primates to cooperate during blood collection: A review. Lab. Primate Newsl. 36, 1–4.

Remington, E.D., Chen, C., Wang, X., 2022. Sound localization acuity of the common marmoset (*Callithrix jacchus*) (preprint). Neuroscience. https://doi.org/10.1101/2022.08.13.503849

Rennie, A.E., Buchanan-Smith, H.M., 2006a. Refinement of the use of non-human primates in scientific research. Part I: the influence of humans. Anim. Welf.-POTTERS BAR THEN WHEATHAMPSTEAD- 15, 203.

Rennie, A.E., Buchanan-Smith, H.M., 2006b. Refinement of the use of non-human primates in scientific research. Part II: Housing, husbandry and acquisition. Anim. Welf.-POTTERS BAR THEN WHEATHAMPSTEAD- 15, 215.

Richardson, W.K., Washburn, D.A., Hopkins, W.D., Savage-rumbaugh, E.S., Rumbaugh, D.M., 1990. The NASA/LRC computerized test system. Behav. Res. Methods Instrum. Amp Comput. 22, 127–131. https://doi.org/10.3758/BF03203132

Roberts, A.C., Robbins, T.W., Everitt, B.J., 1988. The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. Q. J. Exp. Psychol. B 40, 321–341.

Rocchi, F., Dylla, M.E., Bohlen, P.A., Ramachandran, R., 2017. Spatial and temporal disparity in signals and maskers affects signal detection in non-human primates. Hear. Res. 344, 1–12. https://doi.org/10.1016/j.heares.2016.10.013

Röder, E.L., Timmermans, P.J.A., 2002. Housing and care of monkeys and apes in laboratories: adaptations allowing essential species-specific behaviour. Lab. Anim. 36, 221–242.

Roitblat, H.L., 1987. Introduction to comparative cognition. WH Freeman/Times Books/Henry Holt & Co.

Rose, K.D., 2006. The beginning of the age of mammals. JHU Press.

Roy, S., Wang, X., 2012. Wireless multi-channel single unit recording in freely moving and vocalizing primates. J. Neurosci. Methods 203, 28–40.

Rumbaugh, D.M., Hopkins, W.D., Washburn, D.A., Savage-Rumbaugh, E.S., 1989a. Lana Chimpanzee Learns to Count by "Numath": A Summary of a Videotaped Experimental Report. Psychol. Rec. 39, 459–470. https://doi.org/10.1007/BF03395074

Rumbaugh, D.M., Richardson, W.K., Washburn, D.A., Savage-Rumbaugh, E.S., Hopkins, W.D., 1989b. Rhesus monkeys (Macaca mulatta), video tasks, and implications for stimulus-response spatial contiguity. J. Comp. Psychol. 103, 32–38. https://doi.org/10.1037/0735-7036.103.1.32

Ryan, M.J., 2011. Replication in Field Biology: The Case of the Frog-Eating Bat. Science 334, 1229–1230. https://doi.org/10.1126/science.1214532

Savage-Rumbaugh, E.S., 1986. Ape Language: From Conditioned Response to Symbol. Columbia University Press, New York Chichester, West Sussex. https://doi.org/10.7312/sava90374

Schapiro, S.J., Perlman, J.E., Thiele, E., Lambeth, S., 2005. Training nonhuman primates to perform behaviors useful in biomedical research. Lab Anim. 34, 37–42. https://doi.org/10.1038/laban0505-37

Schmitt, V., Fischer, J., 2009. Inferential reasoning and modality dependent discrimination learning in olive baboons (Papio hamadryas anubis). J. Comp. Psychol. 123, 316.

Schmitt, V., Pankau, B., Fischer, J., 2012. Old World Monkeys Compare to Apes in the Primate Cognition Test Battery. PLoS ONE 7, e32024. https://doi.org/10.1371/journal.pone.0032024

Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., Carvalho, S., 2019. Chimpanzee face recognition from videos in the wild using deep learning. Sci. Adv. 5, eaaw0736. https://doi.org/10.1126/sciadv.aaw0736

Schrier, A.M., 1961. A Modified Version of the Wisconsin General Test Apparatus. J. Psychol. 52, 193–200. https://doi.org/10.1080/00223980.1961.9916519

Schwarz, D.A., Lebedev, M.A., Hanson, T.L., Dimitrov, D.F., Lehew, G., Meloy, J., Rajangam, S., Subramanian, V., Ifft, P.J., Li, Z., Ramakrishnan, A., Tate, A., Zhuang, K.Z., Nicolelis, M.A.L., 2014. Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. Nat. Methods 11, 670–676. https://doi.org/10.1038/nmeth.2936

Seed, A., Tomasello, M., 2010. Primate Cognition. Top. Cogn. Sci. 2, 407–419. https://doi.org/10.1111/j.1756-8765.2010.01099.x

Seyfarth, R.M., Cheney, D.L., Marler, P., 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science 210, 801–803.

Shamma, S., 2001. On the role of space and time in auditory processing. Trends Cogn. Sci. 5, 340–348. https://doi.org/10.1016/S1364-6613(00)01704-6

Shettleworth, S.J., 2009. Cognition, evolution, and behavior. Oxford university press.

Shumaker, R.W., Beck, B.B., 2003. Primates in question: the Smithsonian answer book. Smithsonian Books, Washington, D.C.

Sidowski, J.B., Spears, C., 1970. A versatile apparatus for measuring the frequencies and durations of animal and human responses. Behav. Res. Methods Instrum. 2, 235–238. https://doi.org/10.3758/BF03207187

Sinnott, J.M., Owren, M.J., Petersen, M.R., 1987. Auditory frequency discrimination in primates: Species differences (Cercopithecus, Macaca, Homo). J. Comp. Psychol. 101, 126–131. https://doi.org/10.1037/0735-7036.101.2.126

Skinner, B.F., 2005. Walden two. Hackett Publishing.

Skinner, B.F., 1965. Science and human behavior. Simon and Schuster.

Skinner, B.F., 1938. The behavior of organisms: an experimental analysis, The behavior of organisms: an experimental analysis. Appleton-Century, Oxford, England.

Spinelli, S., Pennanen, L., Dettling, A.C., Feldon, J., Higgins, G.A., Pryce, C.R., 2004. Performance of the marmoset monkey on computerized tasks of attention and working memory. Cogn. Brain Res. 19, 123–137. https://doi.org/10.1016/j.cogbrainres.2003.11.007

Suga, N., Ma, X., 2003. Multiparametric corticofugal modulation and plasticity in the auditory system. Nat. Rev. Neurosci. 4, 783–794.

Sun, G., Lyu, C., Cai, R., Yu, C., Sun, H., Schriver, K.E., Gao, L., Li, X., 2021. DeepBhvTracking: a novel behavior tracking method for laboratory animals based on deep learning. Front. Behav. Neurosci. 15.

Tinbergen, N., 1963. On aims and methods of ethology. Z. Für Tierpsychol. 20, 410–433.

Tolman, E.C., 1973. Cognitive maps in rats and men. Image Environ. Cogn. Mapp. Spat. Behav. 1948, 27–50.

Tomasello, M., 2000. Primate cognition: introduction to the issue. Cogn. Sci. 24, 351–361.

Tomasello, M., Call, J., 2011. Methodological Challenges in the Study of Primate Cognition. Science 334, 1227–1228. https://doi.org/10.1126/science.1213443

Vauclair, J., Fagot, J., 1993. Manual and hemispheric specialization in the manipulation of a joystick by baboons (Papio papio). Behav. Neurosci. 107, 210–214. https://doi.org/10.1037/0735-7044.107.1.210

Venkatraman, S., Jin, X., Costa, R.M., Carmena, J.M., 2010. Investigating Neural Correlates of Behavior in Freely Behaving Rodents Using Inertial Sensors. J. Neurophysiol. 104, 569–575. https://doi.org/10.1152/jn.00121.2010

Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P., Fragaszy, D., 2009. Selection of Effective Stone Tools by Wild Bearded Capuchin Monkeys. Curr. Biol. 19, 213–217. https://doi.org/10.1016/j.cub.2008.11.064

Von Frisch, K., 2013. The dance language and orientation of bees, in: The Dance Language and Orientation of Bees. Harvard University Press.

von Uexküll, J., O'Neil, J.D., 2013. A Foray into the Worlds of Animals and Humans: with A Theory of Meaning, Posthumanities. University of Minnesota Press.

Wakeley, J., 2008. Complex speciation of humans and chimpanzees. Nature 452, E3–E4. https://doi.org/10.1038/nature06805

Wakita, M., 2020. Common marmosets (Callithrix jacchus) cannot recognize global configurations of sound patterns but can recognize adjacent relations of sounds. Behav. Processes 176, 104136. https://doi.org/10.1016/j.beproc.2020.104136

Wakita, M., 2019. Auditory sequence perception in common marmosets (Callithrix jacchus). Behav. Processes 162, 55–63. https://doi.org/10.1016/j.beproc.2019.01.014

Ward, R.D., Winiger, V., Higa, K.K., Kahn, J.B., Kandel, E.R., Balsam, P.D., Simpson, E.H., 2015. The impact of motivation on cognitive performance in an animal model of the negative and cognitive symptoms of schizophrenia. Behav. Neurosci. 129, 292–299. https://doi.org/10.1037/bne0000051

Washburn, D.A., Hopkins, W.D., Rumbaugh, D.M., 1989. Automation of learning-set testing: The video-task paradigm. Behav. Res. Methods Instrum. Comput. 21, 281–284. https://doi.org/10.3758/BF03205596

Washburn, D.A., Rumbaugh, D.M., 1992. Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. Behav. Res. Methods Instrum. Comput. 24, 157–164. https://doi.org/10.3758/BF03203490

Wasserman, E.A., Wasserman, E.A., Zentall, T.R., 2006. Comparative cognition: Experimental explorations of animal intelligence. Oxford University Press, USA.

Watson, J.B., 1913. Psychology as the behaviorist views it. Psychol. Rev. 20, 158.

Watson, J.B., Rayner, R., 1920. Conditioned emotional reactions. J. Exp. Psychol. 3, 1.

Weed, M.R., Bryant, R., Perry, S., 2008. Cognitive development in macaques: Attentional set-shifting in juvenile and adult rhesus monkeys. Neuroscience 157, 22–28. https://doi.org/10.1016/j.neuroscience.2008.08.047

Wheeler, W.M., 2015. The social insects: their origin and evolution. Routledge.

Wienicke, A., Häusler, U., Jürgens, U., 2001. Auditory frequency discrimination in the squirrel monkey. J. Comp. Physiol. A 187, 189–195.

Wilson, D.E., Reeder, D.M., 2005. Mammal species of the world: a taxonomic and geographic reference. JHU press.

Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abraira, V.E., Adams, R.P., Datta, S.R., 2015. Mapping sub-second structure in mouse behavior. Neuron 88, 1121–1135.

Winer, J.A., Schreiner, C., 2005. The inferior colliculus: with 168 illustrations. Springer, New York, NY.

Wolfensohn, S., Honess, P., 2008. Handbook of primate husbandry and welfare. John Wiley & Sons.

Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation.

Zeng, F.-G., Galvin III, J.J., 1999. Amplitude mapping and phoneme recognition in cochlear implant listeners. Ear Hear. 20, 60–74.

Zeng, F.-G., Rebscher, S., Harrison, W., Sun, X., Feng, H., 2008. Cochlear implants: system design, integration, and evaluation. IEEE Rev. Biomed. Eng. 1, 115–142.

Zentall, T.R., Wasserman, E.A., 2012. The Oxford handbook of comparative cognition. Oxford University Press.

Zoloth, S.R., Petersen, M.R., Beecher, M.D., Green, S., Marler, P., Moody, D.B., Stebbins, W., 1979. Species-specific perceptual processing of vocal sounds by monkeys. Science 204, 870–873.

Zwicker, E., Fastl, H., 2013. Psychoacoustics: Facts and models. Springer Science & Business Media.

Chapter 2

Flexible auditory training, psychophysics, and enrichment of common marmosets with an automated, touchscreen-based system

Calapai, A.*, Cabrera-Moreno, J.*, Moser, T., Jeschke, M.
***** Equal contribution

Nature Communications 13, 1648 (2022)
DOI: https://doi.org/10.1038/s41467-022-29185-9

Chapter 2 describes a series of experiments that demonstrate that common marmosets (*Callithrix jacchus*) can be trained to solve audio-visual discrimination tasks using a 2 or 3-alternative choice paradigm to collect psychoacoustic data. All this in an autonomous and unsupervised manner directly in their home-cage without the requirement of social separation nor the control for food or water intake.

Author contributions:

AC, MJ conceived the study. AC, JCM, MJ designed and developed hardware and software. JCM performed the experiments with the help of MJ. JCM curated the data and AC analyzed the data and generated the figures. AC, JCM, MJ discussed and interpreted the data. TM provided funding, discussions and feedback on all aspects of the study. MJ supervised the study. AC, JCM, MJ wrote the paper with input from TM

Abstract

Devising new and more efficient protocols to analyze the phenotypes of non-human primates, as well as their complex nervous systems, is rapidly becoming of paramount importance. This is because with genome-editing techniques recently adopted for non-human primates, new animal models for fundamental and translational research were established. One aspect in particular, namely cognitive hearing, has been difficult to assess compared to visual cognition. To address this, we devised autonomous, standardized, and unsupervised training and testing of auditory capabilities of common marmosets with a cage-based standalone, wireless system. All marmosets tested voluntarily operated the device on a daily basis and went from naïve to experienced at their own pace and with ease. Through a series of experiments, here we show, that animals autonomously learn to associate sounds with images; to flexibly discriminate sounds, and to detect sounds of varying loudness. The developed platform and training principles combine in-cage training of common marmosets for cognitive and psychoacoustic assessment with an enriched environment that does not rely on dietary restriction or social separation, in compliance with the 3Rs principle.

2.1 Introduction

In recent years non-human primates (NHP) have seen increased interest as animal models for human diseases due to the advent of transgenic primates and genome-editing technologies (Feng et al., 2020; Sasaki et al., 2009). As NHP are closer to humans with respect to e.g. physiology, cognition, genetics and immunology (Hawash et al., 2021; Rogers and Gibbs, 2014; Tay et al., 2009), results from NHP studies investigating cognition are likely more representative for the situation in humans.

In visual neuroscience, attention, object formation, categorization and other aspects of cognition are extensively studied . In auditory neuroscience, several studies have also used different tasks (e.g. 2-alternative forced choice, go-no go) and cognitive functions (such as memory, categorization, reward processing(Archakov et al., 2020; Brosch et al., 2004; Margiotoudi et al., 2019; Wikman et al., 2019)). In general, though, studies in auditory cognition are lagging behind those of visual cognition with respect to overall sophistication of methods, experiments and task complexities. One factor for this is the common observation that monkeys have been notoriously difficult to train in the auditory domain, and generally display a bias towards vision. For example, it has been shown that baboons can easily learn to locate food items based on visual but not auditory cues (Schmitt and Fischer, 2009). Among other results this surprising failure at such a seemingly simple auditory task has led to the suggestion that inferential reasoning is modality specific.

However, investigations into auditory capabilities and cognition increase in scope as non-human primates have become genetically tractable organisms (Drummer et al., 2021; Feng et al., 2020; Kumita et al., 2019; Park et al., 2016; Sasaki et al., 2009). Notably, the common marmoset (*Callithrix jacchus*) has become a valuable model for biomedical research in general and the neurosciences in particular (Marini et al., 2018; Okano, 2021; Tokuno et al., 2012). Factors such as the relative ease of breeding, early sexual maturation and short life span (Abbott et al., 2003; Schultz-Darken et al., 2016) have contributed to the rapid generation of genetic models of human mental and neurological diseases in marmosets (Okano and Kishi, 2018; Sasaki et al., 2009; Shen, 2013; Tomioka et al., 2020). While generally marmoset training is lacking behind the sophistication of cognitive NHP experiments traditionally performed with macaques, auditory capabilities of marmosets have been investigated extensively (Choi et al., 2015; Osmanski et al., 2016; Osmanski and Wang, 2011; Risueno-Segovia and Hage, 2020; Song et al., 2016; Takahashi et al., 2013). Furthermore, marmosets have now also become the go-to NHP model for hearing loss and cochlear implant research (Hosoya et al., 2016; Johnson et al., 2016, 2012; Keppeler et al., 2021). In the near future many more transgenic primate models will be developed which requires extensive phenotyping such as is standard for rodent models (Brown and Moore, 2012). Phenotyping will need to investigate large number of subjects in a standardized and experimenter/observer independent manner (Aoki et al., 2017; Bohlen et al., 2014; Cibulski et al., 2014; Rivalan et al., 2017; Schmitt et al., 2014; Schubiger et al., 2015; Sorge et al., 2014). In addition, increased awareness for species specific ethical demands asks for refinement of experimentation techniques as much as possible (Prescott, 2020; Prescott et al., 2017). This has led to efforts developing home-cage, computer-based cognitive training of NHPs focusing on the visual domain (Andrews and Rosenblum, 1994; Berger et al., 2018; Bethell et al., 2019; Calapai et al., 2017a; Crofts et al., 2016; Fagot and Paleressompoulle, 2009; Gazes et al., 2013; Kangas and Bergman, 2012; Mandell and Sackett, 2008; Nakamura et al., 2018; O'Leary et al., 2018; Richardson et al., 1990; Sadoun et al., 2018; Spinelli et al., 2004; Takemoto et al., 2011, 2011; Washburn et al., 1989; Washburn and Rumbaugh, 1992).

To achieve comparable efforts in the auditory domain, there is a need for automatic, unsupervised cage-based training and testing of auditory tasks. Towards this goal, we built a standalone wireless device for auditory training and testing of common marmosets, directly in their own cage. The system, termed marmoset experimental behavioral instrument (MXBI), is mostly comprised of off-the-shelf or 3d printed components, is entirely programmed in Python, and based on the Raspberry Pi platform, for maximum flexibility of use, openness, and to allow for easy adaptation by others. The MXBI is set up with a server / client configuration in mind; and capable of animal tagging by means of radio-frequency identification (as in rodent systems (Schaefer and Claridge-Chang, 2012)), which ultimately allows scalable, standardized, automated, and unsupervised training and testing protocols (AUT in short, from (Berger et al., 2018)) in socially housed animals. Moreover, the MXBI and the procedures we describe contribute to the efforts of refining cognitive and environmental enrichments of NHP in human care. Further, we report results from a set of four experiments: 1) an algorithm-based procedure for gradually and autonomously training naïve animals to the basics of a 2-Alternative-Choice (2AC visual task); 2) an audio-visual association experiment where a conspecific call is contrasted to an artificial acoustic stimulus; 3) a generalization experiment assessing the flexibility of the acquired discrimination behavior to other stimuli; 4) and a psychoacoustic detection experiment for quantifying hearing thresholds in a cage-based setting. We show that marmosets can be trained to flexibly perform psychoacoustic experiments on a cage-based touchscreen device, via an automated and unsupervised training procedure that require no human supervision and do not rely on fluid or food control, nor social separation.

2.2 Results

In this study 14 adult common marmosets of either sex and housed in pairs participated across one *initial training* phase and four autonomous cage-based experiments. Animals were generally trained in pairs on auditory tasks with a single MXBI attached to the animals' home cage and without fluid or social restrictions (Figure 2.1A). Aside from the *initial training* (see below) all sessions ran autonomously, while an RFID module identified the animals and an algorithm controlled the individualized, performance-based progression in difficulty (see methods: Automated unsupervised training (AUT)).

Initial training

The goal of the *initial training* was to instruct naïve animals to interact with the touchscreen to receive liquid reward (Arabic gum or marshmallow solution) from the device's mouthpiece. The training was divided into three sequential steps: first, habituation to the device (supplementary video 1); second, forming a mouthpiece-reward association (supplementary video 2), and finally, a touch-to-drink phase (supplementary video 3.1 and 3.2). All animals started exploring the device from the very first session. During the touch-to-drink phase, a mesh tunnel was introduced inside the device (Figure 2.1A), to allow only one animal at a time inside the MXBI. Animals were encouraged to enter the tunnel and reach the touchscreen by placing small pieces of marshmallows or arabic gum along the tunnel, on the mouthpiece, and on the screen. After the *initial training* was concluded (mean = 6 ± 1.4 sessions, median = 6, Table 2.1), animals were introduced to the automated procedure gradually bringing them from naïve to experienced in discrimination as well as detection-based psychoacoustic tasks.

General engagement on the MXBI across all autonomous experiments

Individual animals engaged with the MXBI in different amounts with the median number of trials varying between 31 and 223. On average 116 trials per session (IQR = Q3-Q1 = 192) were performed (Figure 2.1B, Table 2.1). While half of the animals had less than 10 % of sessions without a single trial (median = 10.7 %, IQR = 16.8 %) two animals displayed more than 30 % of sessions without performing a trial. Controlling for session duration (Figure 2.1C). We found no significant correlation between the total number of trials performed by each animal and session number (Partial Pearson correlation controlling for session duration; adjusted $r^2 = 0.05$, p-value: 0.1, N = 802; CI = -0.01, 0.13), suggesting that the level of engagement remained consistent across sessions. Qualitatively, animals tended to engage consistently throughout a session as indicated by the distribution of trial onset times (Figure 2.1E). Consequently, the median time point at which half of the trials were performed was 0.52 of the session’s duration (Figure 2.1D).

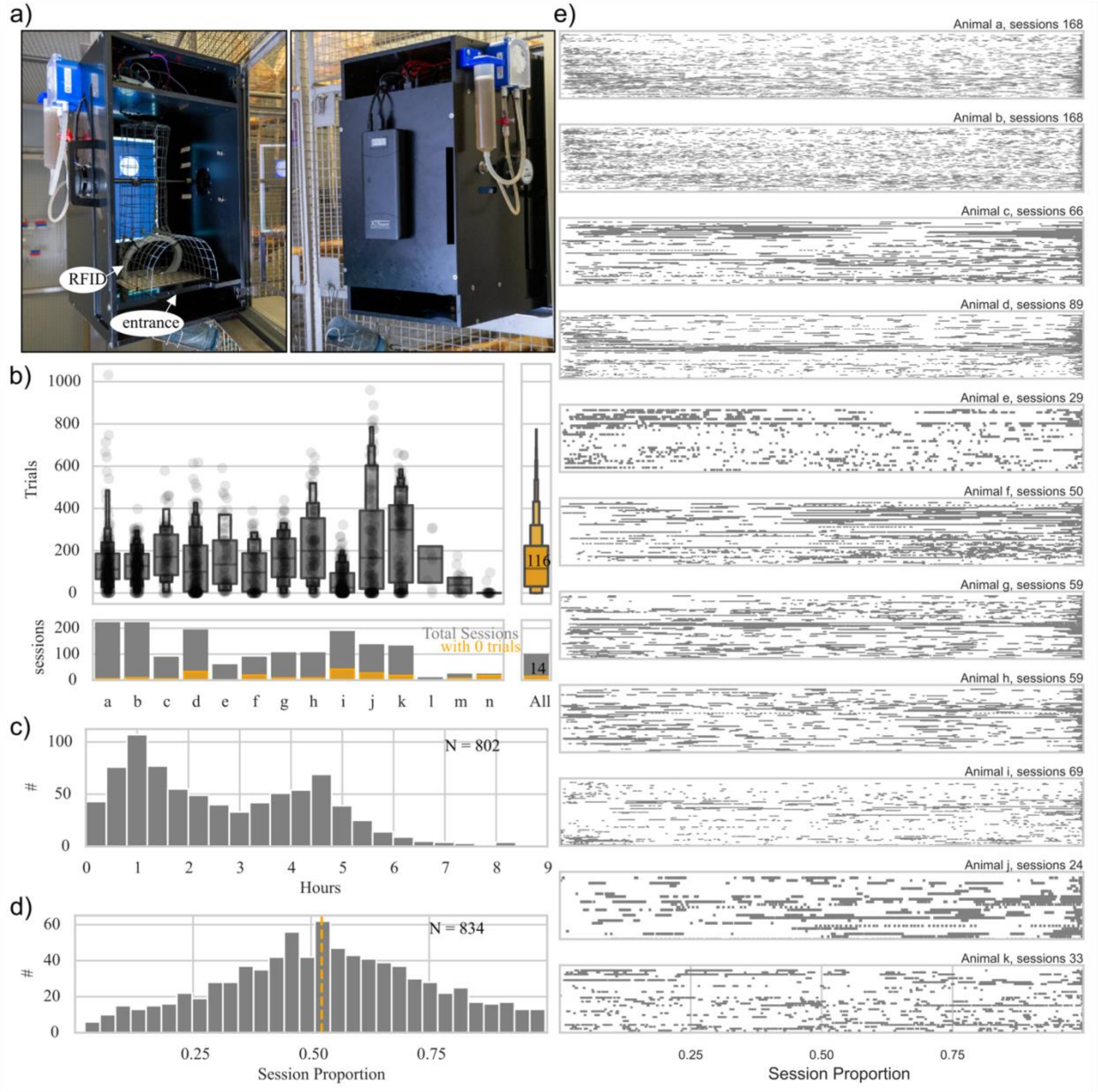


Figure 2.1. General Engagement on the MXBI across all autonomous experiments. ***a)*** *MXBI device attached to a cage in the animal facility. Left—opened for experimenter access to the inside of the device, Right—closed.* ***b)*** *Letter-value plots of the number of trials performed in each session, dark gray: for each animal individually, orange: average distribution with all animals considered. The central box defines the median and 25th up to 75th percentile. Successively narrower boxes are drawn between the 1/8*100th and 7/8*100th, the 1/16th and 15/16*100th, and so on percentile. The total number of trials per animal can be found in Table 1. The number of sessions per animal and the average across animals are plotted below. Sessions without a single trial performed are given in orange.* ***c)*** *Distribution of the duration of all sessions.* ***d)*** *Distribution of all median timestamps as a function of session proportion. The dashed orange line indicates that across sessions half of the trials were performed within 52% of the session duration.* ***e)*** *For each of the animals with more than 3000 trials, each trial of each session with more than 10 interactions (shared ordinated) is plotted with respect to its timestamp normalized by the session duration.*

ID	**Cage mate ID**	**Characteristics**			**Initial training (Sessions)**			**Trials (across all tasks)**	**Sessions (across all tasks)**	**Sessions with 0 trials**
		Sex	**Age [months]**	**Weight [g]**	**Habituation**	**Mouthpiece association**	**Touch to drink**			
a	b	f	41	-	-	-	3	31081	220	5
b	a	m	36	-	-	-	4	28222	220	9
c	f	m	24	415	1	2	3	16181	87	5
d	i	f	84	375	1	2	2	25941	192	33
e	l	f	26	423	2	1	3	9822	58	0
f	c	m	84	386	1	2	2	9901	87	19
g	h	m	29	476	2	3	2	17296	104	8
h	g	m	32	354	2	2	3	23157	104	8
i	d	m	26	390	0.5	0.5	1	10265	186	42
j	k	m	33	446	3	2	2	32585	135	27
k	J	f	31	388	2	2	2	33424	130	18
l	e	f	32	471	2	2	3	1212	8	0
m	n	m	45	-	-	-	-	975	22	7
n	m	f	31	366	2	2	3	168	22	17

Table 2.1. Characteristics and statistics of all animals involved in the experiments. "Characteristics" columns report the sex (S), age in months (A), and weight in grams (W) of each animal at the start of experiment. Initial training columns report the number of sessions required for the shaping stages habituation (H), mouthpiece-reward association (M) and "touch-to-drink" (T). Columns: Trials, Sessions and Session with 0 trials report the statistics for the corresponding panel in Figure 2.1 regarding the total number of trials ("Trials") and the total number of Sessions collected for each animal ("Sessions"). The column "0 Trials" summarizes the amount of sessions without interactions. The Initial Training was not systematically quantified for animals a, b. Weight information at the start of the sessions was also not available for those animals.

Automated unsupervised training (AUT)

An automated and unsupervised training protocol (AUT (Berger et al., 2018)) was implemented to train naive marmosets at their own pace on the basics of a 2AC visually-guided task. In order to identify the appropriate parameters upon which to build such autonomous procedure we first designed and tested multiple AUT versions with a subset of 9 animals (described in supplementary tables S2.1 and S2.2). The resulting final versions of the protocols (AUTs 8, 9, and 10), were then tested with 4 naïve animals (animals f, k, c, and d). The AUT procedure was comprised of 4 milestones – 1) decrease of the size of a visual stimulus (trigger) to be touched for reward, 2) change of position of visual stimulus, 3) introduction of sound and delayed presentation of a visual target, 4) introduction of a second visual target as a distractor – that unfolded through a total of 48 dynamic steps (Figure 2.2; Figure S2.4C). During each session the transitions between steps and milestones was based on the animal's performance in a sliding window of trials (hit rate of 80% to advance, 20% to retreat; Figure S2.4D). Figure 2.2C shows the hit rate across individual steps and milestones for the 4 naïve animals that only performed the final versions of the AUT. While the procedure was designed to encourage a smooth transition from step to step, certain steps (and thus milestones) required more trials to be accomplished. As a consequence, the hit rate calculated across animals varies substantially as function of AUT step (Figure 2.2C). Due to animals learning at different paces and performing different number of trials, we quantified the progression through the AUT as a function of the percentage of total trials completed by each animal (Figure 2.2D). This allowed us to visualize and compare learning progress across animals with inherently different working pace on a common frame of reference. Both the total amount of trials (expressed by line thickness in Fig. 2.2D) needed to complete the AUT and the learning curves throughout the AUT vary substantially across animals (Figure 2.2E) in the middle portion of the AUT, during which the stimulus changed position on the screen and an acoustic stimulus was introduced. Starting from the introduction of sound (milestone 3) we introduced timeouts (grey screen) to provide further feedback on wrong trials. Analysis of inter-trial-intervals (ITIs) trials revealed shorter average ITIs after correct vs. wrong trials suggesting an effect of timeouts on animal behavior (Figure S2.3 and Table S2.4).

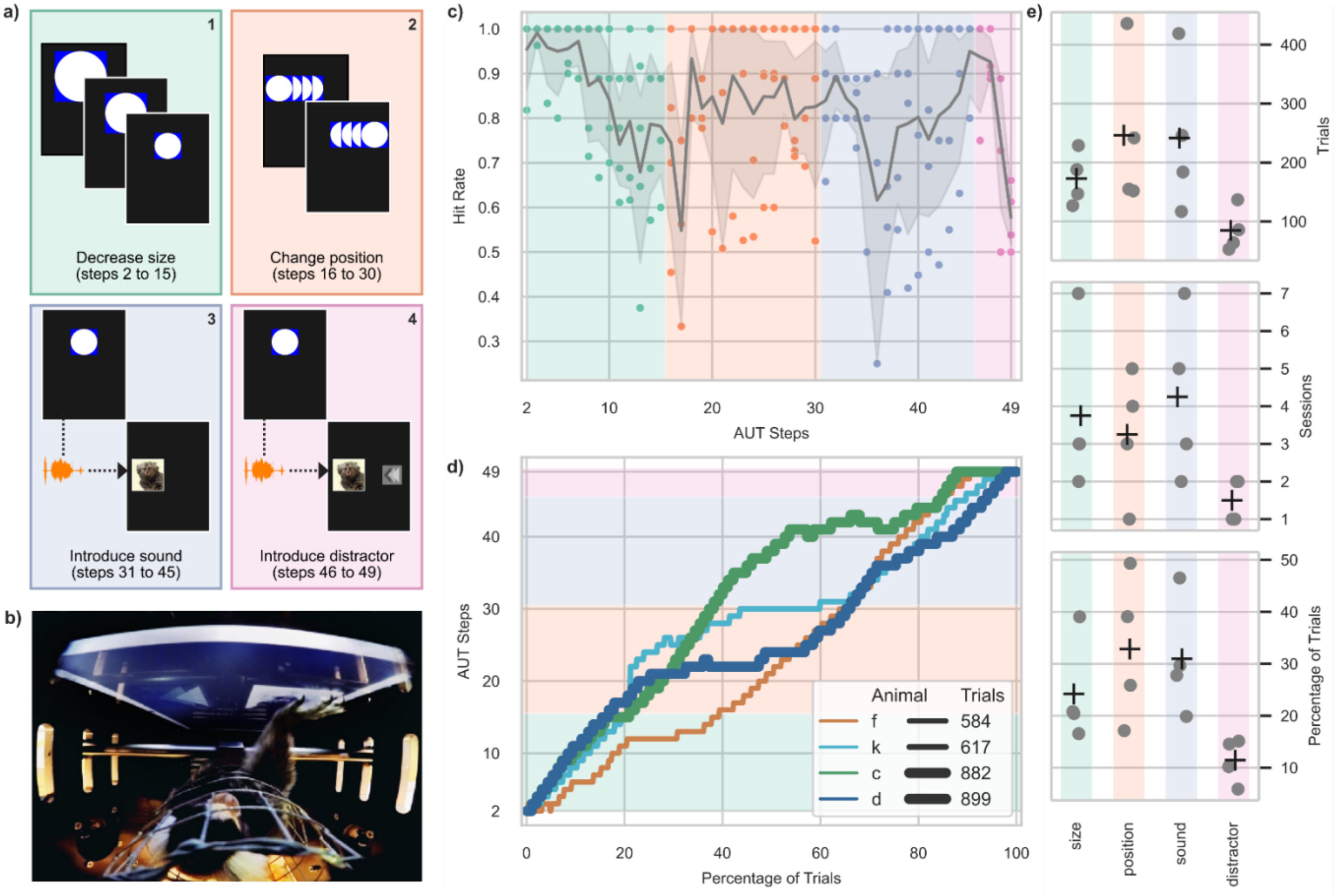


Figure 2.2. Automated unsupervised training (AUT) performance across representative animals. a) *Schematics of the four main milestones of the final AUT protocol.* ***b)*** *Picture taken with an external high-resolution wide-angle camera, where an animal can be seen completing a trial.* ***c)*** *Average hit rate as a function of steps (gray area represents the 95% confidence interval of the mean across animals) across the four animals considered in this analysis.* ***d)*** *Percentage of trials spent on each AUT step and milestones (shaded background) with line thickness indicating the total amount of trials performed by the corresponding animal.* ***e)*** *Distributions of number of trials, number of sessions, and percentage of total trials for each milestone across the four naïve animals (crosses represent average values).*

Audio-visual association

Next, we tested whether animals would generalize from the visually-guided 2AC task introduced via the AUT procedure to an acoustically-guided 2AC discrimination. In this experiment animals were required to discriminate between a conspecific juvenile call (in the following referred to as voc), and a pure tone (simple train – sTr – chosen for individual animals from a range between 1.5 and 3.5 kHz), by selecting one of two visual stimuli permanently associated with each sound (supplementary Video 4). 5 out of 9 animals successfully learned to discriminate between the sTr and the voc by selecting a geometric pattern or a conspecifics face, respectively (Figure 2.3A and C). The remaining 4 animals performed at chance level. To disentangle if these animals were unable to solve the task or maybe were unwilling to perform above chance, we devised a 3 alternative-choice (3AC; see methods) version of the same task (Figure 2.3B and C) and tested 2 of these animals and 2 additional animals who had failed a different control condition (see supplementary material: Artificial Discrimination, Figure S2.1, S2.2). In the 3AC task, all 4 animals performed the task significantly above chance (Binomial test, pot-hoc corrected for multiple comparisons; Table 2.2). Taken together these results demonstrate that 9 out of 11 animals learned the audio-visual association. The remaining two animals that did not learn the 2AC discrimination were assigned to a different project and were not tested on the 3AC version. Additionally, 7 out of 9 animals who accomplished the discrimination task exhibited significantly longer reaction times in responding to the target in voc vs sTr trials (Figure 2.3D; Table 2.2), indicating that the animals behaved differently for different acoustic stimuli.

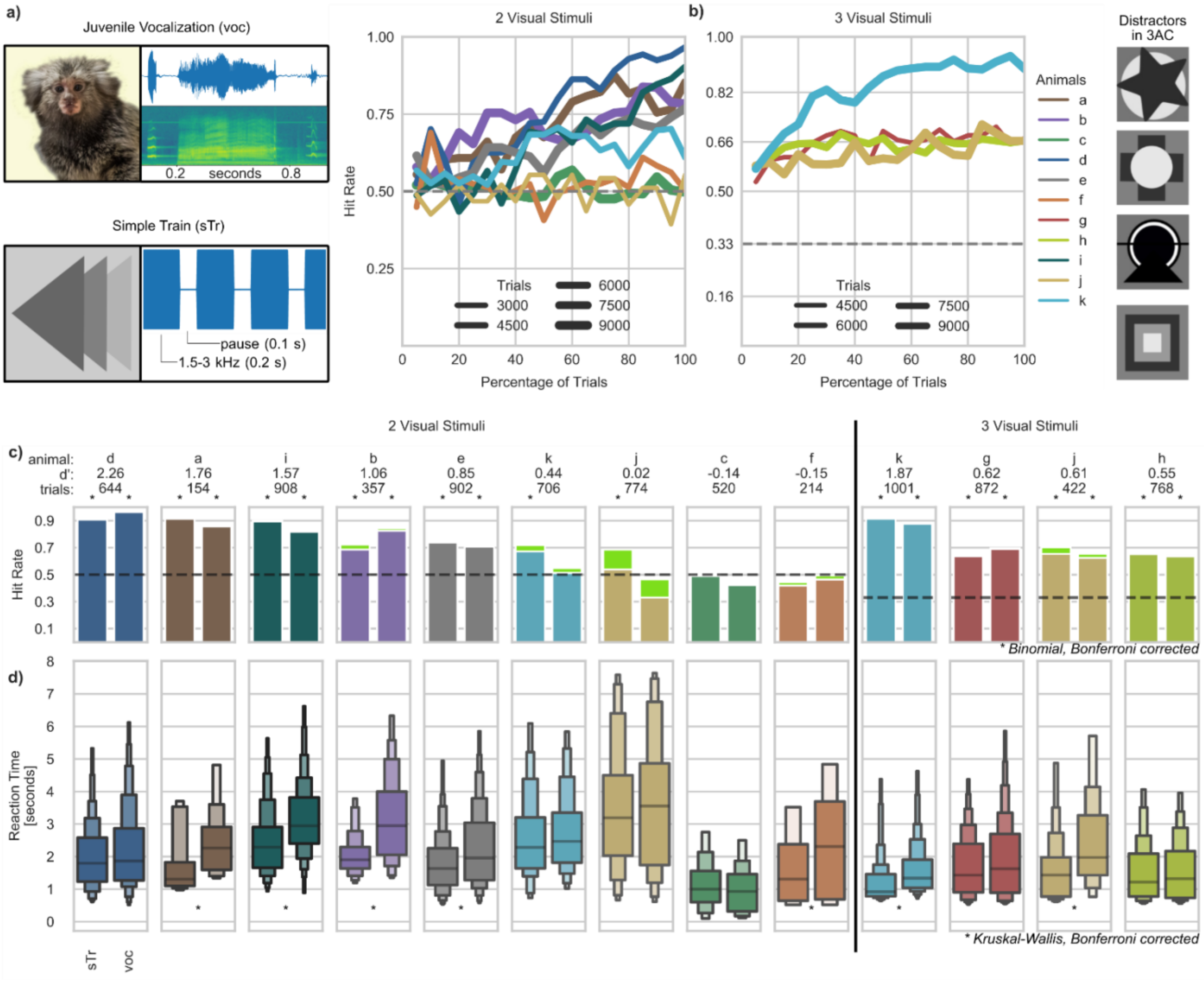


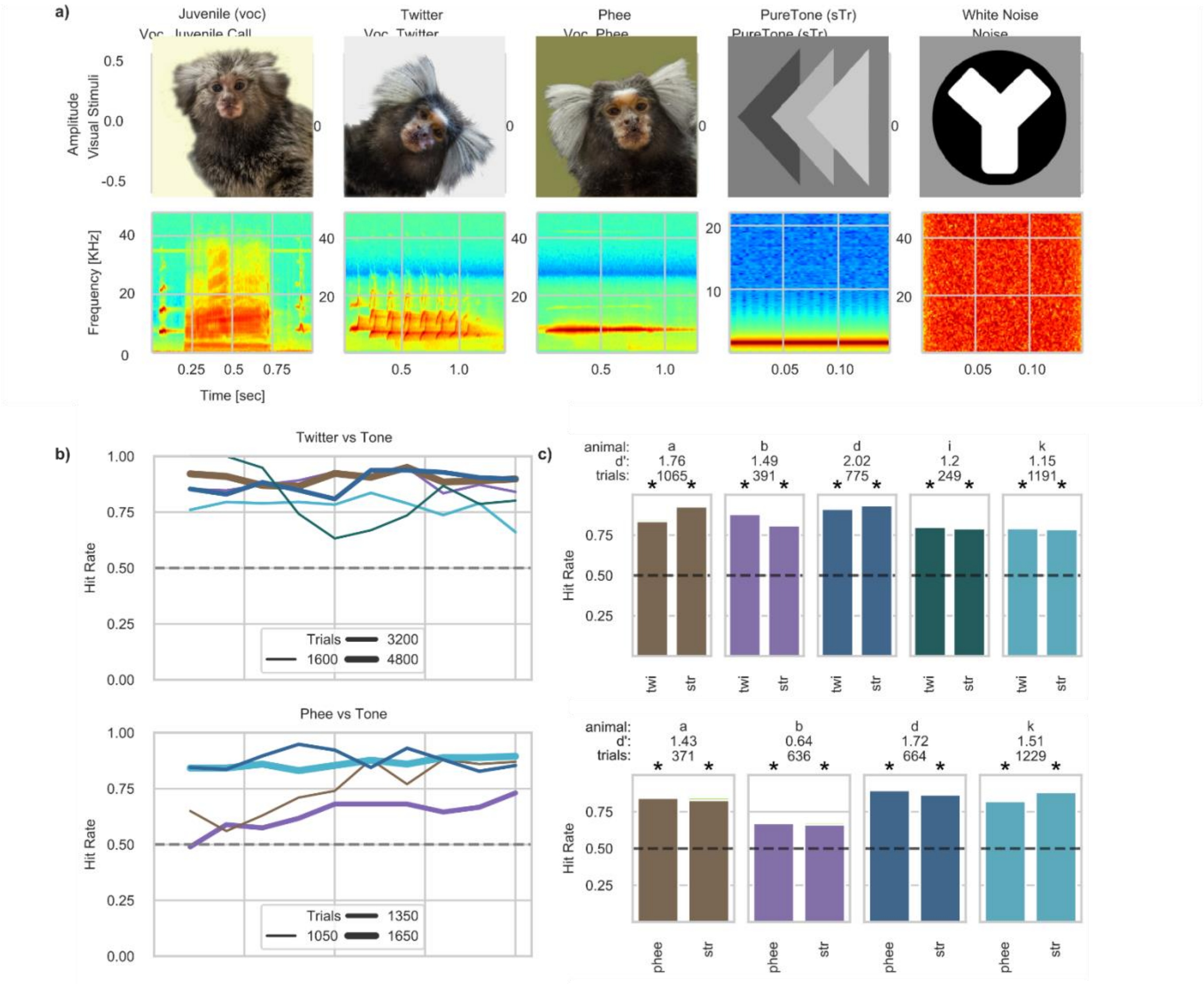
Figure 2.3.Stimuli and results from the audio-visual association experiment. a), b) *Visual and acoustic stimuli combinations used and hit rate as a function of percentage of trials performed, for different animals (colored lines) and across tasks. Hit rate, as a function of the percentage of trials performed by each animal, is grouped into bins of 5%. Line thickness represents the number of trials of each animal in each panel. Dashed lines at 0.5 and 0.33 represent the chance level for the two tasks.* ***c)*** *Hit rate across the last 5 sessions as a function of stimulus type ("sTr" for the pure tone stimulus, "voc." for the juvenile vocalization; green bars indicate ignored trials), with corresponding number of trials and sensitivity index (d'). Stars represent significance reached for the given stimulus at a Bonferroni post-hoc corrected Binomial test (one-sided test).* ***d)*** *Letter-value plots of the reaction times plotted for each stimulus type separately. The central box defines the median and 25th up to 75th percentile. Successively narrower boxes are drawn between the 1/8*100th and 7/8*100th, the 1/16*100th and 15/16*100th, and so on, percentile. Stars represent significant statistical difference in reaction times between the two stimuli at a Bonferroni post-hoc corrected Kruskal–Wallis Test (one-sided test). Statistics and N number for panels (**c**) and (**d**) are given in Table 2.*

Animal	Task	Stimulus	Hitrate	Trials	d'	**Binomial test on performance (Figure 2.3C)**			**Kruskal-Wallis test on Reaction Times (Figure 2.3D)**					
						N (w/o ignored)	Degrees of freedom	Binomial Test (adjusted p-value)	Median	IQR	N	Degrees of freedom	Test Statistics	Kruskal-Wallis (adjusted p-value)
a	2AC	sTr	0.91	70	1.76	70	1	**3.18E-12**	1.3065	0.72725	64	1	28.22	**2.81E-06**
		voc	0.86	84		83	1	**5.09E-11**	2.2635	1.3195	72	1		
b	2AC	sTr	0.69	191	1.06	184	1	**1.08E-07**	1.904	0.6625	131	1	54.65	**3.74E-12**
		voc	0.83	166		163	1	**2.71E-18**	2.95	1.944	137	1		
c	2AC	sTr	0.49	241	-0.14	238	1	1	0.999	0.95925	118	1	1.52	1
		voc	0.42	279		278	1	1	0.928	1.14275	118	1		
d	2AC	sTr	0.91	316	2.26	316	1	**2.01E-53**	1.797	1.3515	287	1	2.73	1
		voc	0.96	328		326	1	**6.38E-79**	1.864	1.60225	316	1		
e	2AC	sTr	0.74	444	0.85	444	1	**2.15E-23**	1.634	1.132	328	1	21.99	**7.14E-05**
		voc	0.71	458		453	1	**2.59E-19**	1.9605	1.76475	324	1		
f	2AC	sTr	0.42	117	-0.15	114	1	1	1.308	1.737	49	1	3.84	1
		voc	0.46	97		94	1	1	2.308	3.01	45	1		
g	3AC	sTr	0.64	437	0.62	432	1	**3.32E-08**	1.4295	1.493	278	1	2.20	1
		voc	0.69	435		432	1	**5.18E-15**	1.626	1.805	300	1		
h	3AC	sTr	0.65	388	0.55	384	1	**6.21E-09**	1.216	1.318	253	1	0.27	1
		voc	0.63	380		375	1	**4.66E-07**	1.319	1.436	241	1		
i	2AC	sTr	0.9	459	1.57	457	1	**3.11E-73**	2.289	1.2525	411	1	83.00	**2.14E-18**
		voc	0.82	449		444	1	**3.19E-45**	2.946	1.4155	367	1		
j	2AC	sTr	0.54	299	0.02	255	1	**0.0004**	3.191	2.469	161	1	0.03	1
		voc	0.33	475		411	1	1	3.554	3.128	157	1		
j	3AC	sTr	0.65	222	0.61	211	1	**7.25E-07**	1.433	1.204	145	1	26.39	**7.26E-06**
		voc	0.63	200		194	1	**0.0009**	1.972	1.841	125	1		
k	2AC	sTr	0.67	320	0.44	305	1	**7.70E-12**	2.282	1.567	215	1	2.64	1
		voc	0.51	386		372	1	1	2.467	1.53675	198	1		
k	3AC	sTr	0.91	503	1.87	498	1	**4.86E-92**	0.919	0.67875	460	1	120.62	**1.20E-26**
		voc	0.88	498		493	1	**4.01E-73**	1.337	0.869	437	1		

Table 2.2, Summary statistics for the audio-visual association across animal and stimuli (Figure 2.3C, D). Columns "Binomial test on performance (Figure 2.3C)" report information regarding the statistical deviations of performance (across stimuli and task type) from a theoretically expected distribution of observations (one-sided), with p-values adjusted with a post-hoc Bonferroni correction for multiple comparisons. D-prime value is provided as indication of the sensitivity of each animal on given task. Columns under "Kruskal-Wallis test on Reaction Times (Figure 2.3D)" report information regarding the statistical difference of the reaction time to the sTr and the voc stimuli, with p-values adjusted with a post-hoc Bonferroni correction for multiple comparisons. Significant values are indicated in bold font.

Generalization to novel stimuli

With the 5 best performing animals in the audio-visual association, we assessed whether animals would be able to generalize the acquired discrimination to three novel stimuli (Figure 2.4A): two different types of vocalizations – an adult marmosets' Phee and a Twitter – and a white noise sound. On one hand, all 5 animals quickly learned to discriminate the Twitter and the Phee when contrasted to the sTr (Figure 2.4B and C). On the other hand, when two new stimuli were contrasted with each other animals displayed lower hit rates. In the white noise vs Twitter condition, 3 animals acquired the discrimination; 1 animal displayed a bias towards the twitter it had previously learned; and for 1 animal the performance fluctuated between 0.6 and 0.75 in the sessions prior to the last 2 in which it was not significantly different from chance. When the juvenile vocalization (voc) was juxtaposed to the Twitter only 1 animal significantly performed above chance and another performed significantly above chance only for the Twitter. Animals seemed to find it more difficult discriminating between vocalizations, despite having already learned and successfully discriminated both from other stimuli extensively (see Table 2.3). We interpret this result as an indication that vocalization stimuli (voc, Twitter, and Phee) carry a distinctive meaning to the animals compared to more artificial stimuli (tones or white noise). This could in fact explain why animals readily discriminate them when contrasted to artificial stimuli but do not display significant discrimination between multiple vocalizations. Note that Animal i was not quantified in the voc vs Twitter and in the Phee vs sTr condition due to a limited number of trials (less than 50 trials in each task).


a)
Juvenile (voc)
Twitter
Phee
PureTone (sTr)
White Noise
Voc Juvenile Call
Voc Twitter
Voc Phee
PureTone (sTr)
Noise
Amplitude
Visual Stimuli
0.5
0.0
-0.5
Frequency [KHz]
40
20
10
0
0.25
0.5
0.75
1.0
0.05
0.10
Time [sec]
b)
Twitter vs Tone
Phee vs Tone
Hit Rate
1.00
0.75
0.50
0.25
0.00
Trials
1600
3200
4800
1050
1350
1650
c)
animal:
d':
trials:
a
1.76
1065
b
1.49
391
d
2.02
775
i
1.2
249
k
1.15
1191
twi
str
a
1.43
371
b
0.64
636
d
1.72
664
k
1.51
1229
phee
str
*

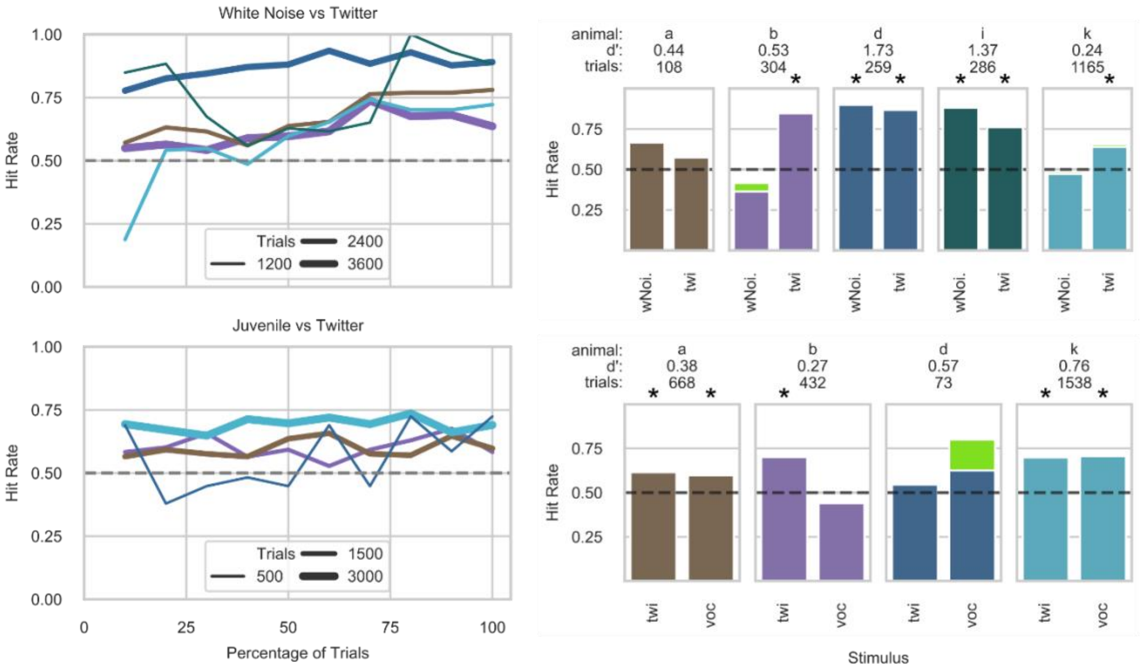


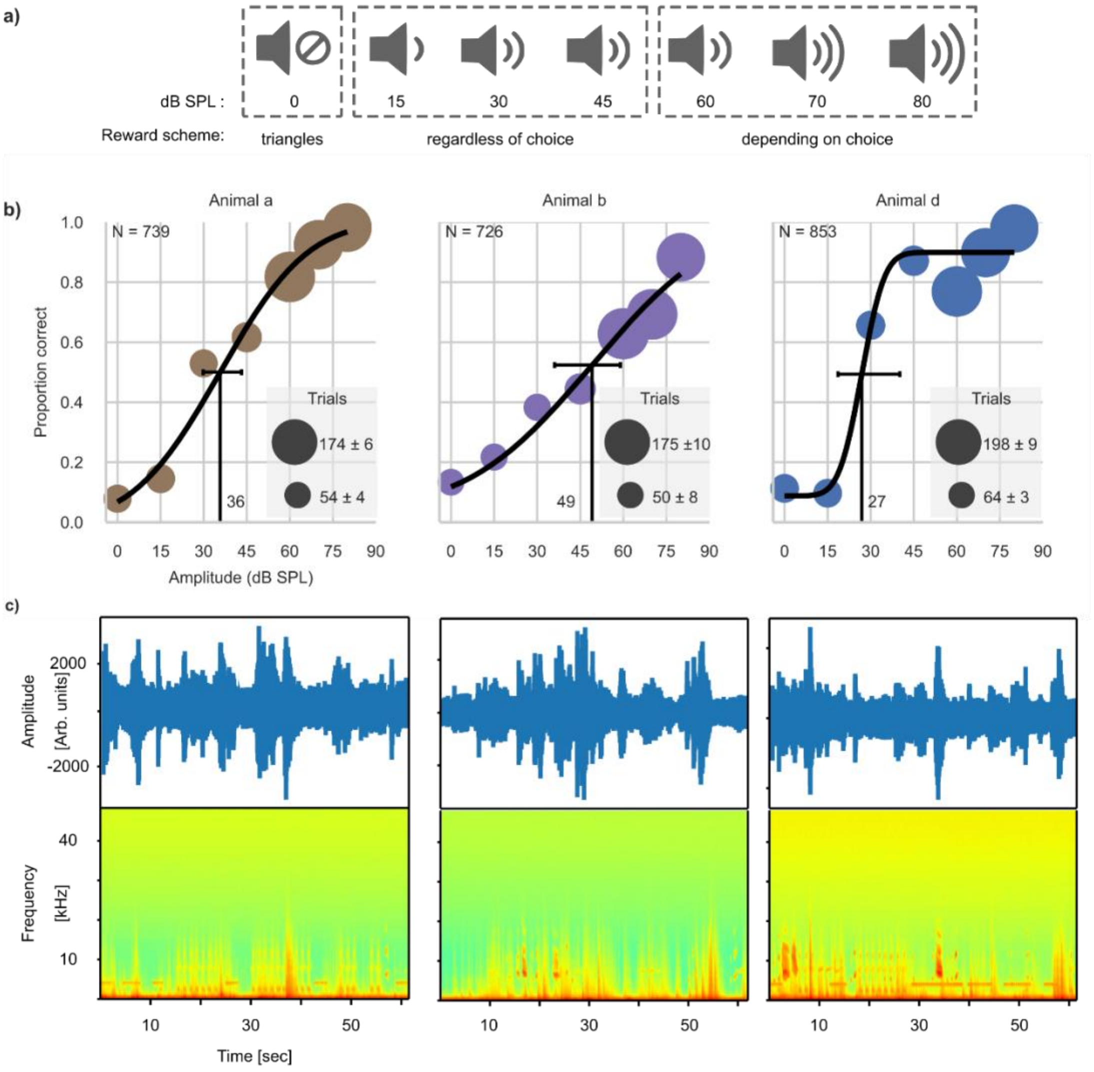
Figure 2.4. Generalization of audio-visual association to novel stimuli. a) *Representative visual stimuli and spectrograms for all five stimuli used in the experiment, paired column wise. The juvenile vocalization (voc) and the pure tone (simple train—sTr) are the same stimuli used in the previous experiments (Fig. 2).* ***b)*** *Hit rate as a function of percentage of trials (10% bin) across of the five animals and the four tasks, with line thickness representing the total amount of trials of each animal at each task.* ***c)*** *Hit rate as a function of stimulus in the last three sessions (eight sessions for animal d in the condition Juvenile vs Twitter and nine for animal k in White Noise vs Twitter), with corresponding number of trials and sensitivity index (d'). Star represents significance reached at a Bonferroni post-hoc corrected Binomial test for the corresponding stimulus (one-sided test). Dashed lines across all plots represent the 50% chance threshold. Green indicates ignored trials. The performance of animal d in the task Juvenile vs Twitter and of animal i in the task Twitter was Tone were based on eight sessions and nine sessions, respectively, instead of 3 (like the rest of the animals and tasks). This was necessary to consider a number of total trials higher than 40 and thus increase the statistical reliability of testing the performance of each animal against chance.*

Animal	Task	Sound	Trials	Hit Rate	Binomial test on performance (Figure 2.4C)		
					N	Degrees of freedom	Adjusted p-value
a	Twitter vs Tone	Tone	530	0.93	529	1	**3.75E-100**
	Phee vs Tone	Tone	186	0.83	183	1	**1.78E-20**
	Juvenile vs Twitter	Juvenile	333	0.6	333	1	**8.72E-03**
	Phee vs Tone	Phee	185	0.84	184	1	**1.94E-21**
	Twitter vs Tone	Twitter	535	0.84	528	1	**5.05E-61**
	Juvenile vs Twitter	Twitter	335	0.61	334	1	**4.65E-04**
	Noise vs Twitter	Twitter	54	0.57	54	1	1
	Noise vs Twitter	Noise	54	0.67	54	1	**3.97E-01**
b	Twitter vs Tone	Tone	200	0.81	198	1	**5.52E-19**
	Phee vs Tone	Tone	320	0.66	315	1	**1.59E-08**
	Juvenile vs Twitter	Juvenile	218	0.44	217	1	1
	Phee vs Tone	Phee	316	0.67	314	1	1
	Twitter vs Tone	Twitter	191	0.88	190	1	**9.98E-28**
	Juvenile vs Twitter	Twitter	214	0.7	214	1	**7.76E-08**
	Noise vs Twitter	Twitter	150	0.85	149	1	**7.60E-18**
	Noise vs Twitter	Noise	154	0.36	146	1	1
d	Twitter vs Tone	Tone	389	0.93	389	1	**7.86E-76**
	Phee vs Tone	Tone	328	0.87	324	1	**3.57E-45**
	Juvenile vs Twitter	Juvenile	40	0.63	33	1	**9.10E-02**
	Phee vs Tone	Phee	336	0.9	335	1	**2.74E-53**
	Twitter vs Tone	Twitter	386	0.91	386	1	**1.87E-66**
	Juvenile vs Twitter	Twitter	33	0.55	33	1	1
	Noise vs Twitter	Twitter	129	0.87	129	1	**4.89E-17**
	Noise vs Twitter	Noise	130	0.9	129	1	**1.71E-21**
i	Twitter vs Tone	Tone	129	0.79	128	1	**1.56E-10**
	Phee vs Tone*	Tone	9	0.89	9	1	**7.81E-01**
	Juvenile vs Twitter*	Juvenile	1	1	1	1	1
	Phee vs Tone*	Phee	10	0.7	10	1	1
	Twitter vs Tone	Twitter	120	0.8	119	1	**1.71E-10**
	Juvenile vs Twitter*	Twitter	4	0.5	4	1	1
	Noise vs Twitter	Twitter	142	0.76	141	1	**3.37E-09**
	Noise vs Twitter	Noise	144	0.88	143	1	**2.50E-21**
k	Twitter vs Tone	Tone	590	0.79	588	1	**7.60E-46**
	Phee vs Tone	Tone	597	0.88	596	1	**5.35E-87**
	Juvenile vs Twitter	Juvenile	751	0.71	749	1	**2.69E-29**
	Phee vs Tone	Phee	632	0.82	629	1	**4.03E-63**
	Twitter vs Tone	Twitter	601	0.79	598	1	**5.14E-49**
	Juvenile vs Twitter	Twitter	787	0.7	785	1	**1.25E-28**
	Noise vs Twitter	Twitter	541	0.64	532	1	**7.54E-11**
	Noise vs Twitter	Noise	624	0.47	615	1	1

Table 2.3, Summary statistics for the Generalization to novel stimuli across animals and the four conditions (Figure 2.4). Columns under " Binomial test on performance (Figure 2.4C)" report information regarding the statistical deviations of performance (across stimuli and task type) from a theoretically expected distribution of observations (one-sided), with p-values adjusted with a post-hoc Bonferroni correction for multiple comparisons. Significant values are indicated in bold font.

Psychoacoustic assessment of stimulus thresholds

Last, we addressed whether the MXBI can be employed for psychoacoustics. We chose to investigate hearing thresholds in a vocalization-detection task and towards this goal trained 3 animals (animal a, b, and d). In this experiment animals that already knew the association between the acoustic and corresponding visual stimuli (see above: section audio-visual association), were now trained to associate the absence of the vocalization with the visual stimulus for the sTr (Figure 2.5). The method of constant stimuli was employed by randomly selecting the sound level from a set of values between 0 and 80dB SPL. The animals were required to report the presence or absence of the vocalization by touching the marmoset face (visual stimulus coupled with the voc) or the triangles (visual stimulus coupled with silence), respectively. Note that due to nature of the task, reward to the animals for stimuli in the range between 15- and 45-dB SPL was provided regardless of the animal's choice. This was instrumental to prevent frustration and thus disengagement from the task when the acoustic stimulus was presented at amplitudes presumably close to the animal's hearing thresholds. In contrast, reward was dependent on the animals' choice for stimuli above 60 dB SPL and at 0 dB SPL. The aim of this reward scheme, illustrated in Figure 2.5A, was to encourage the animals into using the triangles and the marmoset face as yes/no options for the presence/absence of the acoustic stimulation. After two to three sessions with only high amplitude stimuli (70 dB SPL) to stabilize the animals' discrimination performance at 75% or above, test sessions commenced (3 for animal d and 4 sessions for animals a and b - Figure 2.5B). The estimated hearing threshold for the vocalization stimulus (mean 37.3 dB SPL; 36 for animal a, 49 for animal b, 27 for animal d) was below the background noise of the facility of 60 dB SPL (Figure 2.5C).


a)
dB SPL :
0
15
30
45
60
70
80
Reward scheme:
triangles
regardless of choice
depending on choice
b)
Animal a
Animal b
Animal d
N = 739
N = 726
N = 853
Proportion correct
1.0
0.8
0.6
0.4
0.2
0.0
36
49
27
Trials
174 ± 6
54 ± 4
175 ±10
50 ± 8
198 ± 9
64 ± 3
0 15 30 45 60 75 90
Amplitude (dB SPL)
c)
Amplitude [Arb. units]
2000
-2000
Frequency [kHz]
40
10
10 30 50
Time [sec]

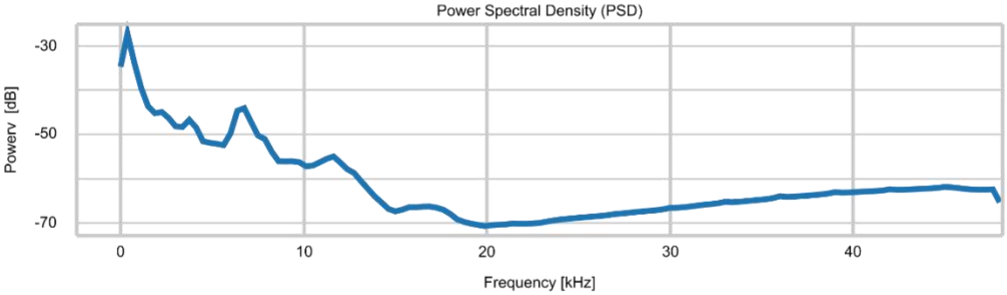


Figure 2.5. Pyschophysical assessment of hearing thresholds for animals a, b, and d after training on the audio-visual association experiment. a) *Schematic representation of the sound levels used and the reward scheme associated with each level. Upon presentation of the vocalization stimulus: at 0 dB SPL, only the selection of the triangles was rewarded. For intensities 15, 30, and 45 dB SPL the animal was rewarded regardless of choice. For intensities at and above 60 dB SPL the reward was delivered depending on the animal's choice.* ***b)*** *Psychometric estimation of hearing thresholds (black vertical lines) based on the proportion of times the animals selected the correct response across the intensities used. 95% confidence intervals (CI) are indicated with black horizontal lines (Animal a: threshold 36 dB SPL; CI between 29 and 43; b: 49 dB SPL, CI between 36 and 58; d: 27 dB SPL, CI between 27 and 40).* ***c)*** *waveforms and spectrograms of 3 one-minute-long snippets recorded inside an MXBI while another MXBI in the same colony room was used to gather data for the audio-visual association experiment. Bottom panel shows the power spectral density of a 5 h long recording with dominant peaks at 6, 12, and 18 kHz caused by vocalizations.*

2.3 Discussion

In this study we report results from four sequential experiments conducted with a stand-alone, touchscreen-based system – termed MXBI - tailored to perform training as well as psychophysical testing of common marmosets in auditory tasks. Animals involved in this experiment operated the device with a consistent level of engagement and for a prolonged time, directly in their own housing environment, without dietary restriction or social separation. All animals navigated an automated, unsupervised training procedure with ease and at their own pace, going from naïve to experienced in a visually-guided discrimination task. In a following audio-visual association experiment, nine out of eleven animals further acquired proficiency in an acoustically-guided 2AC or 3AC discrimination task. Animals also quickly learned to flexibly discriminate three novel sounds they had never encountered before in a Generalization Experiment. Finally, we assessed the hearing thresholds of 3 animals with a spectro-temporally complex sound under potentially distracting auditory conditions. Our results indicate that: 1) marmoset monkeys consistently engage in various psychoacoustic experiments; 2) while performing enough trials and at high performance to allow psychometric evaluations; 3) in a self-paced manner; 4) without the need of dietary restriction or separation from their peers; and 5) with high degree of training flexibility.

Home-cage training of naïve animals

For our experiments we designed a cage-based device and employed an unsupervised algorithm to gradually and autonomously make naïve marmosets accustomed to a 2 or 3 alternative-choice task and a simple detection task in the auditory modality. Each of the 14 animals who participated and successfully completed the first experiment learned 1) to seek and consume reward delivered from the mouthpiece; 2) to operate a touchscreen proficiently; 3) to respond with appropriate timing to abstract sensory stimulation; 4) to understand the concept of a trial structure 5) to tolerate frustration when failing a trial; 6) and ultimately to continuously devise, update, and deploy problem-solving strategies. For practical, experimental, as well as ethical reasons, we aimed at developing an experimental protocol to train many of these aspects directly in the animals' own housing environment, at the animals' own pace (Andrews and Rosenblum, 1994; Evans et al., 2008; Gazes et al., 2013; Crofts et al., 2016; Nakamura et al., 2018; Walker et al., 2020), and without dietary restrictions. Most of these aspects were instructed by a computerized training strategy in which the difficulty was automatically adjusted according to the trial-to-trial performance of the individual animal. The Automated Unsupervised Training (AUT) consisted of a pre-programmed series of steps in which several elements of the task were slowly introduced or adjusted, from trial to trial. The aim of this strategy is to keep animals at a comfortable level of performance to presumably limit frustration, while making the task gradually more difficult and thus making the animals more and more proficient (Berger et al., 2018). Additionally, such subtle, gradual, and constant change in the challenge offered to the animals has been suggested to prevent loss of interest (Bennett et al., 2016; Clark, 2017; Murphy et al., 2003; Tarou and Bashaw, 2007). We indeed observed a long-term rate of engagement, across several hundred sessions across all animals, that would suggest an interest in the experimental sessions that could not be attributed solely to novelty (Murphy et al., 2003). Additionally, animals were always kept together with their cage mate in their home-enclosure and were fed normal colony diet, prior to, after or even during the sessions. Fluid was also available ad libitum. Such generalized and continued interest towards the MXBI , free of any additional coercion, was presumably the result of the combination of a highly preferred primary reinforcer (liquid arabic gum or marshmallow solution), a cognitive, sensory, and interactively rich environment (Bennett et al., 2016; Calapai et al., 2017b; Clarke et al., 2007), and the dynamical adjustments in task level (Tarou and Bashaw, 2007; Berger et al., 2018). Moreover, we did not observe any behavioral alteration that would suggest excessive attachment to our system at the level of the single individual or cage-pair of animals. Rather, 50% of the trials occurred within the first half of the session, in line with a recent report of a steady rate of interactions in voluntary training of motor tasks throughout the waking hours (Walker et al., 2020).

Finally, because we instructed tasks that are typical in cognitive neuroscience and animal cognition (namely a 2 or 3 alternative-choice and a detection task), we believe that similar results would be achieved in training as well as testing other sensory or cognitive domains.

Training flexibility of marmosets

With the exception of two animals who were assigned to a different project and could not be trained further, all animals were successfully trained and tested in audio-visual association experiments reported here. It is important to note that while 2 animals – a and b – readily transferred the knowledge acquired in the visually-guided discrimination (Automated Unsupervised Training) to quickly learn the acoustically-guided discrimination (audio-visual association), the remaining 7 animals required a substantial amount of trials to reach the same level of proficiency. Three animals out the remaining 7 were also tested in Generalization experiment and rapidly generalized the acquired discrimination to novel acoustic stimuli at a comparable rate to animals a and b. Therefore, while the initial transition from the visual to the acoustic domain occurred at variable speed, all tested animals showed a comparable level of flexibility in generalizing to novel stimuli. Finally, all 3 animals tested in the psychoacoustic assessment, quickly learned to reinterpret the discrimination as a detection task as soon as the reward scheme was adjusted accordingly. This allowed for a systematic psychoacoustic assessment of the sound intensity required to detect a vocalization under conditions with background noise.

Together, our results suggest a high degree of training flexibility of common marmosets in general and the auditory modality in particular. Specifically, marmosets can: 1) transfer acquired rules from the visual to the acoustic domain; 2) rapidly learn to discriminate novel acoustic stimuli and 3) flexibly reinterpret a discrimination task as a detection task.

Cognitive hearing in marmosets

The success of the acoustic experiments presented in this study could partly be due to intrinsic properties of the stimuli employed based on the naturalistic connotation in both the visual and the acoustic domain of the juvenile vocalization and juvenile marmoset face association. This ‘natural association’ might then also support the association of the respective other stimuli. Our failed attempts, detailed in the supplementary material, indeed demonstrate the difficulty in having marmosets associate stimuli across the auditory and visual modality. The guiding strategy was that additional properties of the stimuli should match across modalities to support crossmodal association and considered successful concepts from training of rodents and ferrets (Hirokawa et al., 2011; Keating et al., 2013). For example, we presented auditory and visual stimuli together with a reward, or a timeout screen, in a temporally overlapping fashion which leads to strong associations of stimulus components in rodents. Also, the sound was presented from the speaker on which side the correct visual response indicator was located. This has been shown to be a strong cue for ferrets to guide choice towards the respective sound direction. In stark contrast, none of these approaches were successful in marmosets.

Results from the generalization experiment indicate that animals could quickly and flexibly learn to discriminate novel auditory stimuli. However, when two different types of vocalizations were contrasted, only one animal out of 4 performed above chance. Taken together these results indicate that 1) vocalizations might carry a distinctive meaning to the animals that can be exploited to train common marmosets on various psychoacoustic tasks; and 2) the use of a combination of naturalistic and artificial sounds is more likely to instruct marmosets in performing psychoacoustic tasks above chance level.

Psychoacoustic assessment of marmosets in the home enclosure

Performing auditory psychophysics directly in the animals’ colony poses an acoustically challenging environment due to the uncontrolled background noise. The sound pressure needed in order to detect a vocalization of a juvenile marmoset in a cage-based setting - 37.3 dB SPL - was below the sound level of the facility’s background noise - ~60 dB SPL. This might be explained by the adaptation of the auditory system to background sounds which has been documented along the auditory pathway (Dean et al., 2005; Herrmann et al., 2020; Lohse et al., 2020; Rabinowitz et al., 2011; Wen et al., 2009) and has been suggested to optimize perception to the environment (Herrmann et al., 2020; Lohse et al., 2020). Additionally, the juvenile vocalization might have been less affected by background noise (mostly driven by ventilation and marmoset vocalizations) as it minimally overlaps the sound spectrum typically encountered in our colony of adult animals. Nonetheless, our data shows that NHP's psychoacoustic training and assessment is feasible within the animals' home enclosure similar to chair based psychophysics (Osmanski and Wang, 2011). While measurements of hearing thresholds in more classical controlled settings are essential to understand auditory processing and sensitivity, the investigation of audition in more naturalistic environments could provide a closer estimate of real-world hearing capabilities. This might be particularly relevant for auditory processes and mechanisms that involve higher-level, top-down, cortical influences (Fishman et al., 2017; Knyazeva et al., 2018; Schneider et al., 2018) and thus more susceptible to the influence of environmental contextual factors. Environmental sounds produced by conspecifics, for example, could affect how task-relevant sounds are encoded, processed, and interpreted by marmosets that heavily rely on acoustic communication to cooperate, live together, and survive (Morrill et al., 2013).

Towards a high-throughput pipeline for auditory neuroscience

The development of transgenic primate models – and especially marmoset models – for various human diseases (Harris, 2021; Okano and Kishi, 2018; Sasaki et al., 2009; Tomioka et al., 2017) will require phenotyping a large number of animals similar to mouse phenotyping pipelines (Aoki et al., 2017; Brown and Moore, 2012; Poddar et al., 2013). Consequently, cognitive training and testing paradigms, designed around the marmoset model, need to be developed, tested, and implemented (Bennett et al., 2018; Tarou and Bashaw, 2007). Furthermore, in order to allow high-throughput training and testing of common marmosets directly in their own housing environment, our device was designed and built with a series of hardware and software features in mind. First, the use of an inexpensive single board computer as central control unit of the whole device allows for straightforward scaling to more devices and simple adaptation to new experimental requirements. To the best of our knowledge, besides the MXBI introduced here, a fully wireless cage-based system tailored towards visuo-acoustic stimulation and training, capable of ID tagging and set up to be server/client ready has not been presented yet. The wireless connectivity of the MXBI, allowed us to build a network of devices that autonomously interact with a single server node. Upon booting of an MXBI a series of scripts ensures that each device is connected to the central hub where 1) information about animals' ID are stored (used for matching ID codes coming from the implanted chips), 2) data are routinely backed up from the device, and 3) the videos of the sessions are stored. Besides having a unique network address, all devices are essentially identical and can therefore be used on any suitable home cage in our colony. Upon crossing the RFID coil, information coming from the implanted chip will be matched with the database on the server and the local device will load the desired task and AUT step for the given animal. Furthermore, employing a battery-based power solution for the MXBI made the device safer for the animals, due to the exclusively low voltage provided, and easy to handle. While in our case this feature was mostly an add-on, in outdoor cages or on field research sites without direct access to power outlets, this could be a necessary requirement. Combined with image based animal identification (Butler and Kennerley, 2018; Witham, 2018), this would allow for comparative testing of captive and natural populations (Tomasello and Call, 2011). Finally, several structural elements of the MXBI were designed for manufacturability and commissioned to local workshops or locally 3D printed. The combination of structural and electronic hardware elements is particularly well suited, in our opinion, to replicate our device on a large scale. As a result of these built-in features, in the animal's facility of our institute, 6 devices are simultaneously active, training 12 animals in parallel over the course of several hours, and generating on average 1500 trials a day requiring only approximately 35 minutes of human labor.

In conclusion, all of these aspects are to be considered when establishing a successful high-throughput pipelines (across various fields of cognitive neuroscience) because together they ultimately add up to create automated high-throughput protocols for integrating advanced cognitive and behavioral assessments with physiological data recordings (Aoki et al., 2017).

Autonomous devices as cognitive enrichment

Throughout our experiments we found that animals consistently interacted with the device regardless of their performance. In certain occasions animals performed thousands of trials at chance level, across several weeks, despite no social or fluid restriction were applied. While this might seem counterintuitive, we argue that from the animals' perspective our approach, coupled with the appeal of the liquid arabic gum that the device delivered, represents a form of enrichment (Clark, 2017; Tarou and Bashaw, 2007; Murphy et al., 2003). From a psychological standpoint, cognitive enrichment strategies exercise what is known as competence, namely the range of species-specific skills animals employ when faced with various challenges. This, in turn, promotes the sense of agency, described as the capacity of an individual to autonomously and freely act in its environment (Spinka et al., 2001). Promoting both competence and agency has been proposed to be crucial for the psychological wellbeing of captive animals because: 1) animals can better cope and thus better tolerate captivity; and 2) animals can exercise species-specific cognitive abilities that have little opportunity to be expressed in captivity (Brydges and Braithwaite, 2008; Clark, 2017).

Study limitations and caveats

Several animals in the audio-visual association tasks performed at chance level for several thousands of trials. Receiving a reward in half of the trials might be a successful strategy for animals that are not constrained, isolated, or fluid/food restricted. Under these conditions it is unclear whether animals will attempt to maximize their reward - as has been reported in studies where food or fluid regimes are manipulated (Grabenhorst et al., 2019; Wittmann et al., 2020) but see (Jensen et al., 2019) - or are satisfied with chance performance. An animal that is satisfied performing at chance for a certain task will naturally not 'learn' even though it might cognitively be able to. In line with this interpretation, animals that performed at chance level in a 2AC version of an auditory discrimination task, successfully performed the auditory discrimination when the overall chance level was reduced from 50 to 33 % by employing a 3AC version.

Our data demonstrate flexibility of auditory training using natural stimuli and lay the groundwork for further investigations e.g. testing categorical perception of vocalizations by modulating the spectral content of the stimuli used. However, a caveat of our work is that our approaches were not successful in training marmosets on discriminating artificial sounds consistently (see supplementary materials). Among other potential explanations, we attribute this difficulty due to the introduction of auditory cues relatively late in training. This might have biased animals to focus on the visual domain - which is considered the dominant sense in primates (Hirst et al., 2018; Van Essen et al., 2019)- while ignoring other cues. Future studies should therefore explore alternative approaches to train arbitrary acoustic discriminations potentially by introducing reliable auditory cues very early in training.

2.4 Methods

Animal welfare statement

All animal procedures of this study were approved by the responsible regional government office [Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES), Permit No. 18/2976], as well as an ethics committee of the German Primate Center (Permit No. E1-20_4_18) and were in accordance with all applicable German and European regulations on husbandry procedures and conditions. It has to be noted however, that - according to European regulations and implemented in German animal protection law - the procedures described in this study can be considered to be environmental enrichment.

Animals

A total of 14 adult common marmosets of either sex (see Table 2.1) were involved in the experiments carried out in the animal facility of the German Primate Center in Göttingen, Germany. Some of the animals were prepared for neurophysiological and cochlear implant experiments. Animals were pair housed in wire mesh cages of sizes 160cm (H) x 65 cm (W) x 80cm (D) under a light-dark cycle of 12 h (06:00 to 18:00). Neighboring pairs were visually separated by opaque plastic dividers while cloths hung from the ceiling prevented visual contact across the room. Experimental sessions occurred mostly in the afternoon and without controlled food/fluid regimes or social separation from the assigned partner. Liquid arabic gum (Gummi Arabic Powder E414,1:5 dissolved in water Willy Becker GmbH) or dissolved marshmallows (marshmallow juice, 1:4 water dilution) was provided as a reward by the touchscreen device for every correct response in the various experiments. Marshmallow or arabic gum pieces, stuck to the touchscreen, were used during the *initial training* phase.

Apparatus

The marmoset experimental behavioral instrument (MXBI) is directly attached onto the animals' cage and measures 44 cm (H) x 26 cm (W) x 28 cm (D). The device is internally divided into three sections (Figure S4A). The electronics compartment on top contains: a Raspberry Pi 3B+ (raspberrypi.org); a RFID module with a serial interface (Euro I.D. LID 665 Board); two peristaltic pumps (Verderflex M025 OEM Pump), one on each side; a camera module (Raspberry Pi wide angle camera module RB-Camera-WW Joy-IT); and a power bank (Powerbank XT-20000QC3) through which 5 and 12 V (max 2.1A) was provided to the whole system. In our setup and with our tests, the power banks last up to 8 hours before the battery is depleted allowing for continuous training or testing during most of the waking hours of the colony. We chose the Raspberry Pi single board computer instead of more commonly used tablet PCs (Butler and Kennerley, 2018; Issa et al., 2016) for ease of interfacing various external devices. Towards this requirement the Raspberry Pi has various general-purpose input output capabilities allowing to integrate a wide variety of external hardware components such as microcontrollers, touchscreens, etc. with standard communication interfaces (SPI, I2C, I2S). Additionally, new MXBIs can simply be set up by copying the content of the SD card of an existing device into the SD card of the new device. The behavioral chamber in the middle (internal dimensions: 30 cm (H) x 22 cm (W) x 24 cm (D)) hosts: a 10 inch touchscreen (Waveshare 10.1"HDMI LCD [H], later sessions contained a 10" infrared touchscreen attached to the LCD screen, ObeyTec); a set of two speakers (Visaton FR58, 8 ohm, 120–20000 Hz) for binaural acoustic stimulation; a horizontal reward tube with custom-made mouthpiece (placed at 3 cm from the screen but variable between 2 cm and 5 cm); the coil (or antenna) of the RFID and a cylindrical mesh to prevent more than one animal to be inside the device at the same time (Figure 2.1A). Finally, at the bottom of the device, space is left to accommodate a removable tray to collect and clean waste. Hinges on one side allow the device to be opened from the back if cleaning or troubleshooting is needed (Figure 2.1A Left). The MXBI can be anchored to the front panel of the animal's cage via custom designed rails welded to the cage. A removable sliding door at the front panel allows animals to get access to the MXBI when attached. A Python3 based software (Python 3.5.3 with the following modules: tkinter 8.6, numpy 1.12.1, RPi.GPIO 0.6.5, pyaudio 0.2.11) running on the Raspberry Pi records all interaction events (screen touches, RFID tag readings and video recording), manages stimulus presentation (acoustic and visual), controls the reward system and finally backs up the data automatically to a server via wireless local network connection (Figure S4B).

Procedure

Behavioral training and testing sessions were started by connecting the Raspberry Pi and LCD display to power which initiates booting. After booting, a custom script with a series of preconfigured commands was automatically initiated to: 1) connect the device to a central server for automatic, recursive, data logging, as well as main database access; 2) start the local camera server for remote monitoring and video recordings (Figure 2.2B); 3) automatically launch the experimental task when needed. The fluid reward was manually loaded in each device and the pump was primed. The device was then attached to the cage and the sliding door in the front panel removed for the duration of the session. At the end of the session, the sliding door was placed back between the device and the cage so that the device could be detached, cleaned, and stored. The touchscreen surface and the behavioral compartment were thoroughly cleaned to remove odors and other traces. Hot water was used daily to clean the reward system to prevent dried reward from clogging the silicon tubes and mouthpiece. The entire process requires a single person around 35 minutes (15 for setting up and 20 for taking down) with 6 devices.

Sessions

In order to operate the touchscreen at the opposite end from the MXBI's entrance, the animals are required to go through the opening on the front panel and the mesh cylinder (Figure 2.1A). Crossing the antenna inside the mesh cylinder identifies animals via their RFID transponder (Trovan ID-100A) implanted between the animal's shoulders for husbandry and identification reasons. Standing up inside the mesh places the animals' head 3 cm above of the mouthpiece and 4 to 5 cm away from the screen, directly in front of a cut out in the mesh of 3.5 x 8.5 cm (HxW) through which the touchscreen can be operated (Figure 2.1A). Throughout each session, animals were regularly monitored by the experimenter from a remote location (approximately every 15 minutes). Additionally, videos from most sessions were recorded and stored. Fluid (either water or tea) was available *ad libitum* to the animals within their home cage but outside the MXBI. Solid food was provided to the majority of the animals before, after, and during the session, depending on husbandry and / or veterinary requirements.

Experimental paradigm

Throughout the experiments, animals never left their home cage. With the exception of animals a and b, that where pilot subjects and underwent a different initial procedure, all animals were first trained manually to operate the device at a basic level by means of positive reinforcement training and shaping techniques (see methods section: *initial training*). Afterwards, all animals where guided by an unsupervised algorithm through a series of preconfigured training steps (see section Automated unsupervised training (AUT)) to acquire basic proficiency in a standard 2AC discrimination task. The animals' discrimination proficiency was then tested and refined in a next experiment in an acoustically-guided discrimination task (see section audio-visual association). In a third experiment, the acoustic stimuli were replaced with novel stimuli and the animal's ability to generalize was assessed (see section Generalization to novel stimuli). Last, we developed a psychoacoustic detection task to quantify the animal's hearing thresholds (see section Psychoacoustic assessment). It is important to note that not all animals took part in all experiments either because some animals were assigned to different projects or were not always available due to requirements of different experiments.

Initial training

The goal of the *initial training* procedure was to instruct naïve animals to use the touchscreen. To this end, this training was divided into three sequential steps: first, habituation to the device; second, forming a mouthpiece-reward association and finally, a touch-to-drink phase. During the first two steps no wire mesh cylinder was placed inside the MXBI. Unlike the remainder of the training, all *initial training* required the constant surveillance of the experimenter, to remotely access and control the screen of the device from another computer to shape the animal's behavior while monitoring the video feed. The measured round-trip delay between observing the behavior and effectively delivering the reward was approximately 400 ms plus an additional response latency of the observer. Together, we believe that this delay should be sufficient for stimulus – response integration and association (Kobayashi and Schultz, 2008). The *initial training* lasted on average 6 (±2) sessions and was routinely completed within 2 weeks. With the exception of animals a and b, all animals underwent the *initial training*.

Device Habituation: During this first step the device was attached to the cage without the mesh cylinder, to allow the animals to freely explore the behavioral chamber (see supplementary video 1) in sessions lasting on average 40(±20) minutes. Before switching to the next step, the experimenter ensured that both animals would show interest and no aversion (e.g. walking towards and not away from) the device. The number of sessions needed to observe this behavior varied between 1 and 2.

Mouthpiece-reward association*:* Following the habituation, drops of reward of variable magnitude (between 0.3 and 0.5 ml) were remotely triggered by an experimenter in order to direct the interest of the animals towards the mouthpiece (see supplementary video 2). Presumably due to the sudden occurrence of the pump sound while rewarding, the interest towards the MXBI for some animals slightly decreased. To overcome this issue and to increase the likelihood of animals interacting with the device a number of small marshmallow pieces were placed randomly over the mouthpiece. After all pieces were consumed and the animals left the MXBI the experimenter closed the sliding door to place new pieces. Once the animals showed interest in the mouthpiece in the absence of the reward, the association was considered established and the next phase started. This step required between 1 to 5 sessions, with each session lasting 30 to 60 minutes.

Touch-to-drink phase: The aim of this step was to teach the animals to actively seek the reward, by triggering the touchscreen. In order to achieve such behavior efficiently and to make sure the animal used the hand and not e.g. their mouth (which was observed in pilot experiments) to touch the screen, a mesh cylinder was placed inside the device. In turn, this restricted access to one animal at a time, and improved the efficiency of the RFID identification. Additionally, small pieces of marshmallows were placed on the screen within the triggering area, to encourage the animals to retrieve the marshmallow pieces and thereby touch the screen. When all pieces were consumed and the animal had left the MXBI the experimenter closed the sliding door to place new pieces on the screen and resumed the session. While the marshmallow pieces where collected, fluid reward was provided, triggered either remotely by the experimenter or by the animals themselves touching the stimulus on the screen. This procedure successfully allowed all animals to switch from reaching to retrieve the marshmallows to simply touching the screen to trigger fluid reward (see supplementary video 3.1 and 3.2). After 5 to 10 consecutive reaching movements towards the screen in the absence of marshmallows, the behavior was considered acquired and the *initial training* concluded. Between 1 and 4 sessions (each lasting 60 minutes on average ± 10min) were necessary to finish the touch-to-drink phase.

Automated unsupervised training (AUT)

Upon completion of the *initial training* phase, all animals underwent an automated stepwise protocol designed to gradually bring the animals from a *quasi*-naïve state to proficiency in a 2 alternative-choice (2AC) audio-visual association task. Throughout the protocol the performance of the animal was constantly monitored by an algorithm to adjust the task

difficulty, by changing parameters as well as introducing or removing elements in the training (Figure S2.4C). Animals ascended in steps by performing at least 8 trials out of 10 correctly and descended in steps when less than 3 trials out of 10 were correct (Figure S2.4D). Finally, the progress of each animal was automatically stored and retrieved on each trial, so that the animals could navigate the series of steps and resume from the last step they were in when they left, across breaks and sessions. The automated training protocol (AUT – see (Berger et al., 2018)) was comprised of 48 AUT steps, grouped into four milestones: – decrease of the size of a visual stimulus (trigger) to be touched for reward, change of position of said stimulus, introduction of sound and delayed presentation of a visual target, introduction of a second visual target as a distractor –. Through these steps and milestones, the animals were trained on the basics of how to operate a touchscreen within the context of a standard 2AC visually-guided task. The aim of the AUT was to prepare the animals for an audio-visual association experiment, in which they were required to distinguish between different sounds by selecting a corresponding visual stimulus. During the first 15 steps (size milestone) a white circle embedded in a blue rectangle (trigger) placed on the vertical meridian had to be touched to obtain a reward (0.1 – 0.2 ml). From step 2 to 15 the trigger gradually shrunk in size from 6x6 cm to the final size of 3x3 cm. Touching the screen outside the trigger resulted in a 2.5 (earlier sessions) to 5 seconds (later sessions) long timeout indicated by a grey screen during which no new trial could be initiated and touches were ignored. A touch within the boundaries of the trigger resulted in reward administration (as above), followed by a new trial which could be started after 0.8 to 2.5 seconds. In steps 16 to 30 (position milestone) the trigger's position gradually changed by 5 mm at each step, either to the left or to the right of the original central position, until the edge of the screen was reached. From step 31 onwards (delay milestone) the trigger first appeared at the center and upon touch reappeared at the left or the right edge of the screen and had to be touched again. The reward was delivered if both touches were executed correctly. Only touching outside of the second trigger resulted in a timeout. This was done to ease the transition from one stimulus to two different stimuli presented, which was occurring on all steps starting with step 36. Throughout these steps the second trigger was replaced randomly with one of two visual stimuli (targets): either the picture of an infant marmoset face (3x3 cm), or an abstract geometric pattern (3x3 cm) (Figure 2.3A). Starting from step 36 an acoustic stimulus (either a repeated infant marmoset vocalization (Gultekin and Hage, 2017); or a train of pure tones - sTr - chosen for individual animals from a range between 1.5 and 3.5 kHz) was presented 1 to 1.5 seconds before the visual target, with a gradually increasing sound intensity (in steps of 10 dB) from 32 +/- 2 dB SPL on step 36 to a final loudness of 72 +/- 2 dB SPL on step 40. The vocalization was followed by the marmoset face (for 5 seconds) while the sTr was paired with the geometric pattern (Figure 2.3A). From step 41 to step 45 the parameters were kept the same as step 40, to provide prolonged and unchanged exposure to the visuo-acoustic stimulus. At step 46 (visual 2AC milestone) a visual distractor was displayed together with the target but on the opposite edge of the screen. In case of a ‘vocalization’ trial the visual distractor was the geometric pattern and vice versa. The distractor was gradually increased in size from 0.3x0.3 cm on step 46 to 2.8x2.8 on step 49. Thus, from step 46 to 49, animals could exploit the size difference between the visual target and distractor to respond correctly by choosing the larger visual stimulus. Throughout the protocol, if no response was observed within 7s from stimulus presentation the trial was aborted and the trial outcome was labelled as ‘ignored’. The AUT described here (version #10) is the result of several attempts that are described in the supplementary material (Table S2.2).

Audio-visual association

The audio-visual association experiment starts when an animal reaches step 50 in the Autonomous Training (Figure 2.3). Contrary to the AUT, no visual cue could be used to correctly identify the target of a given trial. Here animals had to solely rely on auditory cues to obtain a reward above chance level. In this experiment no AUT algorithm was employed and therefore the trial structure and sequence remained unchanged throughout. This experiment consisted of a two-alternative choice task (2AC), where only one of the two available options was the correct one and the animal’s ability to distinguish the options was assessed from the animal’s relative frequency of choice. We implemented two variants of this task, a 2AC and a 3AC, plus a control condition (see supplementary material). Both variants employed the same stimuli of the Autonomous Training with added visual distractors in the 3AC variant which had no sound associated and were not presented as target, but always as a distractor. While touching the target of a given trial was rewarded, touching a distractor resulted in 5 seconds (later sessions) long timeout indicated by a grey screen during which no new trial could be initiated and further touches were ignored. On the contrary, after a correct response, a new trial could be started 0.8 to 2.5 seconds after reward delivery. A detailed timeline of an example trial from this task is shown in Figure S2.4E and a video of an animal performing a trial in the 2AC variant is available in the supplementary materials (Video 4). Animals who did not perform above chance on the 2AC variant were assigned to the 3AC variant. The 3AC variant was used to lower the chance of obtaining a reward randomly at any given trial from 50% to 33%. Two animals who performed at chance level in the 2AC were assigned to a different experiment and could not be tested on the 3AC.

Generalization to novel stimuli

To evaluate the flexibility of our protocols and determine whether the animals could generalize the already learned 2AC task using different sounds, we performed 4 different variations of the already described 2AC task. Here we tested a twitter vs a pure tone, a phee vs a pure tone, a twitter vs white noise and an infant vocalization vs a twitter. To avoid a high number of changes within every task switch, once the animals learned the first task variation (twitter vs pure tone) they were always brought back to this task to stabilize their performance before moving to the next task variation. Vocalizations were recorded from a different colony (Agamaite et al., 2015). Representative visual indicators that matched every single acoustic stimulus are shown in Figure 2.4.

Psychoacoustic assessment

In order to assess the animals' hearing thresholds, we devised a simple detection task based on the discrimination task used before. In this task animals were trained to choose the grey triangles (previous visual stimulus of sTr) to report the absence of the vocalization (i.e. silence). Once the behavior was stable (after two sessions) and based on the measured background noise of the facility (60 ± 5 dB SPL, see Figure 2.5C) we set the sound intensities to 0, 15, 30, 45, 60, 70, 80db SPL for the vocalization. Given that some of these intensities were below the background noise of the facility, all trials with intensities between 15 and 45 dB SPL were rewarded regardless of the choice of the animal (Figure 2.5A). Moreover, vocalization trials at 0 dB SPL were rewarded if the triangles were selected (visual stimulus for the silence). This was instrumental to first account for both type of trials (silence and vocalization) presented at 0 dB SPL, and second to effectively establish the task as a detection rather than a discrimination task. Finally, all sessions were performed in the afternoon, from 1pm to 4.30pm, when the colony's background noise was the lowest with feeding and personnel's activity occurring mostly in the morning.

In order to measure the sound pressure level in the MXBI, the two devices used to gather hearing thresholds (1 for animals a & b, 1 for animal d) were calibrated inside an insulated sound proof chamber. An amplifier (Hifiberry amp2) coupled to the Raspberry Pi produced the audio signal, while a measuring amplifier (Bruel And Kjaer Measuring Amplifier Types 2610) and a microphone (Bruel And Kjaer Type 4966 1/2-inch) placed at the marmoset ear level pointing towards one speaker, acquired the sound output. Additionally, an oscilloscope (Rigol DS1000Z), attached to the output lines of the amplifier, measured the voltage. We were able to corroborate the step size (0.5 dB SPL) of the amplifier by sampling 5 different frequencies (0.875kHz, 1.75kHz, 3.5kHz, 7kHz, 14kHz) at 10 different sound pressure levels (100dB, 95dB, 90dB, 85dB, 80dB, 75.5dB, 70dB, 65.5dB, 60dB, 50dB). We found a stable and accurate correspondence between the values provided to the amplifier, the sound pressure levels measured by the measuring amplifier, and the voltage values measured by the oscilloscope.

2.5 Data treatment and Statistics

Data acquisition, processing, analysis, and statistical testing was performed in Python 3.5.3 and 3.9. Statistics and significance tests for Figure 2.1 to 2.4 were calculated via the packages *scipy* (Millman and Aivazis, 2011; Oliphant, 2007) and *numpy* (Oliphant, 2006), co-installed upon installation of the package *seaborn*. An alpha level of less than 0.05 was considered significant. Data formatting and visualization for the same figures as well as for Table 2.1 was achieved with the packages *pandas* (McKinney, 2010) and *seaborn* (seaborn.pydata.org). Hit rate's significant difference from chance (Figure 2.3C) was assessed with a Binomial test, post-hoc adjusted for multiple-comparisons with the Bonferroni correction; while reaction time differences between the two presented auditory stimuli (Figure 2.3D) were tested for significance with a Kruskal-Wallis test by ranks. Both tests were adjusted post-hoc for multiple-comparisons with Bonferroni correction (corrected alpha = 0.0019, from the python module *statsmodel.stats.multitest.multipletests*). In Figure 2.2D and Figure 2.3A, B, the variable “percentage of trials” on the abscissa was used to achieve a shared and standardized axis on which multiple animals could be compared and visualized against each other, irrespective of the total amount of trials each individual performed. The assumption behind this choice was that learning occurs through similar mechanism across individuals, but unfolds through a different amount of trials that depend on each animal's engagement level. The resulting process of standardization attenuated the inter-individual variability between animals for parameters such as steps of the AUT (Figure 2.2C) and Hit Rate in (Figure 2.3A, B and 4B).

Psychometric function estimation was achieved with the python module psignifit (Schütt et al., 2016) set to fit a cumulative normal sigmoid function, with all parameters free and with 95% confidence intervals. The resulting function can be expressed as follows:

$$\psi(x; m, w, \lambda, \gamma) = \gamma + (1 - \lambda - \gamma)S(x; m, w) \qquad (1)$$

where m represents the threshold (level at 0.5), w represents the width (difference between levels 0.5 and 0.95), λ and γ represent the upper and lower asymptote respectively (equation 1 in (Schütt et al., 2016)).

2.6 Data availability

The datasets (Calapai, Antonino et al., 2022) generated during and/or analyzed for the current study are available at GitHub (https://github.com/CHiP-Lab/mXBI) and Zenodo (https://doi.org/10.5281/zenodo.6139297).

2.7 Code availability

The code (Calapai, Antonino et al., 2022) to recreate the data figures are available at a dedicated Github repository (https://github.com/CHiP-Lab/mXBI) and Zenodo (https://doi.org/10.5281/zenodo.6139297).

2.8 Acknowledgments

We would like to thank Dr. Steffen Hage and Dr. Michael S. Osmanski for providing juvenile and adult call exemplars, respectively. Additionally, we would like to thank Karin Tilch and Manfred Eberle of the German Primate Center for the pictures of the device and marmosets, respectively, used in this manuscript. The team around R. Schürkötter of the workshop of the Max-Planck-Institute for Multidisciplinary Sciences manufactured the device cases. The work was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 670759 – advanced grant "OptoHear" to T.M.), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) via the Leibniz Program (MO896/5 to T.M.), and by the Max-Planck-Society (fellowship to T.M.). In addition, work by T.M. was supported by Fondation Pour l'Audition (FPA RD-2020-10).

2.9 Supplementary material and extended information

1. AUT versions performed by each animal
    - Table S2.1
2. Overview of all automated unsupervised training protocols (AUT); TableS2.2
3. Acoustic Artificial Discrimination
    - Figure S2.1
    - Figure S2.2
    - Table S2.3
4. Inter-Trial-Intervals analysis
    - Figure S2.3
    - Table S2.4
5. Additional device, training and trial timing information
    - Figure S2.4
6. Example Videos:

Supplementary Movie 1: Habituation Phase. An animal can be seen exploring the behavioral chamber for the first time, file *Supplementary Movie 1.mp4*

Supplementary Movie 2: Mouthpiece-reward association. An animal can be seen approaching the mouthpiece and consuming the fluid reward. After the habituation phase, an experimenter remotely delivered fluid reward through the mouthpiece, file *Supplementary Movie 2.mp4*

Supplementary Movie 3.1: Touch-to-drink phase. An animal inside the mesh can be seen collecting pieces of marshmallows attached to the screen. Accidental touches of the screen during this process resulted in fluid reward delivered through the mouthpiece, file *Supplementary Movie 3_1.mp4*

Supplementary Movie 3.2: Touch-to-drink phase. The same animal of Video 3.1 touches the screen and consumes fluid reward from the mouthpiece, after all marshmallows have been collected, file *Supplementary Movie 3_2.mp4*

Supplementary Movie 4: An animal solves several trials on the final step of the Natural Discrimination task, file *Supplementary Movie 4.mp4*

1. AUT versions performed by each animal

Animal	AUT versions	Order of tasks
a	1, 2, 4, 5, 6, 8, 9,10	Pilot experiments, Natural Discrimination, Artificial Discrimination
b	1, 2, 4, 5, 6, 8, 9,10	Pilot experiments, Natural Discrimination, Artificial Discrimination
c	8, 9, 10	Natural Discrimination
d	9, 10	Natural Discrimination
e	9, 10	Natural Discrimination, Artificial Discrimination
f	8, 9, 10	Natural Discrimination
g	4, 5, 6, 8, 10	Pilot experiments, Artificial Discrimination, Natural Discrimination
h	4, 5, 6, 8, 10	Pilot experiments, Artificial Discrimination, Natural Discrimination
i	8, 9, 10	Natural Discrimination
j	9,10	Natural Discrimination
k	9, 10	Natural Discrimination
l	4	Pilot experiments
m	9, 10	Natural Discrimination
n	9, 10	Natural Discrimination

Table S2.1 – List of AUT versions performed by each animal. Data with versions 3 and 7 are not included in this manuscript due to technical issues with the RFID and to the nature of the experiment (control experiment testing pure visual assessment), respectively.

2 Overview of the AUT versions

AUT version	Description	Visual stimuli	Acoustic stimuli	Changes	notes
1	Protocol made of 40 steps composed of two main sections. The first focusing on training touch precision with a start-stimulus placed in the center of the screen that decreases in size until it reaches the final size of 3x3 cm. The second requiring a total of two interactions to obtain reward, one towards the start stimulus, placed in the center of the screen, and a second one towards either of the two visual stimuli, placed either left or right of the screen. Throughout this second section the distractor stimulus increases in size until it reaches the same size of the target stimulus.	Start-stimulus: White circle without background. Visual stimuli: two triangles, red and blue at opposite orientations, without background	Constant sine wave vs. no sound, matching the red and blue triangles respectively		
2	Same as version 1	Same as version 1	Same as version 1	A white background to the visual stimuli is added.	
3	Protocol made of 47 steps, composed of two main sections. The first section is the same as version 1. The second section is extended in the number of total steps, for finer distractor size increase.	Same as version 1	Same as version 1	17 new steps are added in the second section.	Data not processable due to tecnhical issues with the RFID
4	Protocol made of 54 steps, composed of three main sections. Section 1 is the same as version 1. The second section focusses on training the animal to reach for the target stimulus at different positions on the screen. From trial to trial the target is shown at variable eccentricities. The eccentricity is increased gradually until the edge of the screen is reached. The third section is the same as section 2 in version 3.	Start-stimulus: same as version 1. Visual stimuli: red cross and blue triangle embedded in a white background.	Simple train tone pulse vs. no sound, matching red cross and blue triangle respectively	The position of the visual stimuli is randomly assigned to right and left of the screen center, on a trial by trial basis. The identity of the visual stimuli has changed.	
5	Protocol made of 44 steps, composed of three main sections. The first section is the same as version 4 but with a reduced number of steps. In the second section, the acoustic stimulus, in each trial, is played from the left or right speaker, coherently with the side of the screen in which the visual stimulus was shown. The third section is the same as section 2 in version 3	Same as version 4	Same as version 4	Tthe visual and the acoustic stimuli are coherent in their source location. Decrease in number of steps in the first section.	
6	Protocols made of 44 steps, composed of three main sections. The first section one is the same as version 5. The second section is the same as version 5, but with the termination of the trial in case of no interactions after 7 seconds from stimuli onset (ignored trials). The visual and acoustic stimuli are presented at the same time.	Start-stimulus: White circle embedded in a blue background. Visual stimuli: same as version 4	Same as version 4	Implementation of ignored trials, with visual and acoustic stimuli disappearing after 7 seconds, and a new trial starts. Overlapping of visual and acoustic stimuli during presentation enhance coherence.	
7	Protocol made of 35 steps, comprised of three main sections. All sections are the same as version 6, but section contains less steps.	Start-stimulus: same as version 6. Visual stimuli: gray cross and gray triangle embedded in a gray background.	No sound	Alternative task for assessing visual discrimination. No sound / visual target association	Data now shown. Control experiment to assess pure visual discrimination

8	Protocol made of 44 steps, composed of three main sections. The first section is the same as version 6. In the second section the trial start button is shown at variable eccentricities. The eccentricity is increased gradually until the edge of the screen is reached. The third section is the same as version 6, but the feedback to the animal is enriched. The compound stimulus (acoustic and visual) is presented during the reward. The distractor and the target are shown in isolation after a wrong and correct response, respectively.	Same as version 6	Same as version 6	Overlapping of visual and acoustic stimuli together with reward delivery when a correct response is registered, the distractor stimulus is removed from screen. When wrong response is registered, the distractor stimulus remains and target stimulus is removed from the screen. In the second section the visual stimuli are replaced by the start stimulus.	
9	Same as version 8	Start-stimulus: same a version 6. Visual stimuli: face of baby marmoset vs. gray scale triangles within a gray square	Baby marmoset vocalization vs. train tone pulse, matching the baby face and the gray triangles composite respectively.	Change of visual and acoustic stimuli.	
10	Protocol made of 50 steps, comprised of three main sections. Same as version 8 but with added steps in the second section. For few animals, as a control condition, a different set of compound stimuli (acoustic and visual) are used – see Artificial Discrimination task in the methods sections. This is the final and most successfully protocol, described in the method section of the original manuscript.	Start-stimulus: same as version 6. Visual stimuli: same as version 9. Visual stimuli for the control condition referred as Artificial Discrimination: RGB geometric figure embedded in a yellow background vs. gray scale triangles within a gray square.	Acoustic stimuli: same as version 9. Acoustic stimuli for the control condition referred as Artificial Discrimination: simple train used as acoustic stimulus in version 9 vs. two tones train pulse (complex train), (matching the RGB geometric figure and the gray triangles composite respectively)	A new set of steps are added in the second section. A control condition is implemented for few animals	

Table S2.2; overview of all versions of the automated unsupervised training protocol (AUT)

3. Artificial Discrimination

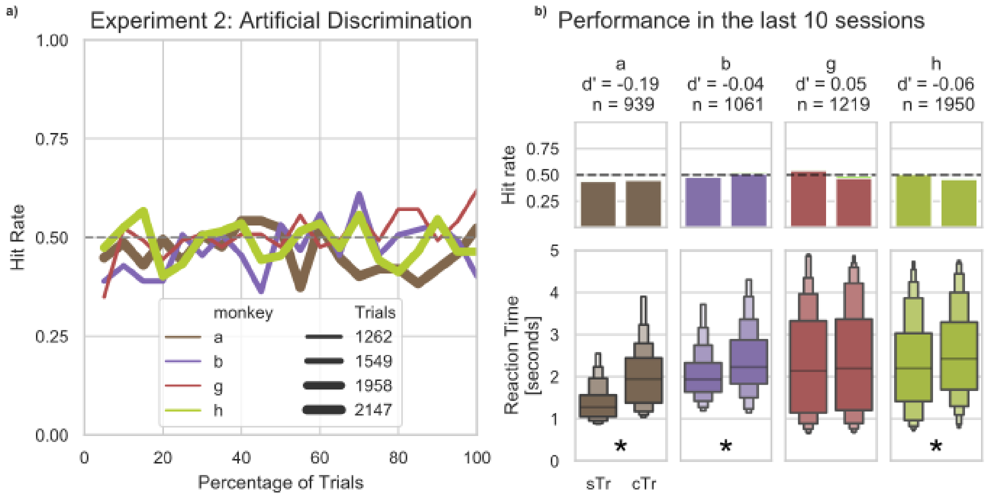


Figure S2.1. Pyschophysical assessment of hearing thresholds for animals a, b, and d after training on the audio-visual association experiment a) *Hit rate as a function of percentage of trials performed in the Artificial Discrimination. Line thickness represents the total number of trials of each animal in this task. Dashed line at 0.5 represents chance level. Hit rate for each animal is represented as a function of the animal's percentage of trials performed and is grouped into bins of 5% of trials.* ***b)*** *hit rate as a function of s timulus type ("sTr" for the train of pure tones, "cTr" for the train of alternating pure tones) in the last 10 sessions, with corresponding number of trials and sensitivity index (d') above each bar plot. Letter-value plots of the reaction times were plott ed for each stimulus type separately. The central box defines the median and 25th up to 75th percentiles. Successively narrower boxes are drawn between the 1/8th and 7/8th, the 1/16th and not significant difference in reaction times between the two stimuli at a Bonferroni post-hoc corrected Kruskal-Wallis Test (one -sided test), for a detail statistics summary see Table S2.3.*

This variant of the audio-visual association experiment (Figure S1) employed two artificially constructed stimuli consisting of trains of Pure Tones: a simple Train (sTr) composed of a repeating pure tone in the range between 1.5 and 3.5 kHz (repetition rate of 3.3 Hz, one frequency chosen per animal); and a complex Train (cTr), consisting of a repeating pattern of 2 pure tones alternating between a fundamental frequency (chosen from 1.5 to 3.5 kHz, one fundamental frequency per animal) and a frequency 42 % higher (which was always larger than the minimum discriminable frequency difference [1]). The sTr was associated with a composite of three grey scaled triangles embedded in a 3x3cm grey square, and the complex train with a colored geometric figure embedded in a 3x3cm yellow square. None of the tested animals performed above chance although differences in reaction times were observed. Note that one marmoset (animal e), that performed on this task was not included into Figure S1 and table S3 due to technical problems during the sessions.

Animals	Stimulus	Hitrate	Trials	d'	Binomial test on performance (Figure S1)			Kruskal-Wallis test on Reaction Times (Figure S1)					
					N (w/o ignored)	Degrees of freedom	Binomial Test (adjusted p-value)	Median	IQR	N	Degrees of freedom	Test Statistics	Kruskal-Wallis (adjusted p-value)
a	sTr	0.45	460	-0.19	456	1	1	1.27	0.52	223	1	89.848	**2.06E-20**
	cTr	0.44	497		489	1	1	1.94	1.06	207	1		
b	sTr	0.51	385	-0.04	380	1	1	1.94	0.69	215	1	21.188	**3.33E-05**
	cTr	0.48	405		404	1	1	2.23	1.04	215	1		
g	sTr	0.47	651	0.05	636	1	1	2.14	2.18	287	1	0.049	**1**
	cTr	0.54	591		585	1	0.15	2.20	2.16	260	1		
h	sTr	0.46	1017	-0.06	1004	1	1	2.20	1.62	457	1	12.141	**0.003**
	cTr	0.50	963		954	1	1	2.42	1.61	419	1		

Table S2.3, Summary statistics for the variant "acoustic discrimination" of the audio-visual association experiment across animal and stimuli (Figure S1). Significant values are indicated in bold font. D-prime value is provided as indication of the sensitivity of each animal on given task. Columns under "Kruskal-Wallis test on Reaction Times (Figure S1)" report information regarding the statistical difference of the reaction time to the sTr and the cTr stimuli, with p-values adjusted with a post-hoc Bonferroni correction for multiple comparisons. Columns “Binomial test on performance (Figure S1)” report information regarding the statistical deviations of performance (across stimuli and task type) from a theoretically expected distribution of observations (one-sided), with p-values adjusted with a post-hoc Bonferroni correction for multiple comparisons.

With the aim of testing further attempts on how to train animals to perform artificial discrimination, we developed an alternative approach (AD_2, Fig. S2) which was designed as a continuation of the acoustic discrimination AUT described in the main text. The idea behind AD_2 was to introduce a previously unknown discrimination by contrasting a new stimulus with stimuli for which a stimulus-response association already exists and then successively reducing the percentage of trials with known contrast while increasing the percentage of trials for the unknown stimulus contrast.

In other words: initially animals know to discriminate the simple train (sTr) from the vocalization (voc), by touching a geometric figure or a marmoset face, respectively. The final goal is to discriminate a simple train from a complex train (cTr) by choosing appropriate geometric patterns (triangles vs. keyhole; for a stimulus description and correct visual response see Fig. S2A). At the beginning of the AD_2, the 2 alternative stimuli available at every trial are the already acquired ones, namely the sTr or the voc. Throughout the steps of the procedure, the voc is replaced, in increasing proportion of trials, by the cTr. Therefore, while the sTr always had 50% chance of being a target, the voc probability decreased throughout the procedure, in favour of cTr, the probability of which increased in steps of 4 % per level. The resulting 12 possible trial types can be seen in Fig. S2A. Moreover, to move between steps of the AD_2 we modified the performance evaluation algorithm such that increases in step occurred after 80 % of trials or more were correct within a window of 24 trials and step-downs already occurred if 45 % of trials or less were correct within a 24-trial window. The AD_2 starts from step 50 (final step of the acoustic discrimination AUT) and gradually increases the percentage of trial types 9, 10, 11 and 12 while decreasing the percentage of trial types with an already trained stimulus-response association (1, 2, 3 and 4). In all steps the animal could correctly perform each trial by selecting a known stimulus-response (trial types 5, 6) or excluding a known response (trial types 7, 8). On each of the 12 steps of the AD_2 the proportion of new trial types (9, 10, 11, 12) increased by 2.1 % per step while trial types with known stimulus-response association (1, 2, 3, 4) decreased over 6 steps by 4.2 % per step. To assist in learning the new stimulus-response association between the cTr and the keyhole pattern, trial types with a vocalization (voc; trial types 5, 6) or the cTr paired with a marmoset's face as distractor (7, 8) were first introduced and increased in likelihood along the stair case until step 56 (2.1 % per trial type and step) after which they were successively eliminated until step 62 (2.1 % per trial type and step). Animal a quickly progressed through the AD_2 reaching the final step (62) for the first time in session 3 (after 1940 trials from the start of the AD_2 procedure, Fig. S2B, C) and stabilized on step 62 from session 9 (after 4222 trials) when the animal quickly recovered from previous step-downs in sessions 5 to 8 (Fig. S2B, C). To assess whether animal a had acquired the final discrimination after stabilizing, we compared only trials where the sTr had to be discriminated from the cTr with their respective visual targets (trial types 9, 10, 11 and 12). Throughout 3552 total trials animal a chose the keyhole visual target after cTr significantly more often than after sTr presentation (Fisher's Exact test, $p = 1.6*10^{-71}$; cTr hit rate = 62 %, sTr hit rate = 68 %, see Fig. S2D).

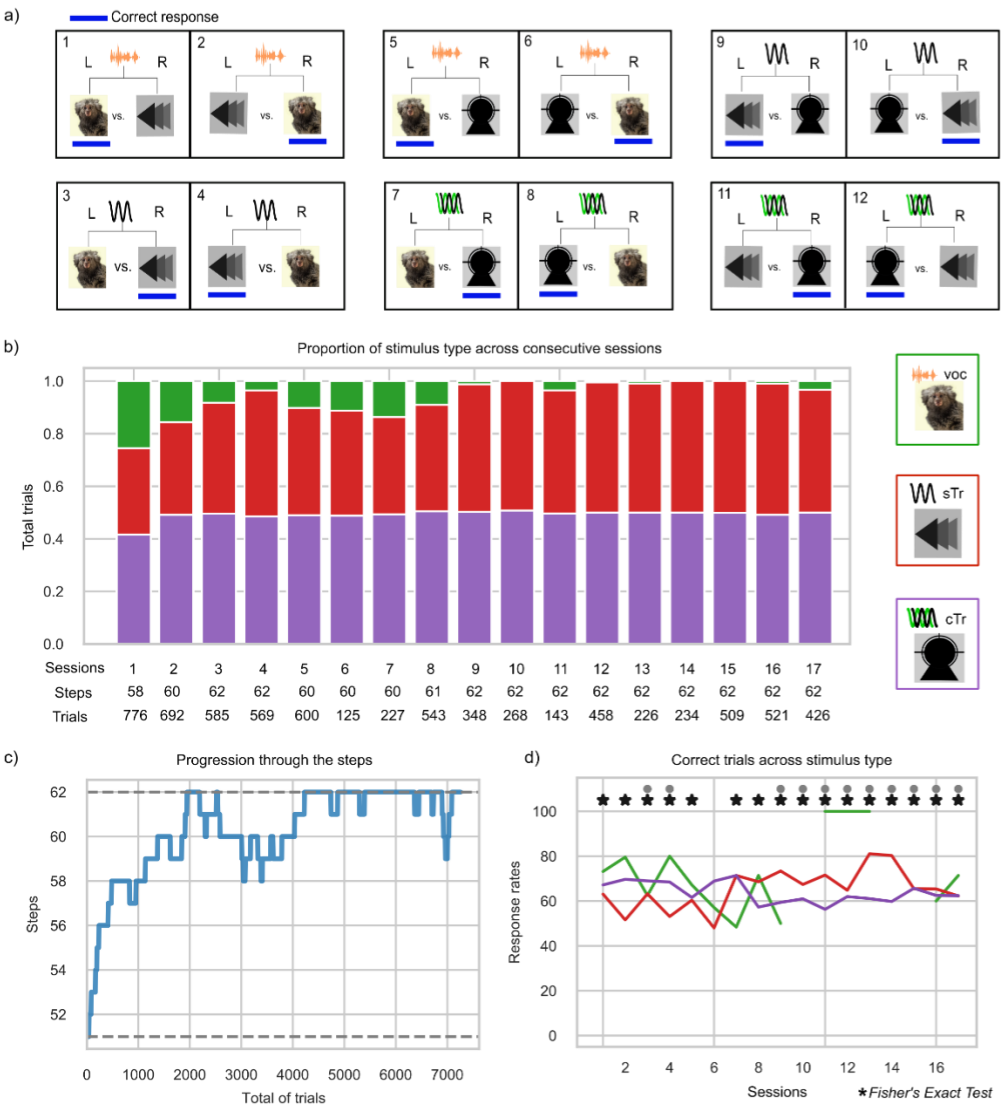


Figure S2.2. Performance of animal a on version 2 of the Artificial Discrimination task. ***a****) Graphical representation of the 12 trial types employed within the AUT, Blue bar under visual indicator shows correct response.* ***b)*** *Shows the percentage of every auditory target across consecutive sessions. Numbers at the bottom indicate the number, highest step reached in that session and the total number of trials for every session. Colors of the panels on the right corresponds to the three targets.* ***c)*** *Learning curve of animal a along the different steps of the task.* ***d)*** *Depicts the hit rates for the three individual targets (sTr, Voc and cTr) across consecutive sessions. The grey dot on top marks, for every session, whether the highest step (step 62) in the AUT was reached. The star represents a significant difference between hit-rate and error-rate between trial types 9, 10 vs 11, 12 or sTr and cTr (Fisher's Exact Test, two sided), p-values listed by session order (3.11e-13, 2.14e-07, 4.24e-14, 1.92e-07, 3.08e-07, 8.05e-02, 2.31e-09, 8.97e-09, 7.26e-10, 3.22e-06, 1.08e03, 5.41e-09, 8.60e-11, 4.16e-10, 1.74e-12, 3.09e-10, 6.90e-08).*

4. Inter-Trial-Intervals analysis

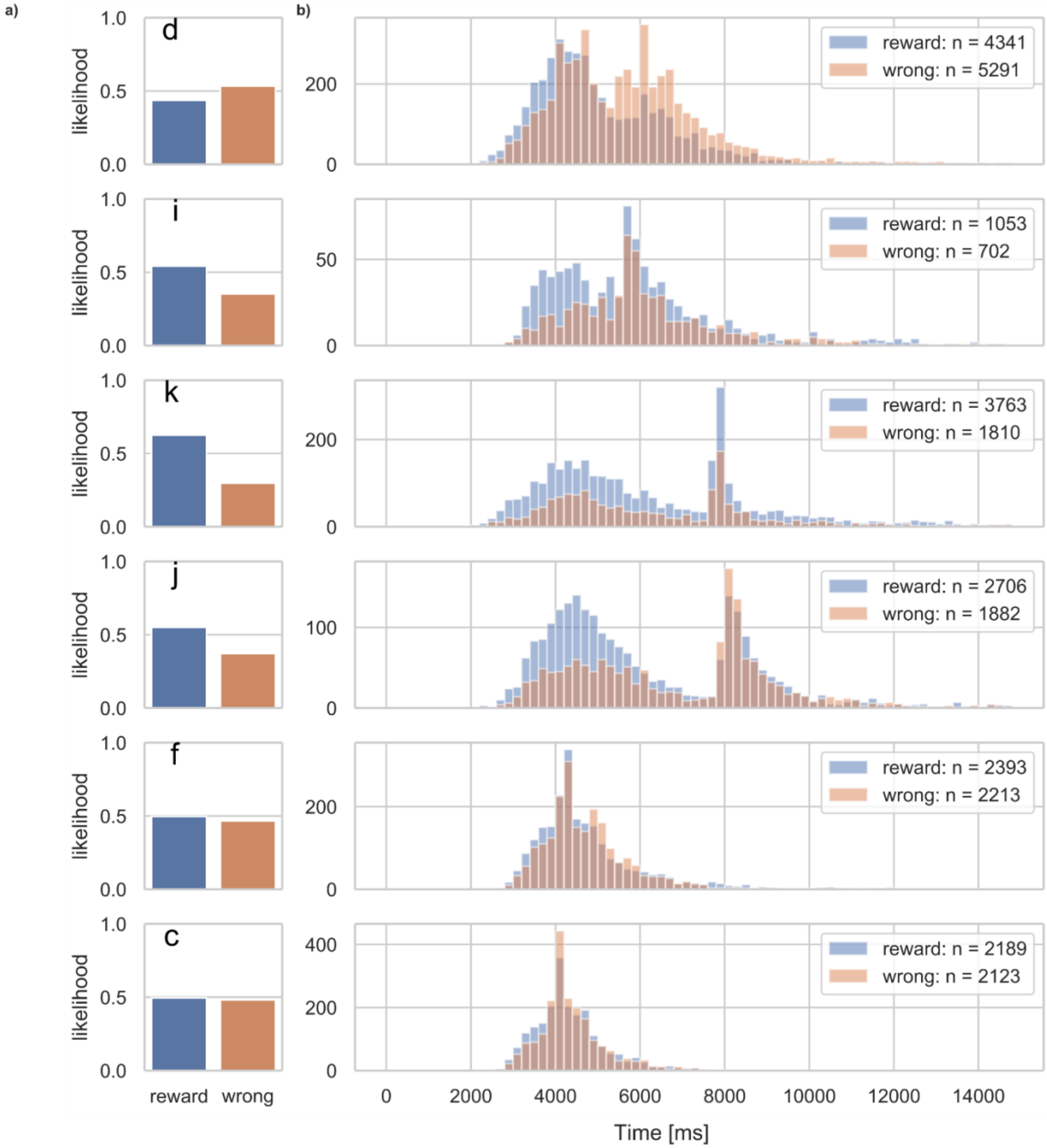


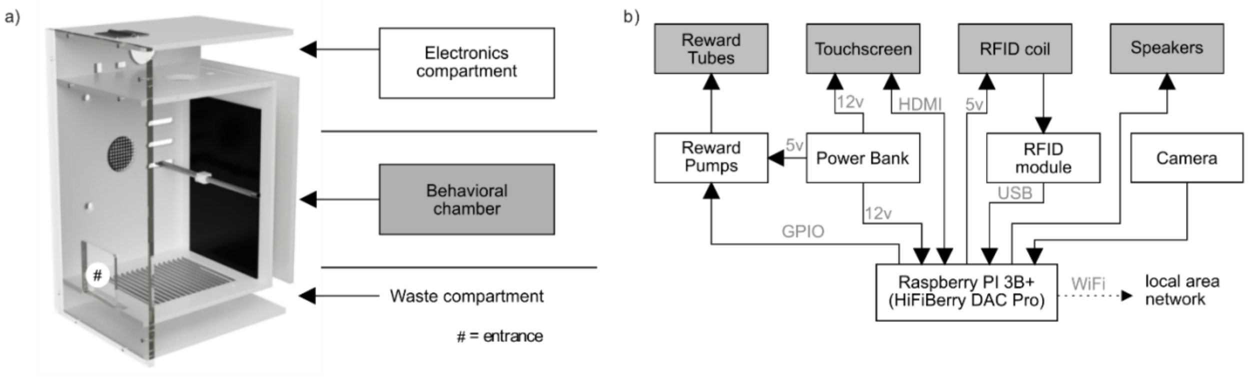
Figure S2.3.Visualization of Inter-Trial Intervals in the automates unsupervised training (AUT – version 10). a) *bar plots indicating, for each animal, the likelihood of a new trial being initiated within 30 seconds from a correct (blue) or a wrong (orange) previous trial. Letters in the bar plots indicate the animals.* ***b)*** *histograms reporting the distribution of inter-trial-interval of each of the 6 animals. Average and standard deviation are given in Table S4.*

For animals that underwent the final AUT (version 10) and performed the acoustic discrimination task described in the main text, we quantified the likelihood of initiating a trial after a correct or a wrong response (within 30 seconds) and analysed the time (in seconds) between consecutive trials for correct and wrong trials separately (Figure S3). Note that after a wrong response a timeout of 2.5 to 5 seconds was used, such that new trials could not be initiated and touches were ignored before the timeout ended. For correct responses instead, the trial was available after a time interval of 0.8 to 2.5 seconds. As a result, we observed differences in likelihood of initiating a new trial (Figure S3A) and in the distribution of inter-trial intervals (Figure S3B) after correct vs. wrong responses in animals who ultimately acquired the acoustic discrimination task (animals d, i, k, j).

Animal	Mean Correct	Mean Wrong	std Correct	std Wrong
c	4604.57	4626.39	3040.61	3010.72
d	5700.33	6474.76	4384.6	4748.27
f	5123.12	5210.68	3533.57	4020.49
i	7041.53	7798.84	6594.45	7522.4
j	7459.02	8142.14	6103.64	5879.97
k	7702.85	7748.64	6614.6	6070.68
total	6271.9	6666.91	5045.24	5208.75

Table S2.4 – Average inter-trial-intervals (in milliseconds) and standard deviation across animals in the AUT (version 10), for correct and wrong trials.

5. Additional device, training and trial timing information



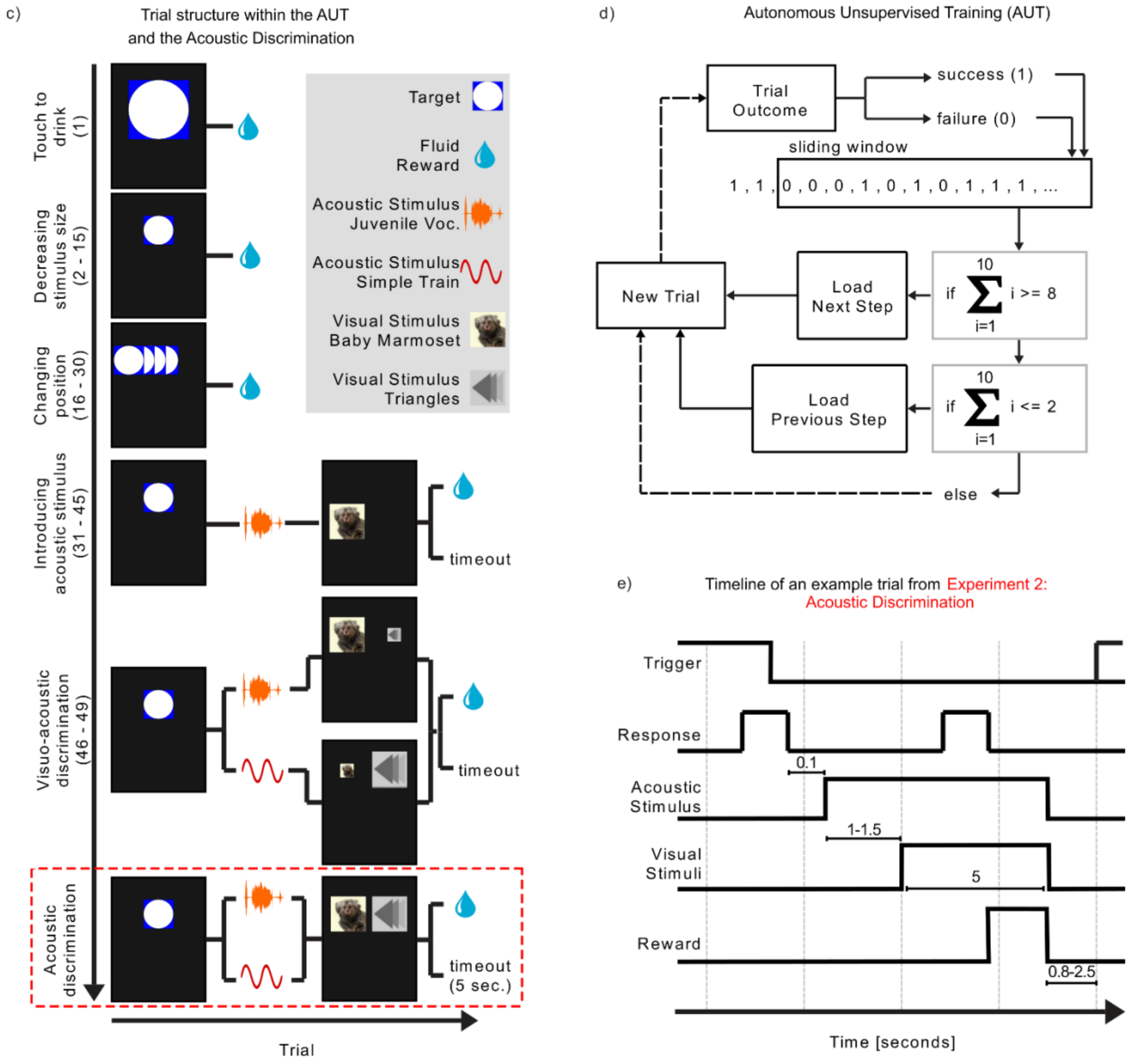


Figure S2.4. a) *rendering of the MXBI without a side panel to show the three internal compartments.* ***b)*** *schematics of the connections and flow of information between individual components of the behavioral chamber (grey background) and of the electronics compartment (white background).* ***c)*** *basic trial structure from the animals' perspective, across six exemplary steps of the automated protocol (AUT).* ***d)****schematic representation of the logic of the recursive algorithm responsible for monitoring the animal's trial-by-trial performance and changing the step accordingly.* ***e)*** *example trial timeline (including the animal's response) for a successful trial at the last step of the AUT protocol, the audio-visual association.*

2.10 References

Abbott, D.H., Barnett, D.K., Colman, R.J., Yamamoto, M.E., Schultz-Darken, N.J., 2003. Aspects of common marmoset basic biology and life history important for biomedical research. Comp. Med. 53, 339–350.

Agamaite, J.A., Chang, C.-J., Osmanski, M.S., Wang, X., 2015. A quantitative acoustic analysis of the vocal repertoire of the common marmoset (*Callithrix jacchus*). The Journal of the Acoustical Society of America 138, 2906–2928. https://doi.org/10.1121/1.4934268

Andrews, M.W., Rosenblum, L.A., 1994. Automated recording of individual performance and hand preference during joystick-task acquisition in group-living bonnet macaques (Macaca radiata). Journal of Comparative Psychology, 108, 358–362.

Aoki, R., Tsubota, T., Goya, Y., Benucci, A., 2017. An automated platform for high-throughput mouse behavior and physiology with voluntary head-fixation. Nat Commun 8, 1196. https://doi.org/10.1038/s41467-017-01371-0

Archakov, D., DeWitt, I., Kuśmierek, P., Ortiz-Rios, M., Cameron, D., Cui, D., Morin, E.L., VanMeter, J.W., Sams, M., Jääskeläinen, I.P., Rauschecker, J.P., 2020. Auditory representation of learned sound sequences in motor regions of the macaque brain. Proc Natl Acad Sci USA 117, 15242–15252. https://doi.org/10.1073/pnas.1915610117

Bennett, A.J., Bailoo, J.D., Dutton, M., Michel, G.F., Pierre, P.J., 2018. Psychological Science Applied to Improve Captive Animal Care: A Model for Development of a Systematic, Evidence-Based Assessment Of Environmental Enrichment for Nonhuman Primates. https://doi.org/10.31234/osf.io/79xky

Bennett, A.J., Perkins, C.M., Tenpas, P.D., Reinebach, A.L., Pierre, P.J., 2016. Moving evidence into practice: cost analysis and assessment of macaques’ sustained behavioral engagement with videogames and foraging devices. Am J Primatol 78, 1250–1264. https://doi.org/10.1002/ajp.22579

Berger, M., Calapai, A., Stephan, V., Niessing, M., Burchardt, L., Gail, A., Treue, S., 2018. Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. Journal of Neurophysiology 119, 796–807. https://doi.org/10.1152/jn.00614.2017

Bethell, E.J., Cassidy, L.C., Brockhausen, R.R., Pfefferle, D., 2019. Toward a Standardized Test of Fearful Temperament in Primates: A Sensitive Alternative to the Human Intruder Task for Laboratory-Housed Rhesus Macaques (Macaca mulatta). Front. Psychol. 10, 1051. https://doi.org/10.3389/fpsyg.2019.01051

Bohlen, M., Hayes, E.R., Bohlen, B., Bailoo, J.D., Crabbe, J.C., Wahlsten, D., 2014. Experimenter effects on behavioral test scores of eight inbred mouse strains under the influence of ethanol. Behavioural Brain Research 272, 46–54. https://doi.org/10.1016/j.bbr.2014.06.017

Brosch, M., Selezneva, E., Bucks, C., Scheich, H., 2004. Macaque monkeys discriminate pitch relationships. Cognition 91, 259–272. https://doi.org/10.1016/j.cognition.2003.09.005

Brown, S.D.M., Moore, M.W., 2012. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. Mammalian Genome 23, 632–640. https://doi.org/10.1007/s00335-012-9427-x

Brydges, N.M., Braithwaite, V.A., 2008. **Measuring Animal Welfare: What Can Cognition Contribute?** Annu Rev Biomed Sci 10, T91–T103. https://doi.org/10.5016/1806-8774.2008.v10pT91

Butler, J.L., Kennerley, S.W., 2018. Mymou: A low-cost, wireless touchscreen system for automated training of nonhuman primates. Behavior Research Methods. https://doi.org/10.3758/s13428-018-1109-5

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., Gail, A., 2017a. A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. Behavior Research Methods 49, 35–45. https://doi.org/10.3758/s13428-016-0707-3

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., Gail, A., 2017b. A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. Behavior Research Methods 49, 35–45. https://doi.org/10.3758/s13428-016-0707-3

Calapai, Antonino, Cabrera-Moreno, Jorge, Moser, Tobias, Jeschke, Marcus, 2022. Flexible auditory training, psychophysics, and enrichment of common marmosets with an automated, touchscreen-based system: dataset and code. GitHub. https://doi.org/10.5281/ZENODO.6139297

Choi, J.Y., Takahashi, D.Y., Ghazanfar, A.A., 2015. Cooperative vocal control in marmoset monkeys via vocal feedback. Journal of Neurophysiology 114, 274–283. https://doi.org/10.1152/jn.00228.2015

Cibulski, L., Wascher, C.A.F., Weiß, B.M., Kotrschal, K., 2014. Familiarity with the experimenter influences the performance of Common ravens (Corvus corax) and Carrion crows (Corvus corone corone) in cognitive tasks. Behavioural Processes 103, 129–137. https://doi.org/10.1016/j.beproc.2013.11.013

Clark, F., 2017. Cognitive enrichment and welfare: Current approaches and future directions. AB&C 4, 52–71. https://doi.org/10.12966/abc.05.02.2017

Clarke, H.F., Walker, S.C., Dalley, J.W., Robbins, T.W., Roberts, A.C., 2007. Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. Cerebral Cortex 17, 18–27. https://doi.org/10.1093/cercor/bhj120

Crofts, H.S., Muggleton, N.G., Bowditch, A.P., Pearce, P.C., Nutt, D.J., Scott, E.A.M., 2016. Home cage presentation of complex discrimination tasks to marmosets and rhesus monkeys: Lab. Anim. 33, 207–214. https://doi.org/10.1258/002367799780578174

Dean, I., Harper, N.S., McAlpine, D., 2005. Neural population coding of sound level adapts to stimulus statistics. Nature Neuroscience 8, 1684–1689. https://doi.org/10.1038/nn1541

Drummer, C., Vogt, E.-J., Heistermann, M., Roshani, B., Becker, T., Mätz-Rensing, K., Kues, W.A., Kügler, S., Behr, R., 2021. Generation and Breeding of EGFP-Transgenic Marmoset Monkeys: Cell Chimerism and Implications for Disease Modeling. Cells 10, 505. https://doi.org/10.3390/cells10030505

Evans, T.A., Beran, M.J., Chan, B., Klein, E.D., Menzel, C.R., 2008. An efficient computerized testing method for the capuchin monkey (Cebus apella): Adaptation of the LRC-CTS to a socially housed nonhuman primate species. Behavior Research Methods 40, 590–596. https://doi.org/10.3758/BRM.40.2.590

Fagot, J., Paleressompoulle, D., 2009. Automatic testing of cognitive performance in baboons maintained in social groups. Behavior Research Methods 41, 396–404. https://doi.org/10.3758/BRM.41.2.396

Feng, G., Jensen, F.E., Greely, H.T., Okano, H., Treue, S., Roberts, A.C., Fox, J.G., Caddick, S., Poo, M., Newsome, W.T., Morrison, J.H., 2020. Opportunities and limitations of genetically modified nonhuman primate models for neuroscience research. Proc Natl Acad Sci USA 117, 24022–24031. https://doi.org/10.1073/pnas.2006515117

Fishman, Y.I., Kim, M., Steinschneider, M., 2017. A Crucial Test of the Population Separation Model of Auditory Stream Segregation in Macaque Primary Auditory Cortex. J. Neurosci. 37, 10645–10655. https://doi.org/10.1523/JNEUROSCI.0792-17.2017

Gazes, R.P., Brown, E.K., Basile, B.M., Hampton, R.R., 2013. Automated cognitive testing of monkeys in social groups yields results comparable to individual laboratory-based testing. Anim Cogn 16, 445–458. https://doi.org/10.1007/s10071-012-0585-8

Grabenhorst, F., Tsutsui, K.-I., Kobayashi, S., Schultz, W., 2019. Primate prefrontal neurons signal economic risk derived from the statistics of recent reward experience. eLife 8, e44838. https://doi.org/10.7554/eLife.44838

Gultekin, Y.B., Hage, S.R., 2017. Limiting parental feedback disrupts vocal development in marmoset monkeys. Nature Communications 8, 14046. https://doi.org/10.1038/ncomms14046

Harris, J.C., 2021. Animal models of neurodevelopmental disorders with behavioral phenotypes. Current Opinion in Psychiatry 34, 87–93. https://doi.org/10.1097/YCO.0000000000000675

Hawash, M.B.F., Sanz-Remón, J., Grenier, J.-C., Kohn, J., Yotova, V., Johnson, Z., Lanford, R.E., Brinkworth, J.F., Barreiro, L.B., 2021. Primate innate immune responses to bacterial and viral pathogens reveals an evolutionary trade-off between strength and specificity. Proc Natl Acad Sci USA 118, e2015855118. https://doi.org/10.1073/pnas.2015855118

Herrmann, B., Augereau, T., Johnsrude, I.S., 2020. Neural Responses and Perceptual Sensitivity to Sound Depend on Sound-Level Statistics. Sci Rep 10, 9571. https://doi.org/10.1038/s41598-020-66715-1

Hirokawa, J., Sadakane, O., Sakata, S., Bosch, M., Sakurai, Y., Yamamori, T., 2011. Multisensory Information Facilitates Reaction Speed by Enlarging Activity Difference between Superior Colliculus Hemispheres in Rats. PLoS ONE 6, e25283. https://doi.org/10.1371/journal.pone.0025283

Hirst, R.J., Cragg, L., Allen, H.A., 2018. Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. Neuroscience & Biobehavioral Reviews 94, 286–301. https://doi.org/10.1016/j.neubiorev.2018.07.012

Hosoya, M., Fujioka, M., Ogawa, K., Okano, H., 2016. Distinct Expression Patterns Of Causative Genes Responsible For Hereditary Progressive Hearing Loss In Non-Human Primate Cochlea. Scientific Reports 6. https://doi.org/10.1038/srep22250

Issa, E.B., Schmidt, K.M., Ohayon, S., DiCarlo, J.J., 2016. A simple, wireless system for remote, high-throughput behavioral testing of nonhuman primates, in: Society for Neuroscience Meeting Abstracts 2016.

Jensen, G., Alkan, Y., Ferrera, V.P., Terrace, H.S., 2019. Reward associations do not explain transitive inference performance in monkeys. Sci. Adv. 5, eaaw2089. https://doi.org/10.1126/sciadv.aaw2089

Jezzini, A., Bromberg-Martin, E.S., Trambaiolli, L.R., Haber, S.N., Monosov, I.E., 2021. A prefrontal network integrates preferences for advance information about uncertain rewards and punishments. Neuron 109, 2339-2352.e5. https://doi.org/10.1016/j.neuron.2021.05.013

Johnson, L.A., Della Santina, C.C., Wang, X., 2016. Selective Neuronal Activation by Cochlear Implant Stimulation in Auditory Cortex of Awake Primate. The Journal of Neuroscience 36, 12468–12484. https://doi.org/10.1523/JNEUROSCI.1699-16.2016

Johnson, L.A., Della Santina, C.C., Wang, X., 2012. Temporal bone characterization and cochlear implant feasibility in the common marmoset (Callithrix jacchus). Hear Res 290, 37–44. https://doi.org/10.1016/j.heares.2012.05.002

Kangas, B.D., Bergman, J., 2012. A novel touch-sensitive apparatus for behavioral studies in unrestrained squirrel monkeys. Journal of Neuroscience Methods 209, 331–336. https://doi.org/10.1016/j.jneumeth.2012.06.028

Keating, P., Nodal, F.R., Gananandan, K., Schulz, A.L., King, A.J., 2013. Behavioral Sensitivity to Broadband Binaural Localization Cues in the Ferret. JARO 14, 561–572. https://doi.org/10.1007/s10162-013-0390-3

Keppeler, D., Kampshoff, C.A., Thirumalai, A., Duque-Afonso, C.J., Schaeper, J.J., Quilitz, T., Töpperwien, M., Vogl, C., Hessler, R., Meyer, A., Salditt, T., Moser, T., 2021. Multiscale photonic imaging of the native and implanted cochlea. PNAS 118. https://doi.org/10.1073/pnas.2014472118

Knyazeva, S., Selezneva, E., Gorkin, A., Aggelopoulos, N.C., Brosch, M., 2018. Neuronal Correlates of Auditory Streaming in Monkey Auditory Cortex for Tone Sequences without Spectral Differences. Front. Integr. Neurosci. 12, 4. https://doi.org/10.3389/fnint.2018.00004

Kobayashi, S., Schultz, W., 2008. Influence of Reward Delays on Responses of Dopamine Neurons. Journal of Neuroscience 28, 7837–7846. https://doi.org/10.1523/JNEUROSCI.1600-08.2008

Kumita, W., Sato, Kenya, Suzuki, Y., Kurotaki, Y., Harada, T., Zhou, Y., Kishi, N., Sato, Kengo, Aiba, A., Sakakibara, Y., Feng, G., Okano, H., Sasaki, E., 2019. Efficient generation of Knock-in/Knock-out marmoset embryo via CRISPR/Cas9 gene editing. Sci Rep 9, 12719. https://doi.org/10.1038/s41598-019-49110-3

Lohse, M., Bajo, V.M., King, A.J., Willmore, B.D.B., 2020. Neural circuits underlying auditory contrast gain control and their perceptual implications. Nat Commun 11, 324. https://doi.org/10.1038/s41467-019-14163-5

Mandell, D.J., Sackett, G.P., 2008. A computer touch screen system and training procedure for use with primate infants: Results from pigtail monkeys (Macaca nemestrina). Dev. Psychobiol. 50, 160–170. https://doi.org/10.1002/dev.20251

Margiotoudi, K., Allritz, M., Bohn, M., Pulvermüller, F., 2019. Sound symbolic congruency detection in humans but not in great apes. Sci Rep 9, 12705. https://doi.org/10.1038/s41598-019-49101-4

Marini, R.P., Wachtman, L.M., Tardif, S.D., Mansfield, K., 2018. The Common Marmoset in Captivity and Biomedical Research.

McKinney, W., 2010. Data Structures for Statistical Computing in Python. Presented at the Python in Science Conference, Austin, Texas, pp. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Mehrpour, V., Martinez-Trujillo, J.C., Treue, S., 2020. Attention amplifies neural representations of changes in sensory input at the expense of perceptual accuracy. Nat Commun 11, 2128. https://doi.org/10.1038/s41467-020-15989-0

Millman, K.J., Aivazis, M., 2011. Python for Scientists and Engineers. Comput. Sci. Eng. 13, 9–12. https://doi.org/10.1109/MCSE.2011.36

Morrill, R.J., Thomas, A.W., Schiel, N., Souto, A., Miller, C.T., 2013. The effect of habitat acoustics on common marmoset vocal signal transmission: Habitat Acoustics of Common Marmosets. American Journal of Primatology 75, 904–916. https://doi.org/10.1002/ajp.22152

Murphy, E.S., McSweeney, F.K., Smith, R.G., McComas, J.J., 2003. Dynamic changes in reinforcer effectiveness: theoretical, methodological, and practical implications for applied research. Journal of Applied Behavior Analysis 36, 421–438. https://doi.org/10.1901/jaba.2003.36-421

Nakamura, K., Koba, R., Miwa, M., Yamaguchi, C., Suzuki, H., Takemoto, A., 2018. A Method to Train Marmosets in Visual Working Memory Task and Their Performance. Front. Behav. Neurosci. 12, 46. https://doi.org/10.3389/fnbeh.2018.00046

Okano, H., 2021. Current Status of and Perspectives on the Application of Marmosets in Neurobiology. Annu. Rev. Neurosci. 44, 27–48. https://doi.org/10.1146/annurev-neuro-030520-101844

Okano, H., Kishi, N., 2018. Investigation of brain science and neurological/psychiatric disorders using genetically modified non-human primates. Current Opinion in Neurobiology 50, 1–6. https://doi.org/10.1016/j.conb.2017.10.016

O'Leary, J.D., O'Leary, O.F., Cryan, J.F., Nolan, Y.M., 2018. A low-cost touchscreen operant chamber using a Raspberry Pi^{TM}. Behav Res 50, 2523–2530. https://doi.org/10.3758/s13428-018-1030-y

Oliphant, T.E., 2007. Python for Scientific Computing. Comput. Sci. Eng. 9, 10–20. https://doi.org/10.1109/MCSE.2007.58

Oliphant, T.E., 2006. A guide to Numpy. USA: Trelgol Publishing.

Osmanski, M.S., Song, X., Guo, Y., Wang, X., 2016. Frequency discrimination in the common marmoset (Callithrix jacchus). Hearing Research 341, 1–8. https://doi.org/10.1016/j.heares.2016.07.006

Osmanski, M.S., Wang, X., 2011. Measurement of absolute auditory thresholds in the common marmoset (Callithrix jacchus). Hear. Res. 277, 127–133. https://doi.org/10.1016/j.heares.2011.02.001

Park, J.E., Zhang, X.F., Choi, S.-H., Okahara, J., Sasaki, E., Silva, A.C., 2016. Generation of transgenic marmosets expressing genetically encoded calcium indicators. Scientific Reports 6. https://doi.org/10.1038/srep34931

Poddar, R., Kawai, R., Ölveczky, B.P., 2013. A Fully Automated High-Throughput Training System for Rodents. PLoS ONE 8, e83171. https://doi.org/10.1371/journal.pone.0083171

Prescott, M.J., 2020. Ethical and Welfare Implications of Genetically Altered Non-Human Primates for Biomedical Research. J. Appl. Anim. Ethics Res. 1–26. https://doi.org/10.1163/25889567-BJA10002

Prescott, M.J., Langermans, J.A., Ragan, I., 2017. Applying the 3Rs to non-human primate research: Barriers and solutions. Drug Discovery Today: Disease Models 23, 51–56. https://doi.org/10.1016/j.ddmod.2017.11.001

Rabinowitz, N.C., Willmore, B.D.B., Schnupp, J.W.H., King, A.J., 2011. Contrast Gain Control in Auditory Cortex. Neuron 70, 1178–1191. https://doi.org/10.1016/j.neuron.2011.04.030

Richardson, W.K., Washburn, D.A., Hopkins, W.D., Savage-rumbaugh, E.S., Rumbaugh, D.M., 1990. The NASA/LRC computerized test system. Behavior Research Methods, Instruments, & Computers 22, 127–131. https://doi.org/10.3758/BF03203132

Risueno-Segovia, C., Hage, S.R., 2020. Theta Synchronization of Phonatory and Articulatory Systems in Marmoset Monkey Vocal Production. Current Biology 30, 4276-4283.e3. https://doi.org/10.1016/j.cub.2020.08.019

Rivalan, M., Munawar, H., Fuchs, A., Winter, Y., 2017. An Automated, Experimenter-Free Method for the Standardised, Operant Cognitive Testing of Rats. PLoS ONE 12, e0169476. https://doi.org/10.1371/journal.pone.0169476

Rogers, J., Gibbs, R.A., 2014. Comparative primate genomics: emerging patterns of genome content and dynamics. Nat Rev Genet 15, 347–359. https://doi.org/10.1038/nrg3707

Sablé-Meyer, M., Fagot, J., Caparos, S., van Kerkoerle, T., Amalric, M., Dehaene, S., 2021. Sensitivity to geometric shape regularity in humans and baboons: A putative signature of human singularity. Proc Natl Acad Sci USA 118, e2023123118. https://doi.org/10.1073/pnas.2023123118

Sadoun, A., Rosito, M., Fonta, C., Girard, P., 2018. Key periods of cognitive decline in a non-human primate model of cognitive aging, the common marmoset (Callithrix jacchus). Neurobiology of Aging 1–40. https://doi.org/10.1016/j.neurobiolaging.2018.10.003

Sasaki, E., Suemizu, H., Shimada, A., Hanazawa, K., Oiwa, R., Kamioka, M., Tomioka, I., Sotomaru, Y., Hirakawa, R., Eto, T., Shiozawa, S., Maeda, T., Ito, M., Ito, R., Kito, C., Yagihashi, C., Kawai, K., Miyoshi, H., Tanioka, Y., Tamaoki, N., Habu, S., Okano, H., Nomura, T., 2009. Generation of transgenic non-human primates with germline transmission. Nature 459, 523–527. https://doi.org/10.1038/nature08090

Schaefer, A.T., Claridge-Chang, A., 2012. The surveillance state of behavioral automation. Current Opinion in Neurobiology 22, 170–176. https://doi.org/10.1016/j.conb.2011.11.004

Schmitt, V., Fischer, J., 2009. Inferential reasoning and modality dependent discrimination learning in olive baboons (Papio hamadryas anubis).

Journal of Comparative Psychology 123, 316–325. https://doi.org/10.1037/a0016218

Schmitt, V., Schloegl, C., Fischer, J., 2014. Seeing the Experimenter Influences the Response to Pointing Cues in Long-Tailed Macaques. PLoS ONE 9, e91348. https://doi.org/10.1371/journal.pone.0091348

Schneider, F., Dheerendra, P., Balezeau, F., Ortiz-Rios, M., Kikuchi, Y., Petkov, C.I., Thiele, A., Griffiths, T.D., 2018. Auditory figure-ground analysis in rostral belt and parabelt of the macaque monkey. Sci Rep 8, 17948. https://doi.org/10.1038/s41598-018-36903-1

Schubiger, M.N., Wüstholz, F.L., Wunder, A., Burkart, J.M., 2015. High emotional reactivity toward an experimenter affects participation, but not performance, in cognitive tests with common marmosets (Callithrix jacchus). Anim Cogn 18, 701–712. https://doi.org/10.1007/s10071-015-0837-5

Schultz-Darken, N., Braun, K.M., Emborg, M.E., 2016. Neurobehavioral development of common marmoset monkeys: Marmoset Neurobehavioral Development. Dev Psychobiol 58, 141–158. https://doi.org/10.1002/dev.21360

Schütt, H.H., Harmeling, S., Macke, J.H., Wichmann, F.A., 2016. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. Vision Research 122, 105–123. https://doi.org/10.1016/j.visres.2016.02.002

Shen, H., 2013. Precision gene editing paves way for transgenic monkeys. Nature News 503, 14–15. https://doi.org/10.1038/503014a

Song, X., Osmanski, M.S., Guo, Y., Wang, X., 2016. Complex pitch perception mechanisms are shared by humans and a New World monkey. Proc Natl Acad Sci U S A 113, 781–786. https://doi.org/10.1073/pnas.1516120113

Sorge, R.E., Martin, L.J., Isbester, K.A., Sotocinal, S.G., Rosen, S., Tuttle, A.H., Wieskopf, J.S., Acland, E.L., Dokova, A., Kadoura, B., Leger, P., Mapplebeck, J.C.S., McPhail, M., Delaney, A., Wigerblad, G., Schumann, A.P., Quinn, T., Frasnelli, J., Svensson, C.I., Sternberg, W.F., Mogil, J.S., 2014. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat Methods 11, 629–632. https://doi.org/10.1038/nmeth.2935

Spinelli, S., Pennanen, L., Dettling, A.C., Feldon, J., Higgins, G.A., Pryce, C.R., 2004. Performance of the marmoset monkey on computerized tasks of attention and working memory. Cognitive Brain Research 19, 123–137. https://doi.org/10.1016/j.cogbrainres.2003.11.007

Spinka, M., Newberry, R.C., Bekoff, M., 2001. Mammalian play: training for the unexpected. The Quarterly review of biology 76, 141–168.

Suriya-Arunroj, L., Gail, A., 2019. Complementary encoding of priors in monkey frontoparietal network supports a dual process of decision-making. eLife 8, e47581. https://doi.org/10.7554/eLife.47581

Takahashi, D.Y., Narayanan, D.Z., Ghazanfar, A.A., 2013. Coupled oscillator dynamics of vocal turn-taking in monkeys. Curr. Biol. 23, 2162–2168. https://doi.org/10.1016/j.cub.2013.09.005

Takemoto, A., Izumi, A., Miwa, M., Nakamura, K., 2011. Development of a compact and general-purpose experimental apparatus with a touch-sensitive screen for use in evaluating cognitive functions in common marmosets. Journal of Neuroscience Methods 199, 82–86. https://doi.org/10.1016/j.jneumeth.2011.04.029

Tarou, L.R., Bashaw, M.J., 2007. Maximizing the effectiveness of environmental enrichment: Suggestions from the experimental analysis of behavior. Applied Animal Behaviour Science 102, 189–204. https://doi.org/10.1016/j.applanim.2006.05.026

Tay, S.-K., Blythe, J., Lipovich, L., 2009. Global discovery of primate-specific genes in the human genome. Proceedings of the National Academy of Sciences 106, 12019–12024. https://doi.org/10.1073/pnas.0904569106

Thiele, A., Bellgrove, M.A., 2018. Neuromodulation of Attention. Neuron 97, 769–785. https://doi.org/10.1016/j.neuron.2018.01.008

Tokuno, H., Moriya-Ito, K., Tanaka, I., 2012. Experimental Techniques for Neuroscience Research Using Common Marmosets. Experimental Animals 61, 389–397.

Tomasello, M., Call, J., 2011. Methodological Challenges in the Study of Primate Cognition. Science 334, 1227–1228. https://doi.org/10.1126/science.1213443

Tomioka, I., Nagai, Y., Seki, K., 2020. Generation of Common Marmoset Model Lines of Spinocerebellar Ataxia Type 3. Front. Neurosci. 14, 548002. https://doi.org/10.3389/fnins.2020.548002

Tomioka, I., Nogami, N., Nakatani, T., Owari, K., Fujita, N., Motohashi, H., Takayama, O., Takae, K., Nagai, Y., Seki, K., 2017. Generation of transgenic marmosets using a tetracyclin-inducible transgene expression system as a neurodegenerative disease model†. Biology of Reproduction 97, 772–780. https://doi.org/10.1093/biolre/iox129

Van Essen, D.C., Donahue, C.J., Coalson, T.S., Kennedy, H., Hayashi, T., Glasser, M.F., 2019. Cerebral cortical folding, parcellation, and connectivity in humans, nonhuman primates, and mice. Proc Natl Acad Sci USA 116, 26173–26180. https://doi.org/10.1073/pnas.1902299116

Walker, J.D., Pirschel, F., Gidmark, N., MacLean, J.N., Hatsopoulos, N.G., 2020. A platform for semiautomated voluntary training of common marmosets for behavioral neuroscience. J. Neurophysiol. 123, 1420–1426. https://doi.org/10.1152/jn.00300.2019

Washburn, D.A., Hopkins, W.D., Rumbaugh, D.M., 1989. Video-task assessment of learning and memory in macaques (Macaca mulatta): effects of stimulus movement on performance. Journal of Experimental Psychology: Animal Behavior Processes 15, 393–400.

Washburn, D.A., Rumbaugh, D.M., 1992. Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. Behavior Research Methods, Instruments, & Computers 24, 157–164. https://doi.org/10.3758/BF03203490

Wen, B., Wang, G.I., Dean, I., Delgutte, B., 2009. Dynamic Range Adaptation to Sound Level Statistics in the Auditory Nerve. J. Neurosci. 29, 13797–13808. https://doi.org/10.1523/JNEUROSCI.5610-08.2009

Wikman, P., Rinne, T., Petkov, C.I., 2019. Reward cues readily direct monkeys' auditory performance resulting in broad auditory cortex modulation and interaction with sites along cholinergic and dopaminergic pathways. Sci Rep 9, 3055. https://doi.org/10.1038/s41598-019-38833-y

Witham, C.L., 2018. Automated face recognition of rhesus macaques. Journal of Neuroscience Methods 300, 157–165. https://doi.org/10.1016/j.jneumeth.2017.07.020

Wittmann, M.K., Fouragnan, E., Folloni, D., Klein-Flügge, M.C., Chau, B.K.H., Khamassi, M., Rushworth, M.F.S., 2020. Global reward state affects learning and activity in raphe nucleus and anterior insula in monkeys. Nat Commun 11, 3771. https://doi.org/10.1038/s41467-020-17343-w

Chapter 3

Automated audiograms, vocal perception assessment, and cochlear implant feasibility testing for common marmosets (*Callithrix jacchus*)

Cabrera-Moreno J. Grunenberg J, Jeschke M.

in preparation

Chapter 3 describes a series of four experiments that expand the flexibility of use of our developed platform, the MXBI. Experiment 1 details an automated unsupervised training protocol that successfully instructed common marmosets into a pure tone detection task. Experiment 2 describes the acquisition of audiogram measurements for five marmosets. Experiment 3 details the implementation of cochlear implant stimulation in freely moving marmosets and their performance in a detection task. Experiment 4 describes a vocal discrimination task where marmosets flexibly identified and grouped a set of 945 acoustic stimuli into two groups.

Author contributions

JCM and MJ conceived and designed the research. JG and JCM design Experiment 1 with the input of MJ. JG and JCM collected the data for Experiment 1. JCM designed Experiments 2, 3, and 4 with the input of MJ and collected the data for Experiments 2, 3, and 4. Finally, JCM analyzed the data for Experiments 1, 2, 3, and 4. MJ supervised the project.

Abstract

Automation in psychoacoustics for common marmosets (*Callithrix jacchus*) represents a step towards optimizing a labor-intensive process. Chapter 2 described a novel paradigm that successfully trained common marmosets in an unsupervised manner to discriminate vocalizations of conspecific from artificial stimuli directly in their home cage without the need for food or water control nor social separation. However, the ability of our system to instruct tasks that did not rely on naturalistic sounds, such as vocalizations, was not yet demonstrated. Chapter 3 reports a series of experiments that demonstrate the flexibility of use of our system by instructing marmosets to solve several auditory tasks relying on pure tones, synthesized vocalizations, and electric stimulation through cochlear implants.

Experiment 1 describes a novel automated unsupervised training protocol used to train 11 common marmosets in a pure tone detection task based on a Go-NoGo paradigm. Subsequently, Experiment 2 details the characterization of hearing thresholds based on pure tones for 5 marmosets, showing similarities to known published audiogram profiles for this species. Experiment 3 demonstrates the feasibility of wireless control of cochlear implants and its use for assessing electrical stimulation thresholds of an example marmoset individual. Finally, Experiment 4 explores the perceptual boundaries of a marmoset vocalization by experimentally assessing the ability of 3 common marmosets to generalize and categorize a set of 945 synthetic stimuli.

Taken together, these results ultimately demonstrate that our novel automated training and testing approach can achieve sophisticated psychophysics measurements, assess the efficacy of cochlear implants wirelessly, and answer critical questions on auditory cognition.

3.1 Introduction

The common marmoset (*Callithrix jacchus*), due to its rich vocal repertoire and communicative use of vocalizations, has emerged as a promising model for auditory neuroscience (Kishi et al., 2014; Mansfield, 2003; Miller et al., 2016). Its auditory capabilities have been extensively described using more traditional methods (Bendor and Wang, 2005; Osmanski et al., 2016; Osmanski and Wang, 2011; Risueno-Segovia and Hage, 2020; Song et al., 2016; Takahashi et al., 2013) and have been suggested to be a fundamental model for hearing loss and cochlear implant research (Hosoya et al., 2021, 2016; Johnson et al., 2016, 2012; Keppeler et al., 2021). Furthermore, with the possibility of transgenic lines (Belmonte et al., 2015; Kaiser and Feng, 2015), it will soon be necessary to phenotype a large number of animals. Despite the high value of marmosets as models for understanding the mechanisms of hearing and other cognitive capabilities, few studies have developed techniques to improve data collection in a more efficient and reliable manner.

The automation of animal training and testing offers the possibility to reduce human labor, improve data replicability, and scale up data collection (King et al., 2009). Furthermore, when combined with a cage-based approach, these protocols can not only create a motivational advantage that might have a positive impact during learning (Andrews and Rosenblum, 1994; Gazes et al., 2013; Washburn et al., 1989) but also expand the range of behaviors that can be investigated (e.g., social behavior, non-restrained motor behavior) and minimize animals stress, improving animal wellbeing (Bonini, 2019; Clark, 2017; Fagot et al., 2014; Fagot and Paleressompoulle, 2009; Mason et al., 2019).

Previously, in Chapter 2, we demonstrated that common marmosets could be trained to perform psychoacoustics experiments on a cage-based using an automated unsupervised training procedure (Calapai et al., 2022). Nonetheless, several other aspects of sound perception remained to be addressed.

In order for our system to be used as an alternative to traditional training and assessing methods, I argue that it must fulfill 3 requirements. 1) show similar engagement as in traditional methods, 2) demonstrate the flexibility of use across various tasks, and 3) collect data that resembles those using traditional training methods.

To this end I report the results of a series of 4 experiments that demonstrates that our approach can 1) automatically train common marmosets to perform a pure tone detection task. 2) assess hearing thresholds of pure tones in several marmosets and obtain comparable measurements to existing reports in the available literature (Osmanski and Wang, 2011). 3) wirelessly control cochlear implants in freely moving marmosets and collect electrical stimulation thresholds. 4) investigate the effects of spectro-temporal feature variations on acoustic perception and classification of 945 stimuli. Establishing the validity of the automated protocols for conducting state-of-the-art psychoacoustic experiments.

3.2 Results

In this study 17 adult common marmosets of either sex and housed in pairs participated across a series of 4 experiments. Animals were generally trained in pairs on auditory tasks with a single MXBI attached to their home cage and without any restriction to fluid or food. Apart from the initial training phase that requires manual training (see below), all sessions ran autonomously, relying on an RFID module to identify the animals.

General engagement

To have an overview of the general engagement of the animals towards our device across experiments, I first assessed the number of trials and their distribution within and across sessions for each animal. Animals *il*, *bc*, *dg*, *bs*, and *dm* did not overcome the *initial training* (see Methods) phase. As such, these animals are not considered in further analysis.

Across all animals, substantial variability was found in the number of trials, with a median of 219 trials (Q1=67 Q2=351) per session (when randomly taking 50 sessions per animal) (Figure 3.1a). Except for pair *wl*-*dc* (where *dc* died for natural reasons during data collection), all marmoset pairs shared an equal number of sessions, determined by the number of days the device was offered to the animals. Animals *iv* had a partner that did not participate in the following experiments but did have access to the device (see Methods). Within a total of 1659 sessions across animals (except *oh* and *pr*), a mean of 32 sessions without trials was found (Figure 3.1b). A detailed number of sessions and trials per individual animal can be found in Table 3.1. The median distribution of session duration was 3.48 h (Figure 3.1c). Additionally, consistent with the data shown in Chapter 2, marmosets remained engaged across sessions regardless of the testing protocol and showed a slight preference to perform more trials within the first third of the session (Figure 3.1d, 4.1e).

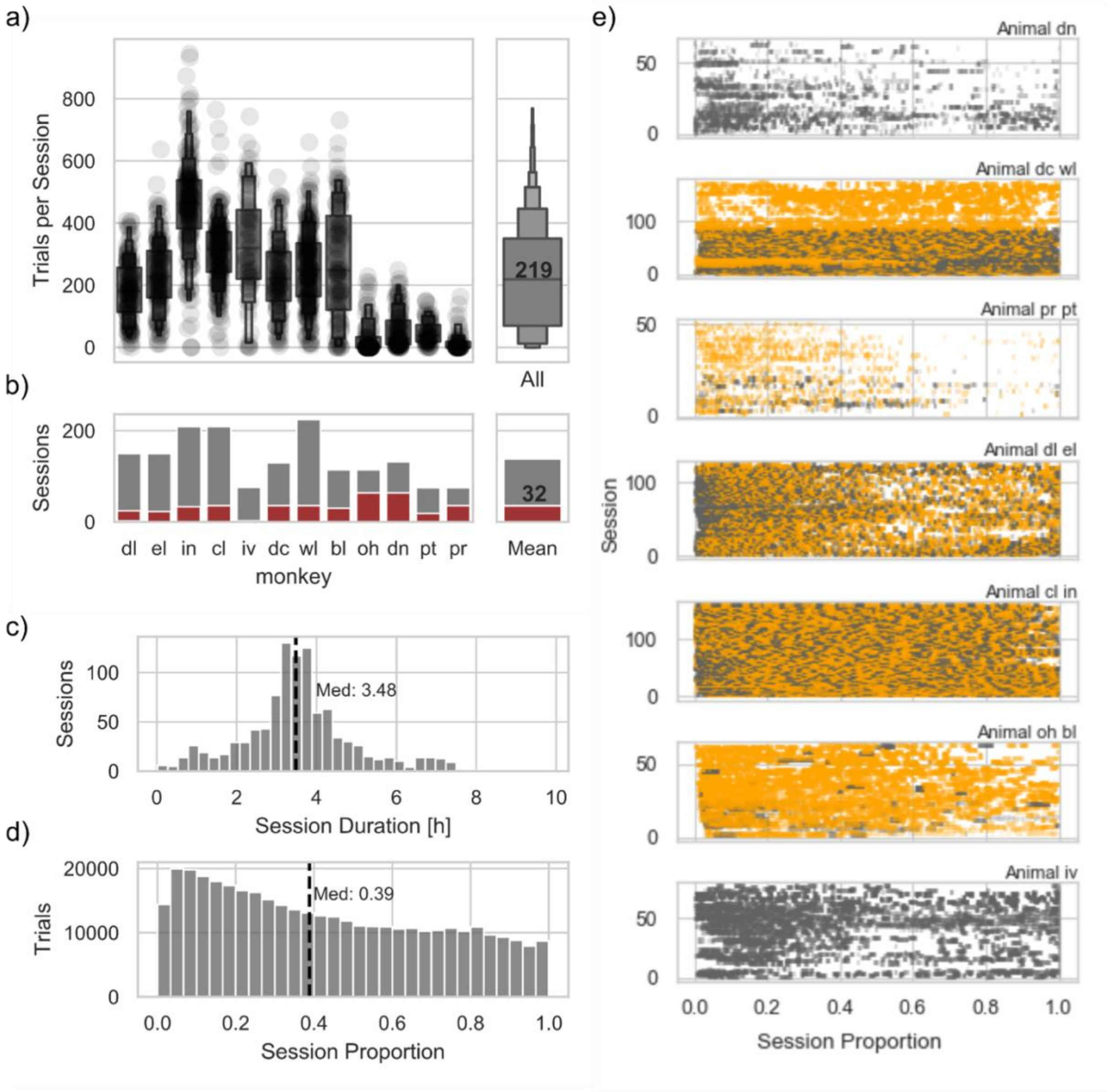


Figure 3.1.General engagement across Experiments 1, 2, 3, and 4. a) *Letter-value plots the number of trials across sessions per animal. The right panel shows the distribution of 50 sessions randomly picked per animal. The central box represents the distribution's first, second, and third quartile (Q1=67, Q2=215, Q3=351 trials). In Table 3.1, the exact number of trials per animal can be found .* ***b)*** *Number of sessions across animals. Red bars represent the number of sessions with 0 trials. The right panel shows the mean across animals.* ***c)*** *Distribution of session duration (Median = 3.44 h).* ***d)*** *Distribution of trials timestamp as a function of session proportion.* ***e)*** *Raster plot for each MXBI showing the trials' timestamp as a function of session proportion. Grey fort he first animal and orange for the second, if available.*

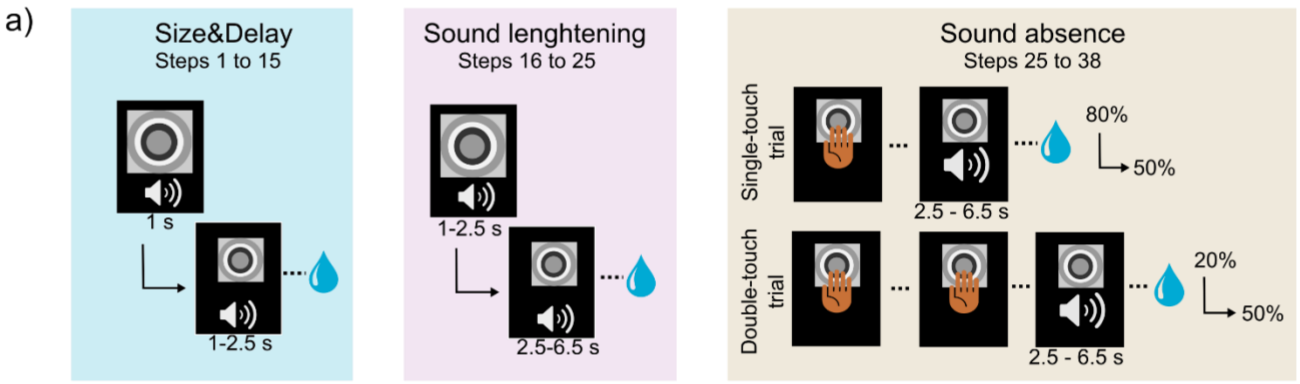
Name	Cage mate	Sex	Age [months]	Weight [g]	Number of trials	Number of sessions	Number of sessions with 0 trials	Naive	Initial training	Experiment order	Cochlear implanted
el	*dl*	m	51	401	31592	150	21			1, 4	
dl	*el*	f	66	485	25433	150	19			1, 4	
wl	*dc, vn**	m	70	410	55618	225	35			1, 2, 3	✓
dc	*wl*	f	51	503	30112	130	36			1, 2	
in	*cl*	m	57	419	95355	209	33			1, 2, 3	✓
cl	*in*	f	48	425	65157	209	31			1, 2	✓
iv	*dr*	f	54	371	24972	76	2			4	
dn	*bs*	m	5	347	7360	132	65	✓	✓	1	
bs	*dn*	m	71	380	n.a.	n.a.	n.a.	✓	✗	1	✓
pr	*pt*	f	16	402	1186	75	36	✓	✓	1	
pt	*pr*	m	16	382	3617	75	15	✓	✓	1	
dm	*dg*	m	21	450	n.a.	n.a.	n.a.	✓	✗	1	
dg	*dm*	*m*	15	362	n.a.	n.a.	n.a.	✓	✗	1	
bl	*oh*	f	40	406	31137	114	27	✓	✓	1, 2	
oh	*bl*	m	72	384	3340	114	64	✓	✓	1	
bc	*il*	m	109	376	n.a.	n.a.	n.a.	✓	✗	1	
il	*bc*	f	31	405	n.a.	n.a.	n.a.	✓	✗	1	

Table 3.1 General animal data*. Each row depicts individual animal data across the fields of each column. The start in animal vn indicates the pair of animal vn with animal wl after the dead of animal dc. The age and weight correspond to the beginning of the experiments. n.a. stands for not available data. Ticks in the Naive column indicate the animals that did not have experience with touchscreen tasks previous to the experiments. Ticks in the column Intial training indicate that the animal overcame the initial training, and the crosses indicate the opposite. Ticks in the column Cochlear implanted indicate those animals that received a cochlear implant unilaterally.*

Experiment 1 – Sound detection

Experiment 1 aimed to train common marmosets in a pure tone detection task relying on the MXBI. Here, 10 out of the 16 total animals tested did not have experience with touchscreens and underwent a non-automated training procedure (see Methods - Initial training) to introduce them to the basics of touchscreen interaction. However, out of these 10 animals, 5 (*bs, dm, dg, bc*, and *il*) did not overcome the initial training and, consequently, did not participate in any of the following experiments. The low interest and motivation these animals showed towards the MXBI might explain the lack of success in learning the interaction rules with the device. The remaining 5 animals (*dn, pr, pt, bl*, and *oh*), together with the 6 with previous touchscreen experience (*el, dl, wl, dc, in*, and *cl*) – for a total of 11 animals, underwent an automated unsupervised training (AUT) protocol comprised of 38 steps, grouped into three milestones (see Methods – Experiment 1) that instructed a pure tone detection task (Figure 3.2a).

Animals navigated the AUT procedure with low variability in the hit rate (Figure 3.2b), suggesting smooth learning progress. Figure 3.2c shows the learning curves for the 11 animals as a function of the percentage of their trials across the 38 steps of the AUT. Except for animals *pr* and *dn* that did not overcome milestones sound to touch and sound lengthening, respectively, 9 animals successfully reached the last step (step 38) with a hit rate above 75%. Animals *el* and *dl* were housed together, and during the first sessions, a problem with the RFID module hindered their identification. Overall, a median of 400, 475, and 900 trials, and 3, 5, and 13 sessions were needed to overcome the three milestones (Figure 3.2d). Finally, the likelihood of trial initiation after a given outcome ('correct' or 'wrong') across steps remained stable, with a clear difference between the likelihood of initiating a trial within the following 30 s after a 'correct' versus a 'wrong' response (mean likelihood after a 'correct' trial = 0.87, after a 'wrong' trial = 0.17) (Figure 3.2e), suggesting that trials' outcome modulate animals' engagement.


a)
Size&Delay
Steps 1 to 15
1 s
1-2.5 s
Sound lenghtening
Steps 16 to 25
1-2.5 s
2.5-6.5 s
Sound absence
Steps 25 to 38
Single-touch trial
2.5 - 6.5 s
80%
50%
Double-touch trial
2.5 - 6.5 s
20%
50%

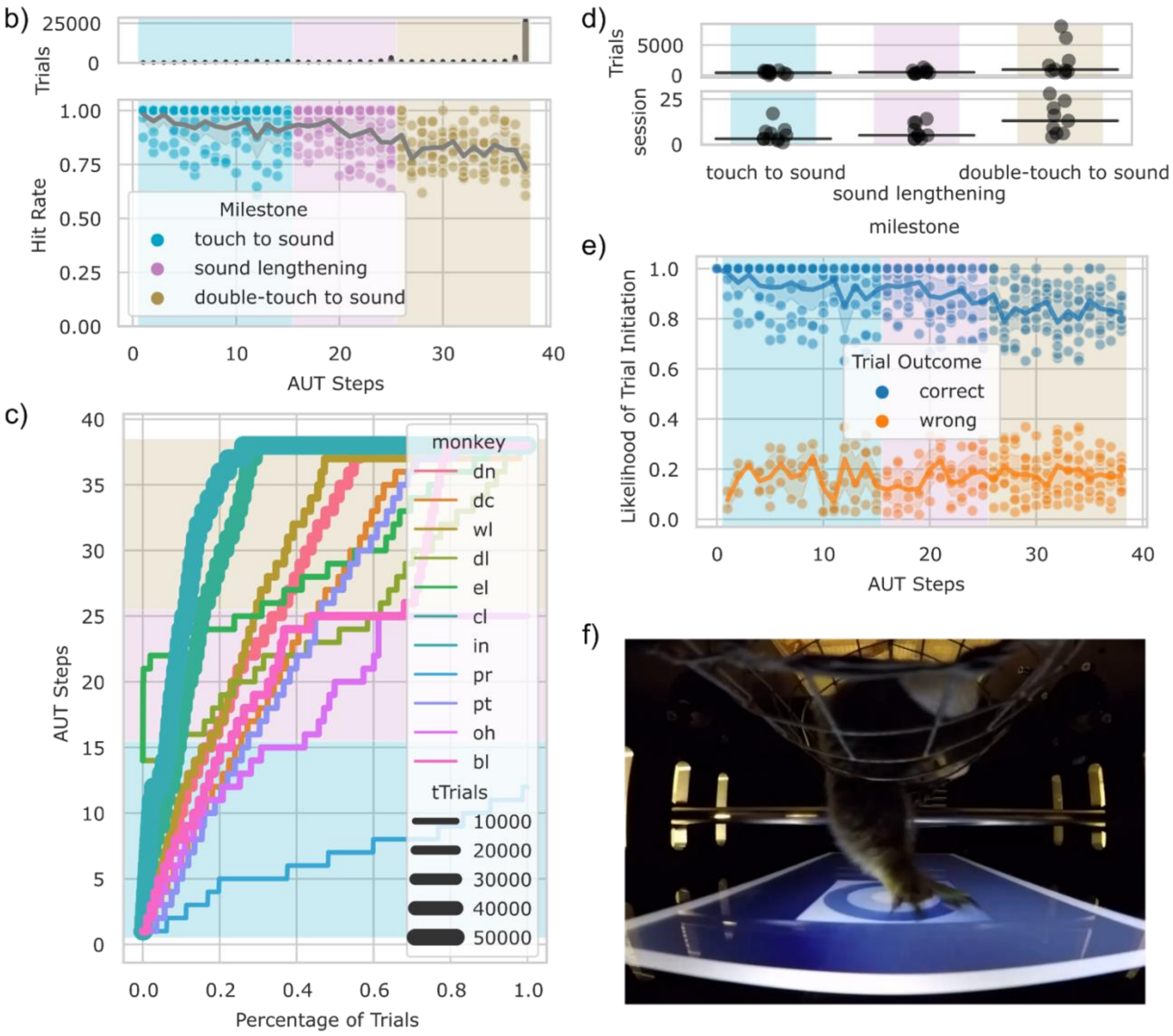


Figure 3.2. Automated unsupervised training protocol for sound detection. a) *Schematic representation of the three milestones of the automated unsupervised training procedure (AUT).* ***b)*** *Upper panel, bar plot showing the number of trials across steps of the AUT. The lower panel shows the average hit rate as a function of steps for all 11 animals. The grey line represents the mean across all animals, and the grey area represents 95% confidence interval of the mean.* ***c)*** *Individual animal progress through the AUT. The thickness of the lines represents the total amount of trials per animal. The background colors indicate the three milestones.* ***d)*** *Distribution of the number of trials and sessions per animal across the three milestones.* ***e)*** *Likelihood of trial initiation as a function of step. The colored background corresponds with the three milestones of the staircase . The lines represent the mean across animals, and the colored area around the lines represents 95% confidence interval of the mean.* ***f)*** *Picture of animal dc performing a trial taken from the surveillance camera in the MXBI.*

Experiment 2 – Hearing thresholds

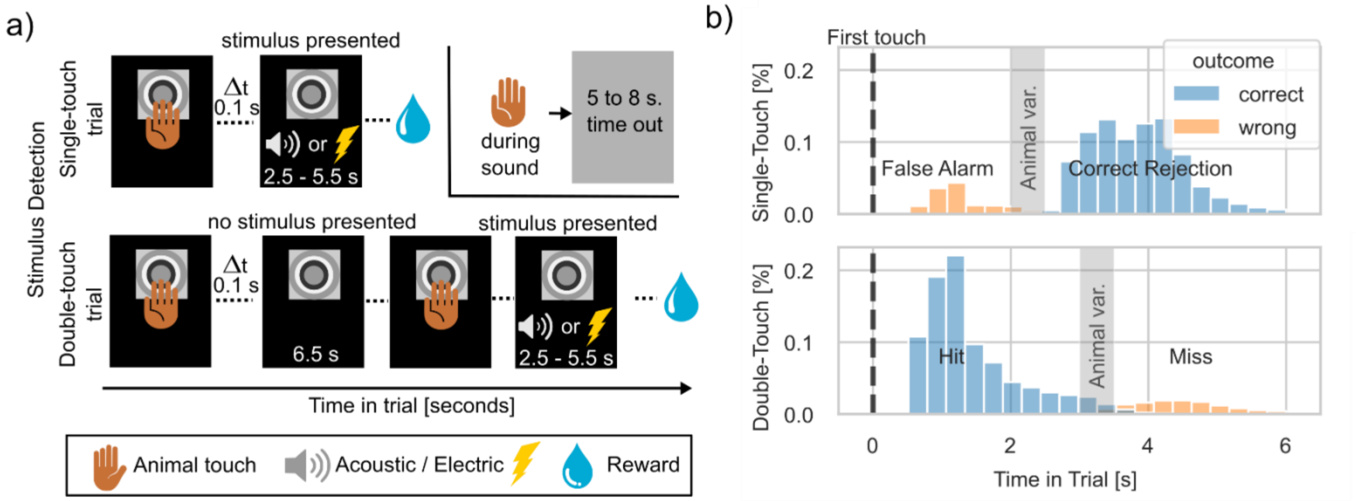
Experiment 2 aimed to measure the audibility curves of 5 marmosets (*in, cl, dc, wl,* and *bl*), relying on the pure tone detection task described in Experiment 1. Shortly, the pure tone detection task was based on a Go-NoGo paradigm comprising two trial types: single and double touch trials. In single-touch trials, a pure tone train was triggered after touching a target on the screen. Double-touch trials required the animals to touch the target on the screen twice to trigger the sound. The animals' task was to detect the presence of a pure tone after the first touch and then decide whether a second touch was required. During the sound presentation (2.5-5.5 s), animals had to withhold additional touches to receive the reward (0.1 ml) delivered upon sound ending. Additional touches during sound presentation aborted the trial, leading to a timeout of 5-8 s, after which a new trial started (Figure 3.1a) (see Methods - Experiment 2).

Animals were required to hold a hit rate above 0.75 and a false alarm rate below 0.20 for at least three consecutive sessions before testing sessions took place. Figure 3.3b shows the response distribution of the 5 animals as a function of time in both trials.

For the testing sessions, pure tone stimuli of 7 frequencies were used (0.2 kHz, 1 kHz, 2 kHz, 5 kHz, 7 kHz, 10 kHz, and 20 kHz). Marmosets were randomly presented with target trials (with intensities of 10 dB SPL, 20 dB SPL, 30 dB SPL, 40 dB SPL, 50dB SPL, and 60 dB SPL) and anchor trials (with intensities of 70 dB SPL and 80 dB SPL). Target trials were held only in single-touch trials and represented 20% of the total.

Across all testing sessions, animals showed a stable engagement and performance with a median of 151 trials per session (Q1=80, Q3=286, trials), 0.88 hit rate (Q1=0.82, Q3=0.92), and 0.06 error rate (Q1=0.04, Q3=0.12) calculated from 50 randomly selected sessions per animal. Out of a total of 597 sessions, 8, 11, 1, 12, and 2 sessions with 0 trials were found for animals *in, cl, dc, wl,* and *bl*, respectively (Figure 3.3c). Figure 3.3d shows the hit rates of pure tones as a function of stimulus amplitude across all 5 frequencies for all animals. Here, as tone intensity decreased, the hit rate also decreased. A minimum of 35 trials per intensity on each frequency were required to estimate the hearing thresholds across animals (Table 3.2). Figure 3.3e shows concomitant changes in the response latency as a function of stimulus amplitude. The latency to touch the screen tended to be longer with pure tones of lower amplitude, which might reflect the difficulty in hearing those stimuli against the background. However, no statistically significant difference was found between the average response latency of the highest (80 dB SPL Mean = 1.3 s) and the lowest (10dB SPL Mean = 1.5 s) amplitude. Figure 3.3f shows the average discriminability indices (d') for all sessions of each frequency. Overall, marmosets showed a lower d' for the frequencies with low sensitivity and higher d' for those frequencies with high sensitivity (median values = 2.5 for 0.2 kHz, 2.6 for 1 kHz, 3.1 for 2 kHz, 3.1 for 5 kHz, 3.0 for 7 kHz, 2.8 for 10 kHz, and 2.7 for 20 kHz).

Across all animals, individual audiograms showed the highest sensitivity for 1 kHz, 2 kHz, 5 kHz, and 7 kHz, and the lowest for 0.2 kHz, 10 kHz, and 20 kHz (upper panel Figure 3.3g). Except for animal *in* that showed a sensitivity increase of 6 dB from 0.2 kHz to 1 kHz, animals *bl, cl, wl,* and *bl* showed a sensitivity increase of ~20 dB. All animals showed a sensitivity drop of ~20 dB from 7 kHz to 10 kHz and 20 kHz. For animals *in* and *dc*, 2 kHz showed the highest sensitivity, while for animals *cl*, *wl*, and *bl*, 7 kHz showed the highest sensitivity. It is essential to highlight that animals *in* and *cl* were single-sided cochlear implanted and deafened (left ear) by the time the hearing thresholds were collected. For animal *wl* just sessions for 5 kHz were collected after implantation, which might explain the low sensitivity at that frequency. The averaged audiogram from the 5 marmosets shows the typical W-shape sensitivity curve for New-world monkeys with two sensitive peaks at 2 kHz with 22 dB SPL and 7 kHz with 20.5 dB SPL, with a decrease in between, at 5 kHz with 27.8 dB SPL (lower panel Figure 3.3g). The average audiogram from the current study shows a general drop of 11 dB in sensitivity across all tested frequencies compared to previously measured marmoset audibility curves (Osmanski and Wang, 2011) (sensitivity decreased – 0.2 kHz = 5.19 dB SPL, 1 kHz = 12.04 dB SPL, 2 kHz = 3.87 dB SPL, 5 kHz = 13.2 dB SPL, 7 kHz = 13.7 dB SPL, 10 kHz = 18.29 dB SPL, 20 kHz = 15 dB SPL).


a)
Stimulus Detection
Single-touch trial
stimulus presented
Δt 0.1 s
or
2.5 - 5.5 s
during sound
5 to 8 s. time out
Double-touch trial
no stimulus presented
stimulus presented
Δt 0.1 s
6.5 s
or
2.5 - 5.5 s
Time in trial [seconds]
Animal touch
Acoustic / Electric
Reward
b)
First touch
Single-Touch [%]
0.2
0.1
0.0
False Alarm
Animal var.
Correct Rejection
outcome
correct
wrong
Double-Touch [%]
0.2
0.1
0.0
Hit
Animal var.
Miss
0
2
4
6
Time in Trial [s]

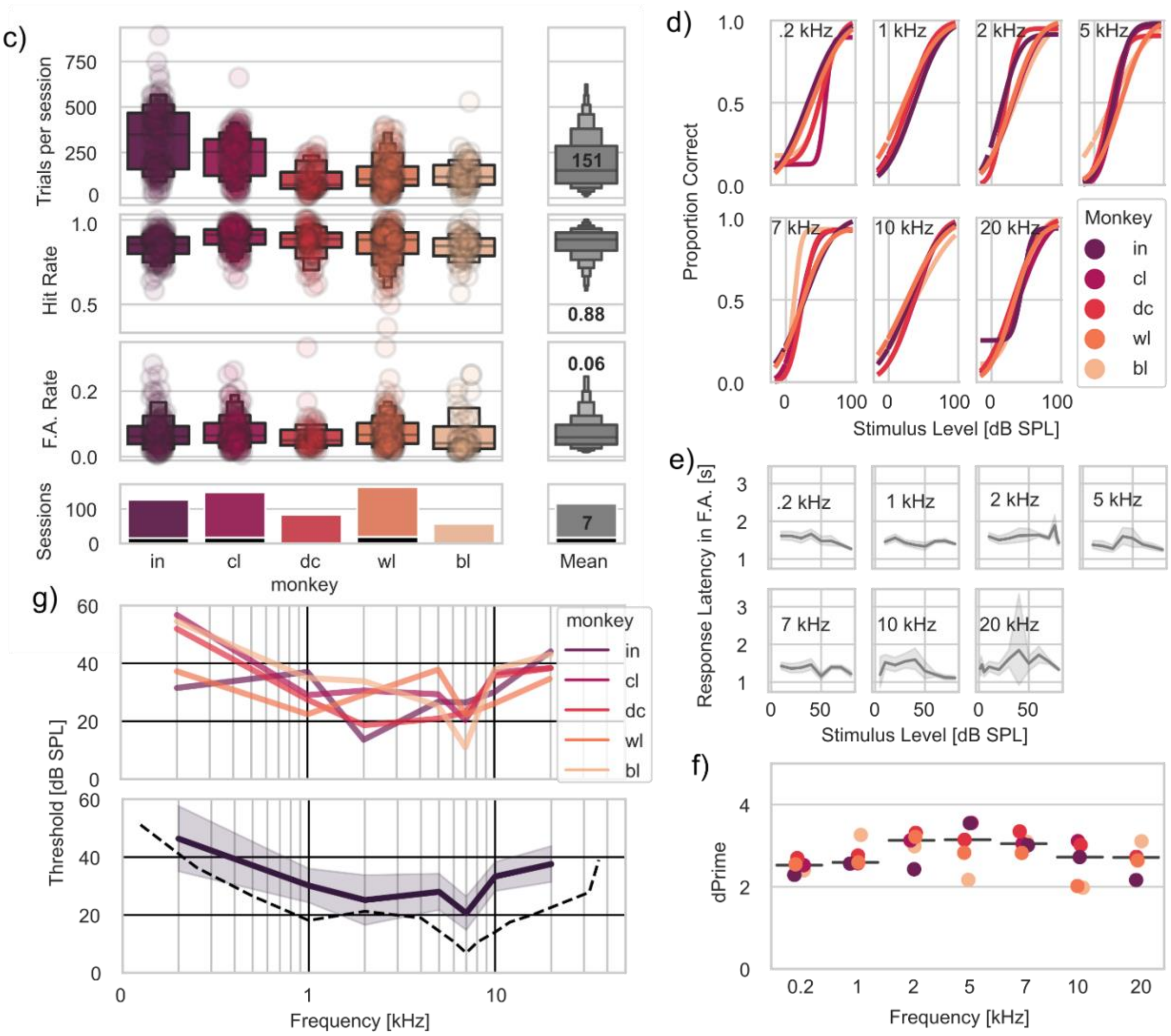


Figure 3.3. Performance during hearing thresholds. a) *Schematic representation of the Go-NoGo task used to collect the hearing thresholds. Two types of trials required the animals to touch either once or twice on the screen depending on the presence or absence of sound. The lightning symbol indicates the possibility of performing the same task using electric stimulation through a cochlear implant (see Methods- Experiment 3).* ***b)*** *Response distribution across all animals for the two types of trials. The Grey area indicates the variability in thresholds for each animal (see Methods – Experiment 2).* ***c)*** *Letter-value plots the number of trials (upper panel), hit rates (mid -upper panel), and false alarm rates (mid -lower) across the total number of sessions (lower panel) per animal. The right grey panels show the distribution of 50 sessions randomly picked per animal. The central box represents the distribution's first, second, and third quartiles. Trials per session False alarm rate* ***d)*** *Psychometric functions for all frequencies tested across animals, calculated as the proportion of correct responses as a function of stimulus intensity.* ***e)*** *Average response latencies across frequencies. The grey area represents the 95% confidence interval.* ***f)*** *Average discriminability indices across all frequencies per animal. e) Individual and average audiograms. The grey area represents the 95% confidence interval. The dotted line shows the average data of 4 marmosets obtained from Osmanski and Wang 2011.*

Animal	.2 kHz	1 kHz	2 kHz	5 kHz	7 kHz	10 kHz	20 kHz
in	38.81	37.01	13.59	27.63	26.45	21.78	43.63
cl	56.72	29.03	30.58	29.21	19.73	56.1	39.97
dc	51.94	27.45	18.61	20.91	22.96	35.64	39.19
wl	37.25	22.41	29.13	51.37	22.24	26.13	36.65
bl	54.54	34.91	18.28	25.29	10.96	37.91	42.38

***Table 3.2 Hearing thresholds**. Shows the exact hearing thresholds for each animal across the tested frequencies.*

Experiment 3. – Cochlear implant assessment

Experiment 3 aimed to expand the use of the MXBI by testing the feasibility of controlling cochlear implants wirelessly to develop automated protocols to assess cochlear implant performance.

Cochlear implants are neuroprostheses that provide users with sound perception by transforming acoustic energy into electrical patterns (relying on a sound processor) that are then delivered through an array of electrodes implanted in the cochlea (Fan-Gang Zeng et al., 2008). Cochlear implants restore hearing by directly stimulating the auditory nerve (Fan-Gang Zeng et al., 2008).

The wireless control of cochlear implants on freely moving marmosets was achieved by relying on the communication protocol described by Jablonski, Harczos, et al. (Jablonski et al., 2020). A custom-coded firmware ESB-USB bridge in the Raspberry Pi enabled wireless communication to the sound processor. Communication stability tests took place in the animal's home cage, with animals carrying the sound processor powered by a fully charged Lithium-ion battery. The communication stability was assessed by requesting the model information from the sound processor every 30 minutes across a session. Successful communication meant receiving the data requested. Four to five hours of stable communication was achieved with our settings, with rare connection losses. However, the connection was usually lost when marmosets were allowed to move farther than the area of a single cage. Animal *wl*, housed in a cage double the size as usual (see - Methods). Therefore for Experiment 3, animal *wl* was restricted to a single cage.

Once the wireless communication stability was established, the flexibility of marmosets relying on electric stimulation to solve auditory tasks was tested. Unilaterally cochlear implanted marmosets *in* and *wl* (see Methods – Experiment 3) underwent a switch of stimulation type, from acoustic to electric, using the same detection task as in Experiment 2. Figure 3.4b shows the hit rate across the sessions where the stimulation type was switched. The hit rate of animal *in* declined promptly upon changing the stimulation type but rapidly rebounded, maintaining a steady level above 0.75 for the remainder of the sessions. In contrast, animal *wl* exhibited no disruption in hit rate when the stimulation switched, indicating that the electrical stimulation through cochlear implants might elicit auditory percept that marmosets swiftly adapt for auditory tasks.

The assessment of the electric stimulation thresholds for all electrodes of the cochlear implant took place once both animals remained with a hit rate above 0.75 and a false alarm below 0.20 for 3 consecutive sessions. During testing sessions, a single electrode was tested at a time with various stimulation intensities (between 10 µA and 150 µA). Target trials with randomized variable stimulation intensities were presented only in the single-touch trials and represented 20% of the total trials.

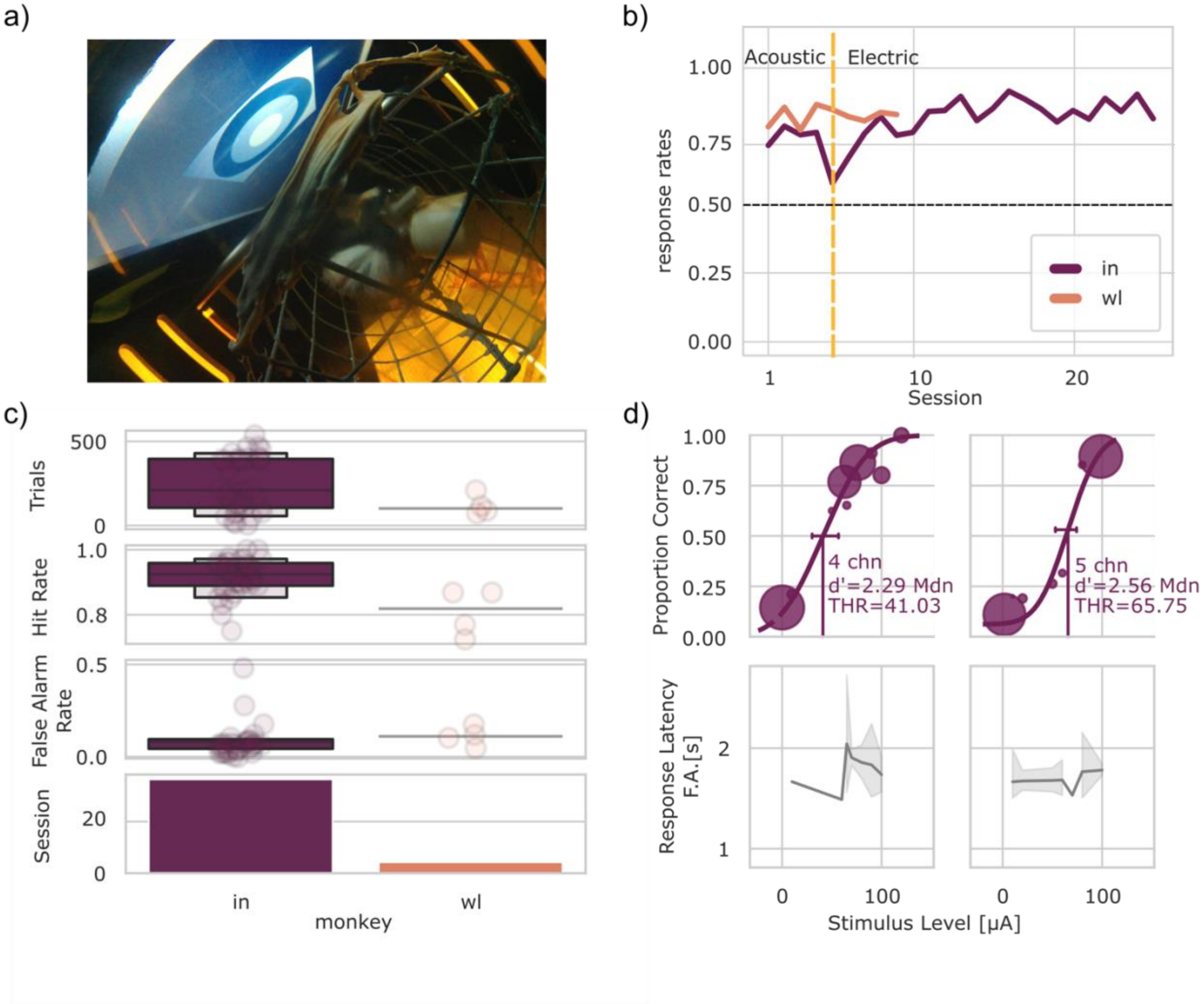
A malfunction of the cochlear implant during data collection led to an incomplete data set for electrodes 4 and 5 for animal *in*. Data for animal *wl* is shown. However, no analysis was performed due to reduce sample size. Data is still under collection.

Across 34 testing sessions for animals *wl* (*in* = 4 sessions) , the engagement (median = 210 trials, Q1=106, Q3=395), hit rate (median = 0.93, Q1=0.89, Q3=0.96), and false alarm rate (median = 0.08, Q1=0.05, Q3=0.09), share strong similarities with Experiment 2 where animals used acoustic stimulation to solve the task (Figure 3.4c). This might suggest the absence of physical discomfort while using the cochlear implant to solve the task. No statistical test was performed due to the low number of sessions.

Electrical stimulation thresholds were calculated as for the number of correct trials across stimulation intensities (upper panel Figure 3.4d). Animal *in* showed a threshold of 41.03µA (95% confidence intervals between 30 µA and 56 µA) in electrode 4 with a median d' = 2.29, and for electrode 5, a threshold of 65.75 µA (95% confidence intervals between 53 µA and 74.3 µA) with a median d' = 2.56. Response latencies for animal *in* showed a steady delay across the tested intensities with a mean difference across the two electrodes of 0.1 s between the lowest and highest intensity (lower panel Figure 3.4d).

The electrical impedances of every electrode were measured at the beginning of every session to monitor their connectivity status (Figure 3.4e,f). High impedance values usually indicate a broken electrode wire or electrode, leading to an open circuit. As such, electrodes with high impedance values were excluded from the testing. For example, electrode 4 of animal *in* showed a drastic increase in impedance values in session 22, going from 3.53 kΩ to 40.37 kΩ, requiring switching the testing electrode to electrode 5. However, 16 sessions after, the impedances of all electrodes increased drastically, above 30 kΩ. This prevented the complete collection of data for animal *in*. Impedances for animal *wl*, remained stable for all electrodes across all tested sessions (Figure 3.3f).

Finally, data from the current study in comparison to those previously collected in our laboratory for animal *in* using a chair-based training procedure in a sound-attenuated chamber showed a slight decrease in sensitivity of 9.82 kΩ and 16.52 kΩ for electrodes 4 and 5, respectively. However, the low amount of data reported in Experiment 3 makes this comparison merely informative.


a)
b)
Acoustic
Electric
1.00
0.75
0.50
0.25
0.00
response rates
in
wl
1
10
20
Session
c)
500
0
Trials
1.0
0.8
Hit Rate
0.5
0.0
False Alarm Rate
20
0
Session
in
wl
monkey
d)
1.00
0.75
0.50
0.25
0.00
Proportion Correct
4 chn
d'=2.29 Mdn
THR=41.03
5 chn
d'=2.56 Mdn
THR=65.75
Response Latency F.A.[s]
2
1
0
100
0
100
Stimulus Level [μA]

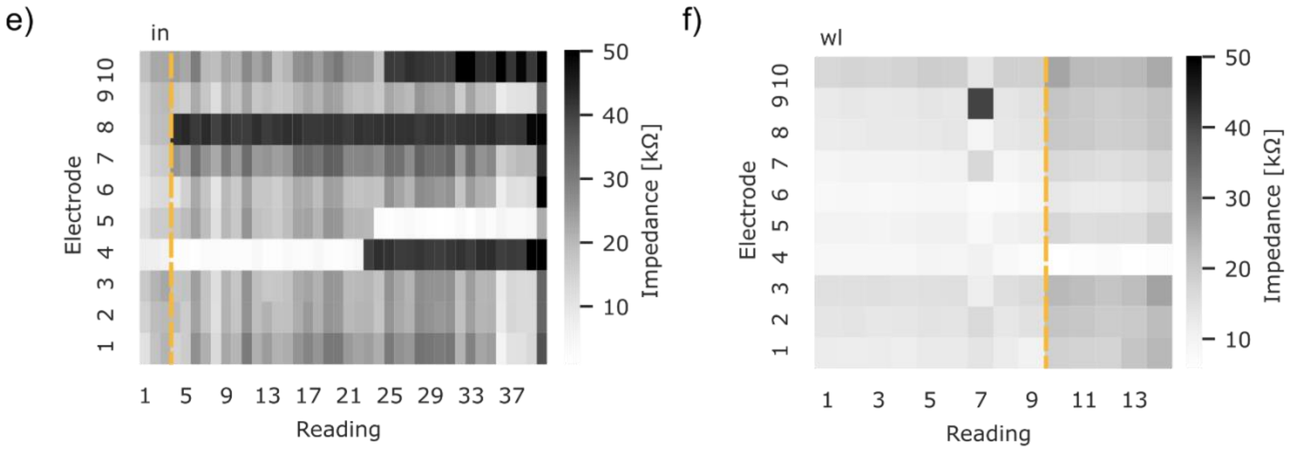


Figure 3.4. Performance during electric hearing thresholds. a) *Picture of animal wl solving the Go-NoGo task using electric stimulation.* ***b)*** *Hit rate of animals in and wl during stimulation switch sessions.* ***c)*** *Letter-value plots in the upper panel the distribution of the number of trials (median = 210, Q1=106, Q3=395), hit rate (median = 0.93, Q1=0.89, Q3=0.96), and false alarm rate (median = 0.08, Q1=0.05, Q3=0.09) across the total number of sessions (barplot lower panel.* ***d)*** *Upper panels show the psychometric function for electrodes 4 and 5 from animal in (electrode 4 = 41.03 µA, 95% confidence intervals = 30 µA - 56 µA, average d' = 2.29, electrode 5 = 65.75 µA, 95% confidence intervals = 53 µA - 74.3 µA, average d' = 2.56). Lower panels show the average response latencies for each electrode. The grey area represents the 95% confidence interval of the mean .* ***e)*** *Heatmap shows the impedance values for each electrode across sessions for animal in. The vertically dotted orange line represents the start of the testing sessions for the collection of electric thresholds.* ***f)*** *Heatmap shows the impedance values for each electrode across sessions for animal wl.*

Experiment 4 – Vocal perception

The common marmoset is an arboreal primate suggested to rely heavily on auditory signals to coordinate complex social behaviors such as group maintenance or territorial defense (Epple, 1968; Hook-costigan and Rogers, 1998; Rosenblum, 1975; Stevenson and Poole, 1976). In the wild and in captivity, marmosets display a variety of spectro-temporally distinct call types which themselves vary between and within individuals. However, how marmosets perceive these signals and how the individual or intertwined combinations of physical features influence their perception is a process that remains to be addressed. Experiment 4 explores the vocal perception capabilities of 3 marmosets *dl*, *el*, and *iv*, by evaluating the behavioral responses to changes in particular acoustic features of synthetic vocalizations (twitters) using a modified version of the 2-alternative choice (2AC) task described in Chapter 2 (see Methods – Experiment 4). Twitter vocalizations are close contact vocalizations uttered typically when visual contact with conspecifics is present and not in the undirect context (Takahashi et al. 2015), which have been previously quantitatively characterized (Agamaite et al. 2015). Twitter calls are intricate stimuli composed of repeated phrases featuring spectrally modulated sweeps. These sweeps exhibit variations in center frequency (the frequency corresponding to the maximum in the spectrum), frequency bandwidth across a call, the number of phrases (representing discernible voicing segments in a call), inter-phrase interval (the average time between consecutive peaks in the envelope), and amplitude-modulated contour (the waveform of the signal), among other factors which combined give rise to an intriguing aspect in vocal communication which is the stochastic variability observed in vocalizations (Agamaite et al. 2015).

To test the importance of call parameters, we generated several hundred synthetic stimuli with specific variations in the spectral and temporal domain. Employing our previously described home-cage automated training and testing system (Calapai et al., 2022), we trained three marmosets to perform an audio-visual discrimination task using a two-alternative choice paradigm. Animals had to classify the synthetic sounds either as a twitter vocalization or as a non-vocalization stimulus.

As demonstrated in Chapter 2 (Results - Generalization to novel stimuli), marmosets *dl, el,* and *iv* could discriminate a twitter vocalization from a pure tone train. Marmosets were first trained to generalize 27 synthetic twitter (see Methods – Experiment 4) and discriminate them from a pure tone (2 s of 2 kHz, with an inter-tone interval of 1 s). Once animals showed a stable discriminability performance (hit rate above 75%) during three consecutive sessions, a series of sequential testing phases took place. Each testing phase held a unique set of target stimuli (referred to now on as complex trains) with particular acoustic physical feature variations (e.g., center frequency, number of harmonics, bandwidth, number of phrases, and inter-phrase interval). Target trials (trials with complex trains with unknown discriminability) represented 10% of the total trials. Anchor trials (27 virtual twitters and 27 complex trains) represented 90% of the total (see Table 3.2 and Table S3.1).

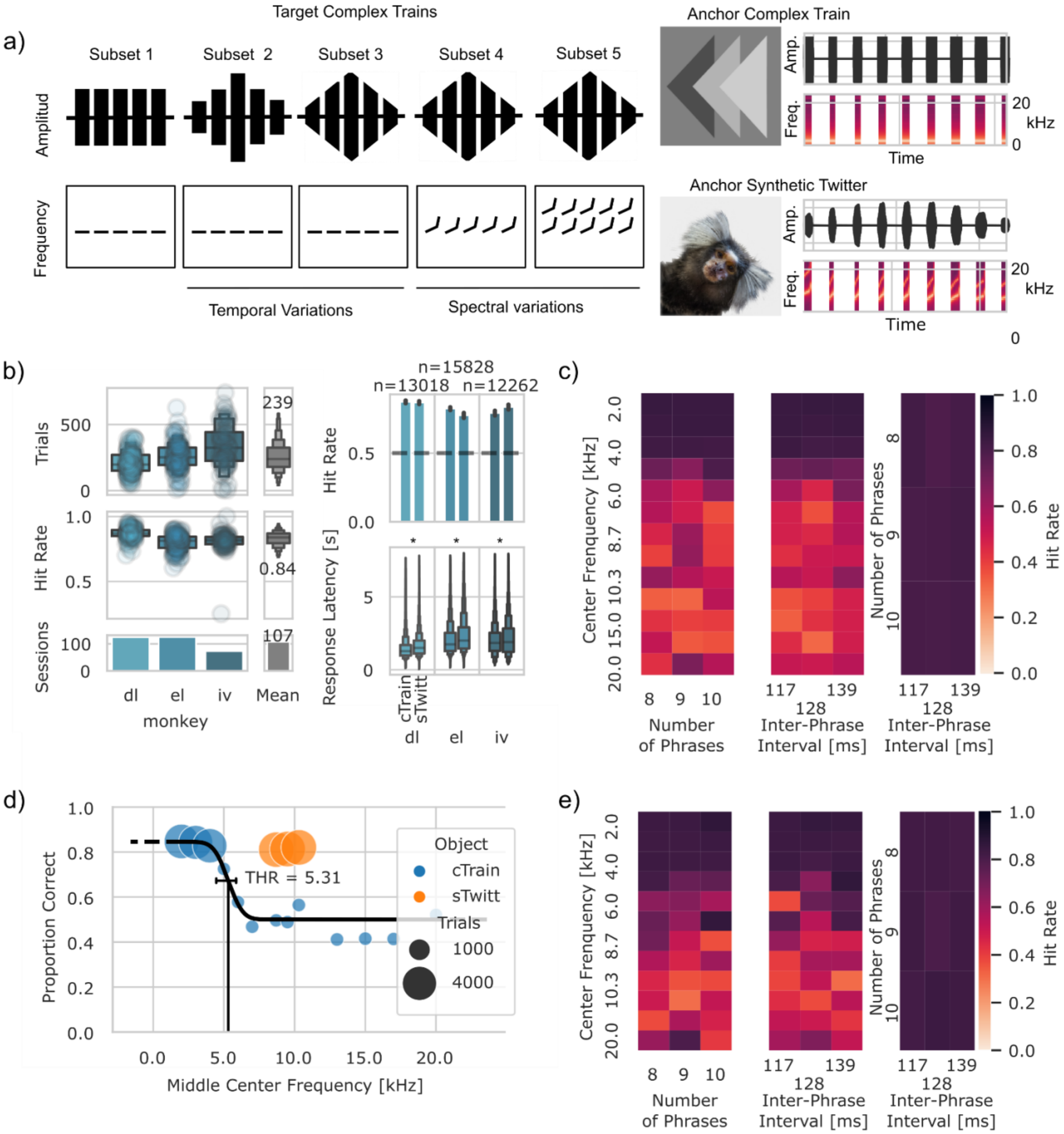
Across the 110, 110, and 73 sessions for animals *dl*, *el*, and *iv*, respectively, showed a stable engagement with a median of 239 (Q1=177, Q3=321) trials per session (based on 50 randomly sampled sessions per animal) (*dl* = 195 Median Q1=150, Q3=278, *el* = 242 Median Q1=189, Q3=314, *iv* = 324 Median Q1=225, Q3=442) (left panel - Figure 3.5b). Similarly, a high and stable hit rate was observed across sessions with a mean of 0.84 (Q1=0.79, Q3=0.86 - based on 50 randomly sampled sessions per animal - *dl* Median = 0.87, Q1=0.86, Q3=0.89, *el* Median = 0.80, Q1=0.75, Q3=0.83, *iv* Median = 0.82, Q1=0.78, Q3=0.84). In 50 randomly sampled sessions, the hit rate per stimulus group (virtual twitter and complex train) remained above 0.75 across all animals (*dl* complex train = 0.87, synthetic twitter = 0.86; *el* complex trains = 0.81, synthetic twitters = 0.76; *iv* complex trains = 0.79, synthetic twitters = 0.83, – right panel Figure 3.5). In addition, the median response latencies showed a significant statistical difference between both stimulus groups for all animals (*dl* complex trains = 1.27 s Q1=1, Q3=1.7, synthetic twitter = 1.5 s Q1=1.1, Q3=2 – Kruskal-Wallis test statistic = 42.02, p = 8.99e-11; *el* complex trains = 1.7 s Q1=1.3, Q3=2.5, synthetic twitters = 2 s Q1=1.5, Q3=2.9 - Kruskal-Wallis test statistic = 25.96, p = 3.47e-7; *iv* complex trains = 1.8 s Q1= 1.3, Q3=2.5, synthetic twitters = 1.8 s Q1=1.2, Q3=2.8 - Kruskal-Wallis test statistic = 4.04, p = 0.044 – right panel Figure 3.5b).

Phase 1 aimed at assessing the discriminability of complex trains (Subset 1 with 108 stimuli) when similarities in the number of phrases, inter-phase interval, and first-phrase center frequency are shared with the synthetic twitters (see Methods – Experiment 4). Here, a significant drop in the hit rate was found when the center frequency of complex trains went above 4 kHz (median hit rate below 5 kHz = 0.84; median hit rate above 4kHz = 0.5; Kruskal-Wallis test statistic=53.84, p=2.1672e-13), reaching a plateau around 7kHz with higher frequencies producing hit rates around 0.5 (Figure 3.5c). Interestingly, the variations in the number of phrases and inter-phrase intervals did not show a significant impact on the performance of the animals (median hit rate NPH 8 = 0.72, NPH 9 = 0.68, NPH 10 = 0.79; Kruskal-Wallis test statistic = [0.005, 0.016, 0.531], p = [0.938, 0.897, 0.817], NPH 8-9, 9-10, 8-10 respectively; median hit rate IPI 117 = 0.72, IPI 128 = 0.74, IPI 139 = 0.71, Kruskal-Wallis test statistic = [0.237, 0.111, 0.552], p = [0.626, 0.738, 0.457], IPI 117-128, 117-139, 128-139 respectively). Fitting a psychometric curve to estimate the center frequency threshold at which marmosets stop grouping the target stimuli as complex trains, a threshold at 5.31 kHz was found (95% confidence intervals = 4.45 and 5.85) (Figure 3.5d).

In Phase 2, two independent subsets of complex trains (Subset 3 and 4) were tested independently but merged in the analysis since both shared similar modifications (amplitude-modulated variations) and had similar outcomes. Subset 2 held stimuli with amplitude-modulated contour but with the relative-phrase amplitude flat to its maximum (108 stimuli), while Subset 3 was composed of stimuli with amplitude-modulated contour plus relative-phrase amplitude modulation (108 stimuli). Here,a similar hit rate pattern as in Phase 1 was found across all the variations, suggesting that the implementation of amplitude modulation features to the complex trains did not affect the discriminability from synthetic twitters (Figure 3.5e) (median hit rate below 5 kHz = 0.83; median hit rate above 4kHz = 0.57; Kruskal-Wallis test statistic=40.68, p=1.1787e-10, median hit rate NPH 8 = 0.80, NPH 9 = 0.79, NPH 10 = 0.81; Kruskal-Wallis test statistic = [0.0008, 0.120, 0.163], p = [0.976, 0.729, 0.686], NPH 8-9, 9-10, 8-10 respectively; median hit rate IPI 117 = 0.81, IPI 128 = 0.76, IPI 139 = 0.80, Kruskal-Wallis test statistic = [0.300, 0.083, 0.653], p = [0.583, 0.772, 0.418], IPI 117-128, 117-139, 128-139 respectively).

For Phase 3, Subset 4 with 297 complex trains were used to test the influence of different bandwidths (from 0 to 6942 kHz). Based on the observed results from previous testing phases suggesting that the center frequency was the main contributor to disrupting the discrimination performance, only complex trains with center frequencies at 8.3, 9.5, and 10.3 kHz were used. Heatmaps in Figure 3.5f illustrate the variability hit rates across stimuli. Surprisingly, variations in bandwidth do not seem to contribute to the further disruption of the performance. Instead, the hit rate across bandwidth variations remained stable at chance level (median hit rate BW 2 kHz = 0.4, 3 kHz = 0.34, 4 kHz = 0.57, 5 kHz = 0.40, 6.94 kHz = 0.37; Kruskal-Wallis test statistic = [0.87, 0.83, 2.33, 1.19], p = [0.76, 0.083, 0.126, 0.257], BW 2-3, 3-4, 4-5, 5-6.94, respectively) which might suggest that frequency content regarding the bandwidth does not influence the discriminability between complex trains and synthetic twitters. Additionally, as in previous phases, the hit rate for target complex trains with lower center frequency remains above 0.75 (median hit rate below 5 kHz = 0.88; median hit rate above 4kHz = 0.5; Kruskal-Wallis test statistic=39.76, p=2.859e-10), and no variations in the hit rate were observed across variations in IPI and NHP (median hit rate NPH 8 = 0.84, NPH 9 = 0.85, NPH 10 = 0.85; Kruskal-Wallis test statistic = [0.2307, 0.1025, 0.0256], p = [0.6309, 0.7487, 0.8727], NPH 8-9, 9-10, 8-10 respectively; median hit rate IPI 117 = 0.81, IPI 128 = 0.76, IPI 139 = 0.80, Kruskal-Wallis test statistic = [0.300, 0.083, 0.653], p = [0.583, 0.772, 0.418], IPI 117-128, 117-139, 128-139 respectively). Hit rates for anchor virtual twitters can be found in supplementary Figure S3.1.

Finally, preliminary data (Phase 4) suggests that hit rates might decrease when the first harmonic is added to the complex trains. Figure 3.5g shows the hit rate across the different stimuli of Subset 5, showing a median of 0.83 across all the tested frequency bandwidths (6.942 kHz, 5 kHz, and 4 kHz). No statistical test was performed due to the low number of trials.


a)
Target Complex Trains
Subset 1
Subset 2
Subset 3
Subset 4
Subset 5
Amplitud
Frequency
Temporal Variations
Spectral variations
Anchor Complex Train
Amp.
Freq.
20
kHz
0
Time
Anchor Synthetic Twitter
Amp.
Freq.
20
kHz
Time
0
b)
Trials
500
0
239
Hit Rate
1.0
0.5
0.84
Sessions
100
0
107
dl
el
iv
Mean
monkey
n=15828
n=13018
n=12262
Hit Rate
0.5
0.0
Response Latency [s]
5
0
*
*
*
cTrain
sTwitt
dl
el
iv
c)
Center Frenquency [kHz]
2.0
4.0
6.0
8.7
10.3
15.0
20.0
8
9
10
Number of Phrases
117
128
139
Inter-Phrase Interval [ms]
Number of Phrases
8
9
10
117
128
139
Inter-Phrase Interval [ms]
Hit Rate
1.0
0.8
0.6
0.4
0.2
0.0
d)
Proportion Correct
1.0
0.8
0.6
0.4
0.2
0.0
THR = 5.31
0.0
5.0
10.0
15.0
20.0
Middle Center Frequency [kHz]
Object
cTrain
sTwitt
Trials
1000
4000
e)
Center Frenquency [kHz]
2.0
4.0
6.0
8.7
10.3
20.0
8
9
10
Number of Phrases
117
128
139
Inter-Phrase Interval [ms]
Number of Phrases
8
9
10
117
128
139
Inter-Phrase Interval [ms]
Hit Rate
1.0
0.8
0.6
0.4
0.2
0.0

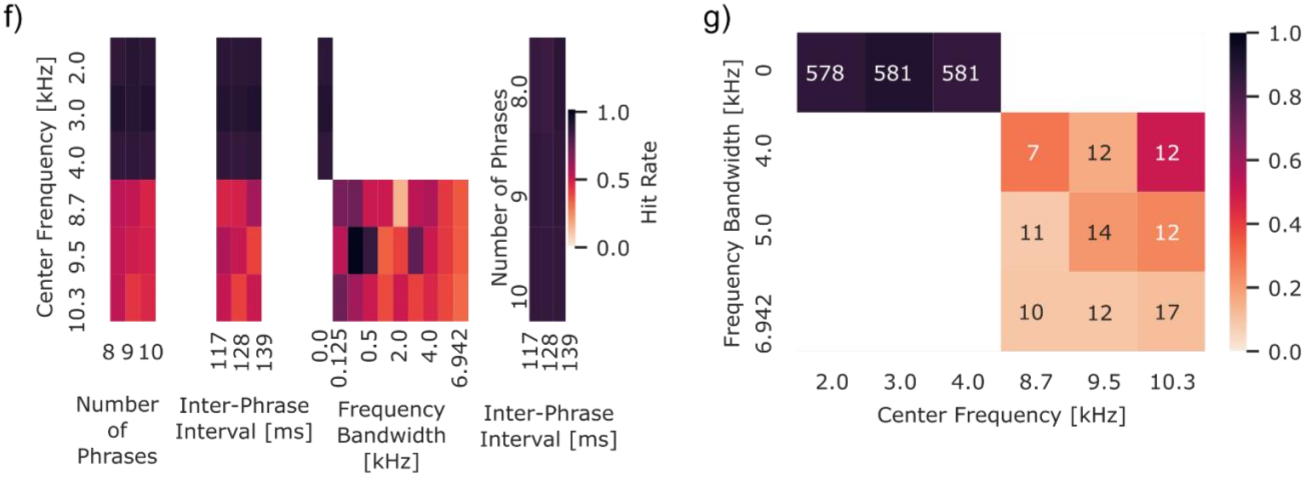


Figure 3.5. Vocal perception. a) *Graphical representation of the 5 subsets of target complex trains. The right panels show the visual stimuli paired to a single anchor virtual twitter and complex train.* ***b)*** *Left panel - letter-value plots the number of trials and hit rates across the total number of sessions per animal. The right grey panels show the distribution of 50 sessions randomly picked per animal. The central box represents the distribution's first, second, and third quartiles. Trials per session Q1=177, Q2=237, Q3=321. Hit rate Q1=0.79, Q2=0.88, Q3=0.86. The right panels show the hit rate and response latencies as a function of the stimulus group within 50 randomly chosen sessions. The central box represents the distribution's first, second, and third quartiles for the letter -value plots. Stars in the letter plots indicate statistical significance (dl complex trains = 1.2 7 s Q1=1, Q3=1.7, synthetic twitter = 1.5 s Q1=1.1, Q3=2 - Kruskal-Wallis test statistic = 42.02, p = 8.99e -11; el complex trains = 1.7 s Q1=1.3, Q3=2.5, synthetic twitters = 2 s Q1=1.5, Q3=2.9 - Kruskal-Wallis test statistic = 25.96, p = 3.47e -7; iv complex trains = 1.8 s Q1= 1.3, Q3=2.5, synthetic twitters = 1.8 s Q1=1.2, Q3=2.8 - Kruskal-Wallis test statistic = 4.04, p = 0.044).* ***c)*** *Heatmaps showing the hit rate across stimuli variations in Phase 1. From left to right: center frequency as a function of the number of phrases, the inter-phrase interval, and the number of phrases as a function of the inter-phrase interval.* ***d)*** *Psychometric estimation of discrimination threshold (black line vertical) based on the proportion of times the animals selected the correct response across variations in center frequency. 95% confidence intervals are indicated with a horizontal line. (Threshold = 5.31 kHz, confidence intervals = 4.45 and 5.85). Synthetic twitters hit rates (orange dots) do not form part of the threshold estimation.* ***e)*** *Heatmaps showing the hit rate across stimuli variations in Phase 2. From left to right: center frequency as a function of the number of phrases, the inter-phrase interval, and the number of phrases as a function of the inter-phrase interval.* ***f)*** *Heatmaps showing the hit rate across stimuli variations in Phase 3. From left to right: center frequency as a function of the number of phrases, the inter-phrase interval, the bandwidth, and the number of phrases as a function of the inter -phrase interval.* ***g)*** *Heatmaps showing the hit rate across stimuli variations in Phase 4. From left to right: center frequency as a function of the number of phrases, the inter-phrase interval, the bandwidth, and the number of phrases as a function of the inter-phrase interval. Numbers inside the boxes indicate the number of collected trials.*

Stimulus	CF [kHz]	NPH	IPI [ms]	BW [kHz]	NHR	RPA	AMC
ST	8.7	8	117	6.942	1	True	True
ST	8.7	8	128	6.942	1	True	True
ST	8.7	8	139	6.942	1	True	True
ST	8.7	9	117	6.942	1	True	True
ST	8.7	9	128	6.942	1	True	True
ST	8.7	9	139	6.942	1	True	True
ST	8.7	10	117	6.942	1	True	True
ST	8.7	10	128	6.942	1	True	True
ST	8.7	10	139	6.942	1	True	True
ST	9.5	8	117	6.942	1	True	True
ST	9.5	8	128	6.942	1	True	True
ST	9.5	8	139	6.942	1	True	True
ST	9.5	9	117	6.942	1	True	True
ST	9.5	9	128	6.942	1	True	True
ST	9.5	9	139	6.942	1	True	True
ST	9.5	10	117	6.942	1	True	True
ST	9.5	10	128	6.942	1	True	True
ST	9.5	10	139	6.942	1	True	True
ST	10.3	8	117	6.942	1	True	True
ST	10.3	8	128	6.942	1	True	True

ST	10.3	8	139	6.942	1	True	True
ST	10.3	9	117	6.942	1	True	True
ST	10.3	9	128	6.942	1	True	True
ST	10.3	9	139	6.942	1	True	True
ST	10.3	10	117	6.942	1	True	True
ST	10.3	10	128	6.942	1	True	True
ST	10.3	10	139	6.942	1	True	True

Table 3.2 Espectro temporal characteristics of the 27 synthetic twitters. *Number of phrases (NPH), middle center frequency (CF), inter-phrase interval (IPI), middle frequency bandwidth (BW), number of harmonics (NHR), relative-phrase amplitude (RPA), and amplitude modulated contour (AMC).*

3.3 Discussion

In Chapter 3, I report the results from a series of 4 experiments conducted with a stand-alone system described in Chapter 2, the MXBI. Animals involved in these experiments actively interacted with our device throughout sessions without controlling food or water intake. In Experiment 1, 11 marmosets navigated an automated, unsupervised procedure to learn a pure tone detection task that was later used in Experiment 2 to collect audiograms for 5 animals. Experiment 3 demonstrated the feasibility of wireless control of cochlear implants in freely moving marmosets and showed that unilaterally cochlear implanted marmosets flexibly learned to use electric stimulation to solve an auditory detection task. Electric hearing thresholds are still under data collection for 2 animals. Finally, in Experiment 4, 3 marmosets underwent a vocal discrimination task where it was found that variations in the center frequency and harmonic content of synthetic stimuli (complex trains) contribute the most for vocal recognition. These results ultimately confirm the conclusions reported already in Chapter 2, regarding the stable engagement of marmosets across sessions and additionally demonstrate that our novel automated training and testing protocols can achieve sophisticated psychophysics measurements, assess the efficacy of cochlear implants wirelessly, and answer critical questions on auditory cognition.

Manual training variability

For all of our experiments, marmosets were required to interact with a touchscreen to receive liquid reward from a mouthpiece. In Chapter 3, I described a procedure to introduce common marmosets to touchscreen interaction (see Methods – Initial training). Although a similar protocol in Chapter 2 was successfully tested (training 14 out of 14 complete naive marmosets to interact with a touchscreen), our approach in Chapter 3 successfully trained only 5 out of 10 naive marmosets. Several reasons could explain this discrepancy with our previous reports. First, all of the 5 animals that showed low interest in our device also showed little interest in humans during manual feeding, as opposed to the animals from the study described in chapter 2. It has been reported, for instance, that the marmosets' inter-individual differences in behavior, such as exploration and boldness, might also account for consistent variations in learning speed and cognitive performance (Bugnyar and Huber, 1997; Burkart et al., 2009; Day et al., 2003; Šlipogor et al., 2022; Voelkl and Huber, 2000). Second, all naive animals of Experiment 1 described in this chapter were manually trained by a new experimenter from experiments in Chapter 2, which might have added an uncontrolled source of variations in training and shaping strategies across animals. It has been shown that variations in the experimenter's expertise and identity (sex, familiarity with the animals, odor, and individual training strategies) influence animals' behavior during training (Bohlen et al., 2014; Cibulski et al., 2014; Rivalan et al., 2017; Schmitt et al., 2014; Schubiger et al., 2015; Sorge et al., 2014a, 2014b). Coincidentally, automating the initial training as well, similar to all other experiments described here, could prevent these issues.

Automated audiograms directly in the home-cage

Audiograms are mainly collected within a sound-controlled chamber (Osmanski and Wang, 2011; Pfingst et al., 1978; Stebbins, 1973), where specific frequency bands are attenuated to avoid any disturbance of the subject's performance during data collection. One big challenge of collecting hearing thresholds directly in the home cages is represented by the acoustic interferences of a marmoset facility's uncontrolled and highly variable acoustic environment. While collected in a marmoset colony, our audiograms show, on average, the typical w-shape reported for New-world monkeys with the highest sensitivity peaks at 2 kHz and 7kHz and a region of reduced sensitivity between these two peaks. However, we indeed observe a decrease in sensitivity across all tested frequencies compared with the average audiogram collected in a sound-attenuated chamber (Osmanski and Wang, 2011). Three factors might have contributed to this shift. First, sounds in the background during data collection could have masked pure tones of ongoing trials and negatively impacted the animals' performance. Masking is a well-understood phenomenon that could affect and compromise the perception of one sound in the presence of another (Egan and Hake, 1950; Moore, 2013). Several studies have shown how noise affects frequency sensitivity (Rocchi et al., 2017), making, consequently, sound perception more challenging. Second, more biologically salient sounds (conspecific vocalizations) arising from the background could have potentially switched the attention of the animals in ongoing trials. Auditory attention has been defined as the cognitive process that allows the listener to focus selectively on the stimulus of interest while ignoring irrelevant competing stimuli (Choudhury and Gorman, 2000; Fritz et al., 2007; Gomes, 2000). Attention can be voluntarily directed (top-down) or captured by salient stimuli (bottom-up) (Ahveninen et al., 2006; Johnson and Zatorre, 2006; Knudsen, 2007). Bottom-up attention reflects the effect of silience filters at the level of the central nervous system that might select for stimuli of instinctive (looming stimuli) or learned (conspecific vocalizations) biological importance. These stimuli are perceived as "popping out" from the scene, and in some cases, their relevance is so strong that they might interrupt other kinds of ongoing information processing (Baddeley, 2003; Bisley and Goldberg, 2003; Egeth and Yantis, 1997; Knudsen, 2007; Miller and D'Esposito, 2005; Remington et al., 1992). Anecdotally, marmosets were observed several times vocalizing inside the MXBIs, as a possible response to a conspecific vocalization coming from another animal in the colony. Third, contrary to the chair-based setup from Osmanski et al. (Osmanski and Wang, 2011), where marmosets were kept with their heads in a fixed position in space relative to the loudspeaker, our devices allowed the animals to move during the delivery of the sound, which might have modulated the sound intensity itself.

Another notable difference with previous reports (Osmanski and Wang, 2011) is the proficiency of the animals. In our experiments, animals were better at operating the acoustic discriminations across all frequencies and tasks (*% improvement compared to other studies*). While the reported average d-Prime for the reported session was 1.68 +/- 0.18, our discriminability indices were all, in general, above 2.0 (Figure 4.3d). Additionally, our animals showed higher discriminability indices for those frequencies with higher sensitivity. This greater performance was kept across all sessions performed, as depicted in Figure 3.3c (hit rates and false alarm rates across sessions). Previous reports showed greater performance (faster reaction times) when acoustic and visual stimulation are paired compared to those where visual or acoustic were presented individually (Cappe et al., 2010; Miller et al., 2001). Multisensory redundancy increase stimulus saliency, which might partially explain this improvement in performance.

Finally, although the response latencies of this work showed a similar tendency as those reported by Osmanski and Wang (Osmanski and Wang, 2011), with a gradual decrease in the response latencies, as the acoustic stimuli became less intense, no significant difference in response latencies was found between the highest and lowest intensities as opposed to reports by Osmanski and Wang (Osmanski and Wang, 2011) (Figure 3.3d). Differences in task design might account for these differences. Motor responses and task structure While the animals in Osmanski et al. (Osmanski and Wang, 2011) had to lean forward a tube that was placed in front of the animals' mouths, our animals had to reach the screen through a mesh aperture using one of the two hands which might have added additional variability in the reaching behavior. As an anecdotal observation, in some trials, animals required more than one attempt to reach the target on the screen before the response was registered.

Wireless cochlear implant control and assessment in freely moving marmosets

Experiment 3 reports the results from 1 unilateral cochlear implanted marmoset (animal *in*) and preliminary results from a second animal (*wl*). Although the electrical hearing thresholds were collected in a single animal (*in*), testing the wireless communication with the sound processor and the control over the cochlear implant was done in both animals. The ability to wirelessly control cochlear implants in freely moving marmosets represents a noteworthy advancement in the development of automated protocols for phenotyping and characterizing various cognitive mechanisms in cochlear implant research. This innovation is particularly aimed at evaluating the performance of a new generation of cochlear implants that rely on optogenetic stimulation. These implants, developed within our laboratory, allow for the assessment of their functionality in a behavioral context, marking a significant stride in understanding and refining cochlear implant technologies (Dieter et al., 2019; Dombrowski et al., 2019; Hernandez et al., 2014; Jablonski et al., 2020; Jeschke and Moser, 2015; Keppeler et al., 2018; Moser, 2015; Wrobel et al., 2018). Additionally, animals *in* and *wl* smoothly transitioned between the two types of stimulation, suggesting an equivalent level of comfort of these animals in using the electric stimulation

Data for animal *in* completely stopped after all electrodes (first electrode 4, then 5, and then all) showed unexpectedly high impedance values (Figure 3.4e). Current efforts are trying to unfold the reasons behind this circuitry anomaly. A similar phenomenon happened with a third animal (*cl*) which could not perform Experiment 3. In this case, animal *cl* showed connectivity issues in the cochlear implant electrodes at the beginning of the testing procedure. A further evaluation led to the conclusion that the reference electrodes did not have contact anymore with the tissue (due to a possible break of the line), making the electrical stimulation impossible due to the open circuit circumstances. Although, changes in impedance values over time are a well-documented phenomenon in cochlear-implanted human patients (due to fibrous growth around the electrodes and changes in the surface chemistry of the electrodes) (Fayed et al., 2020; Lin et al., 2019; Neuburger et al., 2009; Wei et al., 2021; Wilk et al., 2016). Our experimental design requires a substantial improvement to warranty the stable connectivity of the cochlear implant over time to reliably interpret further data.

Spectro-temporal features for vocal perception

In Experiment 4, 3 common marmosets were trained to perform a vocal discrimination task using a 2-alternative choice paradigm. Here, marmosets were presented with a total of 945 synthetic stimuli with variations in the spectro-temporal domain across a series of 4 testing phases in which they were required to classify the given stimuli in one of two categories: twitter-like or non-twitter-like sounds. The aim of Experiment 4 was to investigate the effects of vocal (twitter call) discriminability of 3 marmosets by assessing their behavioral responses to twitter-like stimuli with diverse spectro-temporal variations. The use of the twitter call as the testing vocalization was encouraged by the accessibility to marmosets with experience in twitter discrimination (Chapter 2 – Generalization to novel stimuli), the availability of a quantitative description of the spectro-temporal features of twitters (Agamaite et al., 2015), and the capability of a software to freely manipulate a vast number of acoustic features independently to synthesize twitters (DiMattina and Wang, 2006).

Across our experiments, the response patterns observed suggest that the combination of spectral characteristics such as center frequency and harmonic content have a greater impact on the twitter discrimination performance than temporal variations like the number of phrases, inter-phrase intervals, or amplitude modulations (Figure 3.5d,e,f,g).

Several groups have investigated vocal perception in NHPs using a wide variety of sounds and methods. In cognitive neuroscience, there seems to be an agreement that suggests that NHPs possess a specific cortical area responsible for integrating conspecific vocalizations (Belin, 2006; DiMattina and Wang, 2006; Newman and Wollberg, 1973; Ortiz-Rios et al., 2015; Petkov et al., 2008; Rauschecker et al., 1995; Romanski et al., 2005). However, how different acoustic parameters contribute to the recognition of sounds is not a fully understood process. Neurophysiology and behavioral literature point to periodicity and temporal envelope as possible cues for vocal recognition (Brewer and Barton, 2016; Chandrasekaran et al., 2011; Mesgarani et al., 2014; Stevens, 1983) in combination with spectral features such as spectral envelope and harmonic content (Ackermann et al., 2014; Fischer, 1998; Furuyama et al., 2017; Ghazanfar and Rendall, 2008; Kaplan et al., 1978; Lieberman and Blumstein, 1988; May et al., 1989; Owren et al., 1997; Remez et al., 1981; Rendall et al., 2004). A caveat when comparing these findings is the various manipulation of the acoustic stimuli used for testing. I argue that in order to ultimately demonstrate which and how specific features might be behaviorally relevant for NHPs (specifically common marmosets) for call identification and discrimination, a scrutinous manipulation of individual acoustic elements in the stimuli, as we did in Experiment 4, must be an essential requirement.

Exploring the spectro-temporal features that account for making a sound a vocalization is an arduous task that requires testing and controlling for a significant amount of variabilities within the vocalizations. As for our results, further investigations must unwind to what extent and how the first harmonic and its variations interact with the rest acoustical features to make marmosets categorize a sound as a twitter-like sound.

Finally, an important aspect to highlight in our findings is that the 3 marmosets reported in Experiment 4 showed a discrimination behavior that comes from training. Categorical discrimination is a flexible and context-dependent process, and it is influenced by the experience with the stimuli

(Repp et al., 1984). As such, it remains unclear whether the marmosets use the reported discrimination features in their natural conditions or whether these discrimination patterns are just the result of training. This is paramount to take into consideration when future plans aim to compare the performance of the novel optogenetic cochlear implant over the electrical cochlear implant behaviorally in an ecological validity context.

Conclusion

Overall, the results described in this chapter aligned with the findings in Chapter 2, where I report that marmosets steadily engage with the MXBI to solve auditory tasks voluntarily directly in their home cages and further expand by demonstrating the possibility of collecting precise, pure tone audiograms, assess the efficacy of wirelessly-controlled cochlear implants in freely moving animals and address essential questions on auditory cognition. Ultimately proving the high versatility and robustness of our system, the MXBI, to optimize experimental protocols through automation in primate auditory cognition.

3.4 Methods

Animal welfare statement

All experiments presented in this chapter were approved by the responsible authorities from the regional government office [Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES), Permit No. 14/1572 and 18/2976] as well as the ethics committee of the German Primate Center (Permit No. E1-20_4_18) and were in accordance with all applicable German and European regulations on husbandry procedures and conditions. It has to be noted, however, that—according to European regulations and implemented in German animal protection law — the practices described in Experiments 1, 2, and 4 can be considered environmental enrichment.

Animals

Seventeen common marmosets were tested across all experiments, see Table 3.1. in the facilities of the German Primate Center. Similar conditions as for the animals described in Chapter 2 were kept. With the exception of animal pairs *wl-dc, iv-dr,* and *dn-bs,* which were housed in double-size cages, animal pairs were housed in wire mesh cages of 160 cm (H) x 65 cm (W) x 80 cm (D) with a daylight cycle of 12 hours (06:00 to 18:00). Opaque plastic dividers between neighboring pairs and cloths hung from the ceiling across the room prevent visual contact. Animals *bs*, *wl*, *in*, and *cl* were prepared for neurophysiological and cochlear implant experiments. Experimental sessions occurred either in the morning (09.00 to 13:00) or in the afternoon (13:00 to 17:00). Food was provided daily at noon, and water was available ad libitum. Liquid arabic gum (Gummi Arabic Powder E414, Willy Becker GmbH) was dissolved in water (1:5) and delivered as a reward for correct trials across all experiments. During the initial phase of the training, pieces of marshmallows were stuck to the screen to encourage the animals to interact with the touchscreen.

Apparatus

The data was collected using 7 MXBIs. A detailed description of the devices can be found in the Methods – Apparatus section of Chapter 2, even though the reproducibility of the MXBI is intended to produce identical devices. Minor variations among the 7 MXBI could be found in the mesh position that divides the animals from the touchscreen (mean difference between the 7 MXBIs 0.7 cm, std +/- 0.2), the diameter of the RFID antenna (10 cm and 8 cm), and the mouthpiece's material (stainless steel and plastic). These variations across MXBIs were the product of updating the new models to best-fitting components. Although, these variations might account for slight differences in the accessibility or comfortableness of interaction with our device. I would argue that these differences would be of minimum concern regarding the performance of the animals since no significant behavior change was observed between these variations.

Sessions

Experiments 1 and 4, took place in the mornings (from 09:00 to 13:00), and afternoons (from 13:00 to 17:00), or across the whole day (from 09:00 to

17:00). However, Experiments 2 and 3 took place just in the afternoons. For Experiment 3, animals *in* and *wl* were manually handled at the beginning and end of the sessions to place and remove the mobile sound processor. All sessions were video monitored by the experimenter, and in some cases, recordings of the sessions were saved. Animals never leave their home cage, and separation between cagemates happened on rare occasions during the initial training phase, where the animals were taught to interact with the screen. Animals *iv* had a cage mate that did not participate in any of the experiments but had access to the MXBI. No task was offered to this animal. Except for animals *el*, *dl*, *wl*, *dc*, *in*, *cl*, and *iv* that had previous experience in touchscreen tasks, all remaining animals underwent a manual training procedure (see initial training of Experiment 1) to learn the principles of touchscreen interaction. An overall description of the different experiments that every animal underwent can be found in Table 3.1. It is essential to highlight that the order in which the experiment numbers appear in the table is the actual succession in training.

Experiment 1 – Sound detection

Experiment 1 comprised manual and automated training phases. The manual training phase aimed to introduce naive marmosets into touchscreen interaction using positive reinforcement techniques. In contrast, the automated training phase aimed to instruct marmosets in a pure tone detection task using a Go-NoGo paradigm.

Only 10 (*dn*, *bs*, *pr*, *pt*, *dg*, *dm*, *bl*, *oh*, *il* and *bc*) out of the 16 marmosets that underwent Experiment 1 did not have experience with touchscreen tasks and therefore went through the initial training phase.

Initial training

The inial training was divided into four stages that, in sequence, aimed to instruct marmosets to touch the screen in exchange for a reward.

Habituation to the device – Animals had access to a single MXBI across several sessions until individual marmosets were observed to remain in the MXBI for at least 30 seconds. During these sessions, the screen remained turned off, and animals were monitored via video surveillance to assess their behavior.

Mouthpiece-reward association – Reward drops were remotely triggered (by the experimenter) to attract the animals' attention toward the mouthpiece. Simultaneously, a pure tone train was played (2 kHz, 0.1 s long with 0.1 s inter-tone interval). This stage was accomplished when individual animals were observed approaching the mouthpiece without a reward being delivered (presumably seeking for reward).

Sound-reward conditioning – Across a series of sessions, the reward was systematically delayed from the sound onset, going from 0 to 1 s. This phase was considered completed when individual animals directed their gaze or approached the mouthpiece right after sound onset but before reward delivery in 5 consecutive trials.

Touch-to-reward – Pieces of marshmallows were stuck on the touchscreen to stimulate hand-reaching toward the touchscreen. A touch on the screen started the pure tone train, which triggered the reward upon ending. Once individual animals were observed touching the screen in the absence of marshmallows to initiate the sound across 10 consecutive trials, the initial training was considered over.

Pure tones used in Experiments 1 and 2 were generated and stored offline using a custom Python 3 script, using Scipy, Numpy, and Pylab libraries. Pure tones were 100 ms long with a ramp (in and out) duration of 5 ms.

Automated unsupervised training (AUT)

After the completion of the initial training, animals went through a series of 38 preprogrammed steps that aimed to train marmosets in a pure tone detection task. The progress of the animals through the AUT was constantly monitored by the algorithm described in Chapter 1 (Figure S4e). Shortly, the algorithm evaluated the performance of individual animals in a sliding window of 20 trials. A hit rate equal to or above 80% made individual animals ascend the staircase, while a hit rate equal to or below 45% made an animal descend. The progress of individual animals was stored and retrieved within and across sessions, providing animals the possibility to navigate through the

AUT at their own pace. The AUT protocol comprised 38 steps grouped into three milestones (Figure 3.1a).

Touch to sound – Comprised steps 1 to 15 aimed to train naïve animals into precise touch behavior while also conditioning them to expect the delivery of reward after the presentation of pure tone trains of variable duration. A geometric figure (Four circles of different sizes and tonality of grey embedded in a grey square) on the screen (named trigger) served as a visual target to start a pure tone train upon touch (2 kHz, 0.1 s long with 0.1 s inter-tone interval). 0.1 ml of arabic gum was delivered through the mouthpiece as a reward at the sound end. From step 1 to step 15, the size of the trigger shrank from 6 x 6 cm to 3 x 3 cm, while the pure tone train duration range increased from 1 to 1 s to 1 to 2.5 s. Animals were required to touch the trigger and wait for the pure tone train to end to obtain the reward (correct trial). Additional touches during the sound presentation stopped the sound and aborted the trial (wrong trial). A timeout of 5 to 8 s was given, where the screen turned grey, and no interaction with it was possible. *Since this trial structure required the animals to touch a single time on the screen to start the sound, I referred to these trials as single-touch trials.*

Sound-lengthening – Comprised steps 16 to 25 aimed to strengthen the sound-ending-reward association by further increasing the variability in sound duration from 1 to 2.5 s to 2.5 to 6.5 s.

Double-touch to sound – Comprised steps 25 to 38 aimed to train animals to react to the absence of sound upon the first touch by touching the target on the screen a second time. Trials that required two touches (double-touch trials) to trigger the sound represented 20% at step 16 and gradually increased to 50% at step 38. In double-touch trials, animals were given the maximum sound duration (5.5 s) plus 3 s to react to the absence of sound by touching the screen a second time, triggering the sound. A second touch within the first 6 s was counted as correct (hit), whereas a second touch after 6 s was counted as wrong (miss). The absence of a second touch within the time given resulted in an ignored trial. Single-touch trials were treated as already explained. The withholding of additional touches during sound presentation was counted as a correct rejection, while a touch during the sound presentation was counted as a false alarm. Once step 38 was reached and a hit rate above 75% and a false alarm rate below 25% were kept in 4 consecutive sessions, Experiment 1 was considered completed (Figure 3.2a)

Experiment 2 - Hearing thresholds

Five marmosets from Experiment 1 (*in*, *cl*, *wl*, *dc*, and *bl*) underwent the collection of hearing thresholds. For these sessions, pure tone stimuli ranging from 0.2 kHz to 20 kHz (0.2 kHz, 1 kHz, 2 kHz, 5 kHz, 7kHz, 10kHz, and 20 kHz) were generated and stored offline, as described in Experiment 1 – Initial training. One frequency was tested at a time, and while all animals started with 2 kHz, the remaining frequency order was randomized. Three MXBIs were used to collect the data, assigning each to a pair of marmosets (*dc*-*wl*, *it*-*cl*, and *bl*-*oh*). It is essential to highlight that although animal *oh* was not tested for Experiment 2, it still had access to the device and was offered a different task. Sound intensity calibration for the 3 MXBIs occurred three times across all experimental sessions, one before and two during data collection. Chapter 2 details the calibration process (Methods, Psychoacoustic assessment). Experimental sessions took place in the afternoon (from 13:00 to 17:00) when the colony's background noise was the lowest, with feeding and personnel activity occurring mainly in the morning. The Animals' cages were located in different rooms of the animal facility to avoid overlapping stimuli while performing the task.

Sensitivity thresholds were tested using step 38 from the AUT procedure explained in Experiment 1 (Figure 3.3.a). Frequencies from 0.2 kHz to 20 kHz were tested one at a time(0.2 kHz, 1 kHz, 2 kHz, 5 kHz, 7 kHz, 10 kHz, and 20 kHz). Within experimental sessions, marmosets were randomly presented with target trials (with intensities of 10 dB SPL, 20 dB SPL, 30 dB SPL, 40 dB SPL, 50dB SPL, and 60 dB SPL) and anchor trials (with intensities of 70 dB SPL and 80 dB SPL). Target trials were held only in single-touch trials and represented 20% of the total.

A minimum of 35 trials per intensity was required to estimate the hearing threshold of individual frequencies. Psychometric estimation of hearing thresholds was based on the times the animals selected the correct response across the intensities. Once all frequency thresholds were collected, individual animal audiograms were built by collapsing the hearing thresholds for unique frequencies.

Lastly, for animal *in*, sessions of Experiment 3 (electrical stimulation through the cochlear implant) were interleaved with Experiment 2. See Methods – Experiment 3.

Experiment 3 - Electric stimulation thresholds in cochlear implanted marmosets

Cochlear implant

Animals *bs*, *cl*, *in*, and *wl* were chronically, unilaterally implanted with cochlear implants (MEDEL Marmoset-Electrode, 10 channels in 10mm 0.3mm Array, 14 Pin connectors, (2 wires as a reference electrode were inserted underneath the musculus temporalis), in November 2018, November 2020, June 2020, and June 2022, respectively) into the scala tympani via a cochleostomy of the basal turn. In addition to the implantation, animals were single-sided deafened with 0.2 ml neomycin injection into the cochlea. Animals *cl, in,* and *wl* received a virus injection (An AAV based viral vector including a hSyn promotor and a ChR2 variant with a EYFP tag) for parallel experimental reasons. All three procedures were performed in the same day in the left ear under anesthesia.

Only animals *in* and *wl* underwent Experiment 3. Animal *bs* did not learn the task of Experiment 1, and the cochlear implant of animal *cl* was found out of use by the time Experiment 3 took place.

Wireless control and communication protocol

To achieve wireless control of the cochlear implant. We used the communication protocol described by Jablonski, Harczos, et al. (Jablonski et al., 2020) for freely moving rats. The two marmosets were equipped with a head enclosure (placed on the top of the head) to host during experimental sessions: 1) a male pin (6 x 2) connector from the cochlear implant; 2) A custom-made printed circuit board (PCB) containing the sound processor (built upon a digital signal controller – nRF52832, Nordic Semiconductor) and an antenna (FXP23 BlueDiaminf 2.4 GHz Band Antenna) wrapped around the board; and, 3) A powering PCB containing a rechargeable battery (3.7 V Lithium-ion battery CP1654A3, VARTA Microbattery GmbH). The sound processor used a communication protocol based on Enhanced ShockBurst (ESB) over its 2.4 GHz radio transceiver to connect with a custom-coded firmware ESB-USB bridge in the Raspberry Pi (Figure 3.4a).

Experimental sessions

All sessions occurred in the afternoon (from 13:00 to 17:00 h). The sound processor and the battery board were placed and removed from the animal's head enclosure before and after every session. Animal *in* and *wl* underwent a series of testing sessions with the aim of testing 1) the communication stability between the mobile sound processor and 2) the control of the cochlear implant. For animal *in* these sessions were interleaved with sessions of Experiment 2. Therefore, the animal *in* was presented with different types of stimulation (acoustic and electric) in a successive manner for several weeks while collecting data for Experiment 2. For animal *wl* al testing sessions with electric stimulation took place after Experiment 2. Once stable control was achieved for 3 consecutive days (electrical stimulation and wireless control of the sound processor), only sessions of Experiment 3 took place.

Experiment 2's Go-NoGo task was adapted to deliver electric (Monopolar electrical stimulation (100 repetitions of 100 µs stimulus duration (50 µs positive and 50 µs negative phase) with an inter-stimulus interval of 1000 µs – charge-balanced biphasic cathodic-first pulses) instead of acoustic stimulation. Target trials held stimulation intensities of 10 µA, 30 µA, 50 µA, and 70 µA, while anchor trials held 90µA and 110 µA. The amplitude ranges were based on previously collected data for animal *in* (using a chair base training procedure). Simulation intensities and channel order were randomized for each animal. A single channel was tested at a given time. The transition from acoustic to electric stimulation happened within a single session and without the use of target trials. Animals started the session with acoustic stimulation, and after 100 trials, the stimulation type was switched and kept to electric. Once the animal showed a stable performance (>= 75% hit rate and <= 25% false alarm rate) using anchor trials in electrical stimulation, the collection of the electrical stimulation thresholds took place.

A minimum of 35 trials per intensity were required to estimate the electrical stimulation thresholds. Similarly, as in Experiment 2, psychometric estimation of the stimulation thresholds was based on the times the animals selected the correct response across the stimulation intensities.

Impedance values measure the resistance to the current flow and give essential information about the electrical state of the electrode, which might change over time due to changes in tissues and liquids of the inner ear around the electrode. Across all experimental sessions, each electrode's impedance was measured at the beginning of every session to evaluate the electrodes' connectivity status. Testing sessions were aborted when an unexpected increase in impedance values was observed.

Experiment 4 – Vocal perception

Experiment 4 aimed to explore the effects of vocal perception in marmosets across specific spectro-temporal variations. Thereby, 945 different synthetic stimuli were created relying on a Python adaptation from the software reported by DiMattina & Wang (DiMattina and Wang, 2006), using the population statistics variability values reported by Agamaite and colleagues (Agamaite et al., 2015) as reference. All stimuli had variations in the number of phrases, center frequency, inter-phrase interval, middle-frequency bandwidth, number of harmonics, relative phrase amplitude, and amplitude-modulated contour (Table 3.2 and Ttable S3.1).

First, marmosets *dl*, *el*, and *iv* were trained to generalize 27 synthetic twitters (later named anchor twitters) and discriminate them from 27 pure tone trains (later named anchor complex trains) using the same visuo-acoustic discrimination paradigm reported in Chapter 2 (Results - Generalization to novel stimuli). Once animals showed a stable discriminability performance (hit rate above 75%) during three consecutive sessions, a series of sequential testing phases took place.

Each testing phase held a unique set of synthetic target stimuli (referred to as complex trains) with particular spectro-temporal variations (e.g., center frequency, number of harmonics, bandwidth, number of phrases, and inter-phrase interval), which were randomly interleaved between the anchor stimuli. Similarly, as in the 2AC audio-visual association from Chapter 2, animals were required to touch the image of an adult marmoset on the screen upon the presentation of a synthetic twitter and touch a geometric pattern (three grey triangles embedded in a grey square) when an anchor complex train was played. A correct response led to the delivery of 0.1 ml liquid arabic gum, while a wrong response led to a timeout of 5 to 8 s, where a grey screen was presented, and no interaction with the touchscreen was possible. A trial with no screen interaction during the sound presentation was labeled as ignored, and a new trial started after an inter-trial interval of 2 - 3 s. Trials holding target stimuli represented 10% of the total and were rewarded 50% regardless of the animals' response.

Phase 1 used 108 target complex train stimuli (Subset 1) with variations in the number of phrases (9, 10, and 10), inter-phrase interval (117ms, 128ms, and 139 ms), and center frequency (2 kHz, 3 kHz, 4 kHz, 5 kHz, 6 kHz, 7 kHz, 8.7 kHz, 9.5 kHz, 10.3 kHz, 13 kHz, 15 kHz, 17 kHz, and 20 kHz).

Phase 2 used two independent subsets of complex trains (Subset 3 and 4). Although they were tested independently, they were grouped into the same testing phase due to their outcome similarities. Both subsets implemented amplitude-modulated variations. Subset 2 added to the stimuli of Subset 1 amplitude-modulated contour but kept the relative-phrase amplitude flat to its maximum. Subset 3 stimuli added relative-phrase amplitude modulation to stimuli of Subset 2.

Phase 3 used 297 stimuli (Subset 4) with variations in the number of phrases (9, 10, and 10), inter-phrase interval (117ms, 128ms, and 139 ms), center frequency (8.7 kHz, 9.5 kHz, 10.3 kHz), and frequency bandwidth (0 kHz, 0.03 kHz, 0.065 kHz, 0.125 kHz, 0.25 kHz, 0.5 kHz, 1 kHz, 2 kHz, 3 kHz, 4 kHz, 5 kHz, and 6.942 kHz).

Finally, Phase 4, still under data collection, used 297 stimuli from Subset 4 with the first harmonic added (Subset 5).

3.5 Data treatment and Statistics

Data acquisition, formatting, analysis and plotting were performed using Python 3.5.3 and 3.7.7 using NumPy (Oliphant, 2006), SciPy (Millman and Aivazis, 2011; Oliphant, 2007), Pandas (McKinney, 2010), and Seaborn. The threshold for statistical significance was set to 0.05. Reaction time differences between the two stimuli groups in Experiment 4 (Figure 3.5b) and hit rates between groups of complex trains across the testing sessions in Experiment 4 were tested for significance with a Kruskal-Wallis test.

Psychometric functions for Experiments 2 and 3 were calculated using the module psignifit (Schütt et al., 2016), set to fit a cumulative normal sigmoid function. With all parameters free and with 95% confidence intervals. Resulting in a function that can be expressed as follows:

$$\psi(x; m, w, \lambda, \gamma) = \gamma + (1 - \lambda - \gamma) S(x; m, w)$$

Where m represents the threshold, w represents the width λ, and γ represents the upper and lower asymptote, respectively (Eq. (1) in ref. (Schütt et al., 2016)).

3.6 Acknowledgments

We would like to thank Prof. Xioaqin Wang for providing us with the MATLAB source code for synthetizing the vocalization stimuli used in Experiment 4.

3.7 Supplementary material

The supplementary material contains the following data:

Figure S3.1 Hit rates of anchor synthetic vocalizations across all testing phases of Experiment 4.

Table S3.1 Spectro-temporal characteristics of the five subsets of complex trains used across each testing phase of Experiment 4.

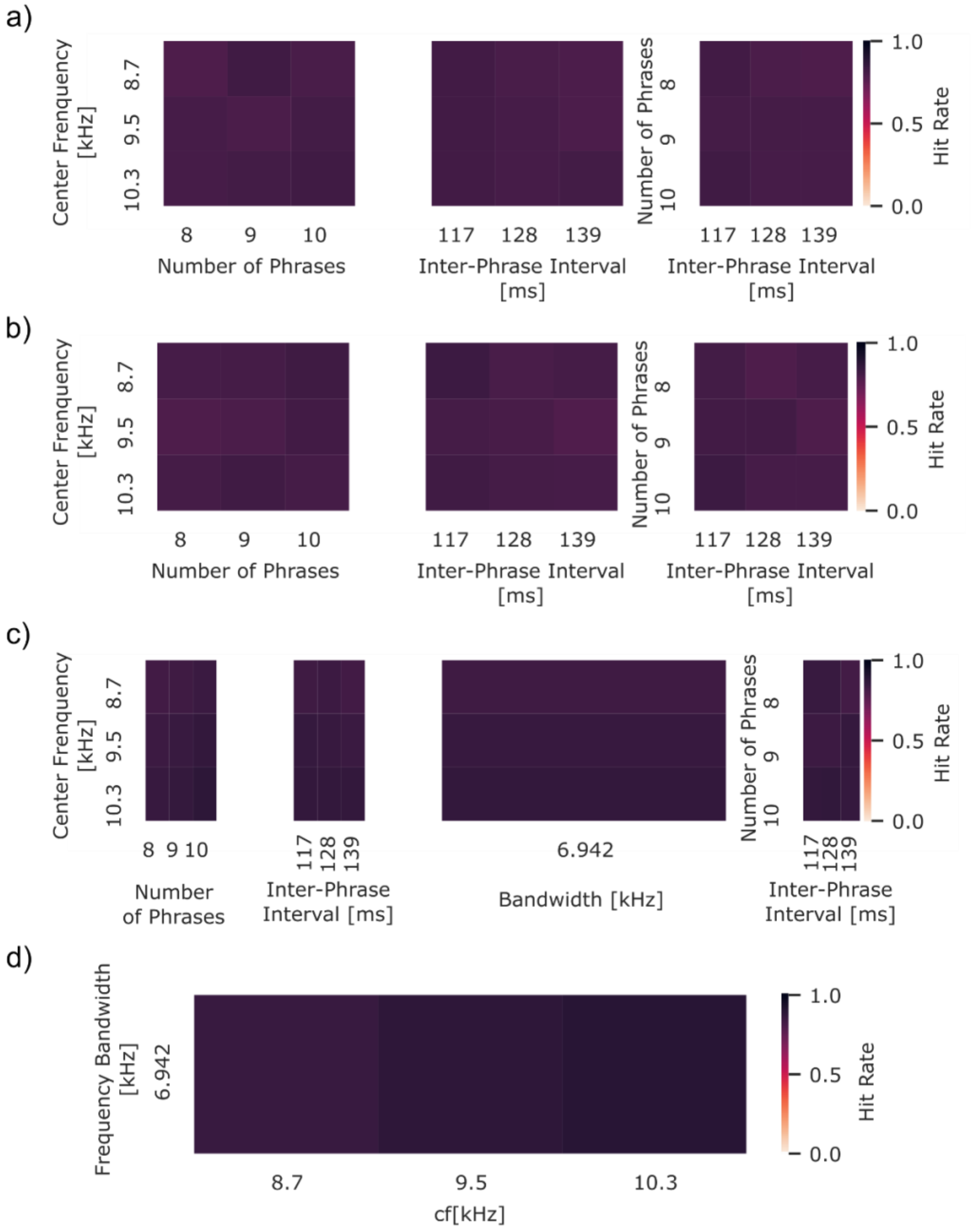


Figure S3.1. Hit rates for synthetic vocalizations. a) *Heat map showing the hit rate of the anchor synthetic twitters across stimulus sprectro-temporal characteristics of Phase-1.* ***b)*** *Heat map showing the hit rate of anchor synthetic twitters across stimulus spectro-temporal characteristics of Phase-2.* ***c)*** *Heat map showing the hit rate of anchor synthetic twitters across stimulus spectro-tempora characteristics of Phase-3.* ***d)*** *Heat map showing the hit rate of anchor synthetic twitters across stimulus spectro-temporal characteristics of Phase-4.*

Subset 1 of complex trains

Stimulus	CF [kHz]	NPH	IPI [ms]	BW [kHz]	NHR	RPA	AMC
CTT	2	8	117	0	1	0	0
CTT	2	8	128	0	1	0	0
CTT	2	8	139	0	1	0	0
CTT	2	9	117	0	1	0	0
CTT	2	9	128	0	1	0	0
CTT	2	9	139	0	1	0	0
CTT	2	10	117	0	1	0	0
CTT	2	10	128	0	1	0	0
CTT	2	10	139	0	1	0	0
CTT	3	8	117	0	1	0	0
CTT	3	8	128	0	1	0	0
CTT	3	8	139	0	1	0	0
CTT	3	9	117	0	1	0	0
CTT	3	9	128	0	1	0	0
CTT	3	9	139	0	1	0	0
CTT	3	10	117	0	1	0	0
CTT	3	10	128	0	1	0	0
CTT	3	10	139	0	1	0	0
CTT	4	8	117	0	1	0	0
CTT	4	8	128	0	1	0	0
CTT	4	8	139	0	1	0	0
CTT	4	9	117	0	1	0	0
CTT	4	9	128	0	1	0	0
CTT	4	9	139	0	1	0	0
CTT	4	10	117	0	1	0	0
CTT	4	10	128	0	1	0	0
CTT	4	10	139	0	1	0	0
CTT	5	8	117	0	1	0	0
CTT	5	8	128	0	1	0	0
CTT	5	8	139	0	1	0	0
CTT	5	9	117	0	1	0	0
CTT	5	9	128	0	1	0	0
CTT	5	9	139	0	1	0	0
CTT	5	10	117	0	1	0	0

CTT	5	10	128	0	1	0	0
CTT	5	10	139	0	1	0	0
CTT	6	8	117	0	1	0	0
CTT	6	8	128	0	1	0	0
CTT	6	8	139	0	1	0	0
CTT	6	9	117	0	1	0	0
CTT	6	9	128	0	1	0	0
CTT	6	9	139	0	1	0	0
CTT	6	10	117	0	1	0	0
CTT	6	10	128	0	1	0	0
CTT	6	10	139	0	1	0	0
CTT	7	8	117	0	1	0	0
CTT	7	8	128	0	1	0	0
CTT	7	8	139	0	1	0	0
CTT	7	9	117	0	1	0	0
CTT	7	9	128	0	1	0	0
CTT	7	9	139	0	1	0	0
CTT	7	10	117	0	1	0	0
CTT	7	10	128	0	1	0	0
CTT	7	10	139	0	1	0	0
CTT	8.7	8	117	0	1	0	0
CTT	8.7	8	128	0	1	0	0
CTT	8.7	8	139	0	1	0	0
CTT	8.7	9	117	0	1	0	0
CTT	8.7	9	128	0	1	0	0
CTT	8.7	9	139	0	1	0	0
CTT	8.7	10	117	0	1	0	0
CTT	8.7	10	128	0	1	0	0
CTT	8.7	10	139	0	1	0	0
CTT	9.5	8	117	0	1	0	0
CTT	9.5	8	128	0	1	0	0
CTT	9.5	8	139	0	1	0	0
CTT	9.5	9	117	0	1	0	0
CTT	9.5	9	128	0	1	0	0
CTT	9.5	9	139	0	1	0	0
CTT	9.5	10	117	0	1	0	0
CTT	9.5	10	128	0	1	0	0

CTT	9.5	10	139	0	1	0	0
CTT	10.3	8	117	0	1	0	0
CTT	10.3	8	128	0	1	0	0
CTT	10.3	8	139	0	1	0	0
CTT	10.3	9	117	0	1	0	0
CTT	10.3	9	128	0	1	0	0
CTT	10.3	9	139	0	1	0	0
CTT	10.3	10	117	0	1	0	0
CTT	10.3	10	128	0	1	0	0
CTT	10.3	10	139	0	1	0	0
CTT	13	8	117	0	1	0	0
CTT	13	8	128	0	1	0	0
CTT	13	8	139	0	1	0	0
CTT	13	9	117	0	1	0	0
CTT	13	9	128	0	1	0	0
CTT	13	9	139	0	1	0	0
CTT	13	10	117	0	1	0	0
CTT	13	10	128	0	1	0	0
CTT	13	10	139	0	1	0	0
CTT	15	8	117	0	1	0	0
CTT	15	8	128	0	1	0	0
CTT	15	8	139	0	1	0	0
CTT	15	9	117	0	1	0	0
CTT	15	9	128	0	1	0	0
CTT	15	9	139	0	1	0	0
CTT	15	10	117	0	1	0	0
CTT	15	10	128	0	1	0	0
CTT	15	10	139	0	1	0	0
CTT	20	8	117	0	1	0	0
CTT	20	8	128	0	1	0	0
CTT	20	8	139	0	1	0	0
CTT	20	9	117	0	1	0	0
CTT	20	9	128	0	1	0	0
CTT	20	9	139	0	1	0	0
CTT	20	10	117	0	1	0	0
CTT	20	10	128	0	1	0	0
CTT	20	10	139	0	1	0	0

Subset 2 and 3 of complex trains

Stimulus	CF [kHz]	NPH	IPI [ms]	BW [kHz]	NHR	RPA	AMC
CTT	2	8	117	0	1	1	0
CTT	2	8	128	0	1	1	0
CTT	2	8	139	0	1	1	0
CTT	2	9	117	0	1	1	0
CTT	2	9	128	0	1	1	0
CTT	2	9	139	0	1	1	0
CTT	2	10	117	0	1	1	0
CTT	2	10	128	0	1	1	0
CTT	2	10	139	0	1	1	0
CTT	3	8	117	0	1	1	0
CTT	3	8	128	0	1	1	0
CTT	3	8	139	0	1	1	0
CTT	3	9	117	0	1	1	0
CTT	3	9	128	0	1	1	0
CTT	3	9	139	0	1	1	0
CTT	3	10	117	0	1	1	0
CTT	3	10	128	0	1	1	0
CTT	3	10	139	0	1	1	0
CTT	4	8	117	0	1	1	0
CTT	4	8	128	0	1	1	0
CTT	4	8	139	0	1	1	0
CTT	4	9	117	0	1	1	0
CTT	4	9	128	0	1	1	0
CTT	4	9	139	0	1	1	0
CTT	4	10	117	0	1	1	0
CTT	4	10	128	0	1	1	0
CTT	4	10	139	0	1	1	0
CTT	5	8	117	0	1	1	0
CTT	5	8	128	0	1	1	0
CTT	5	8	139	0	1	1	0
CTT	5	9	117	0	1	1	0
CTT	5	9	128	0	1	1	0

CTT	5	9	139	0	1	1	0
CTT	5	10	117	0	1	1	0
CTT	5	10	128	0	1	1	0
CTT	5	10	139	0	1	1	0
CTT	6	8	117	0	1	1	0
CTT	6	8	128	0	1	1	0
CTT	6	8	139	0	1	1	0
CTT	6	9	117	0	1	1	0
CTT	6	9	128	0	1	1	0
CTT	6	9	139	0	1	1	0
CTT	6	10	117	0	1	1	0
CTT	6	10	128	0	1	1	0
CTT	6	10	139	0	1	1	0
CTT	7	8	117	0	1	1	0
CTT	7	8	128	0	1	1	0
CTT	7	8	139	0	1	1	0
CTT	7	9	117	0	1	1	0
CTT	7	9	128	0	1	1	0
CTT	7	9	139	0	1	1	0
CTT	7	10	117	0	1	1	0
CTT	7	10	128	0	1	1	0
CTT	7	10	139	0	1	1	0
CTT	8.7	8	117	0	1	1	0
CTT	8.7	8	128	0	1	1	0
CTT	8.7	8	139	0	1	1	0
CTT	8.7	9	117	0	1	1	0
CTT	8.7	9	128	0	1	1	0
CTT	8.7	9	139	0	1	1	0
CTT	8.7	10	117	0	1	1	0
CTT	8.7	10	128	0	1	1	0
CTT	8.7	10	139	0	1	1	0
CTT	9.5	8	117	0	1	1	0
CTT	9.5	8	128	0	1	1	0
CTT	9.5	8	139	0	1	1	0
CTT	9.5	9	117	0	1	1	0
CTT	9.5	9	128	0	1	1	0
CTT	9.5	9	139	0	1	1	0

CTT	9.5	10	117	0	1	1	0
CTT	9.5	10	128	0	1	1	0
CTT	9.5	10	139	0	1	1	0
CTT	10.3	8	117	0	1	1	0
CTT	10.3	8	128	0	1	1	0
CTT	10.3	8	139	0	1	1	0
CTT	10.3	9	117	0	1	1	0
CTT	10.3	9	128	0	1	1	0
CTT	10.3	9	139	0	1	1	0
CTT	10.3	10	117	0	1	1	0
CTT	10.3	10	128	0	1	1	0
CTT	10.3	10	139	0	1	1	0
CTT	13	8	117	0	1	1	0
CTT	13	8	128	0	1	1	0
CTT	13	8	139	0	1	1	0
CTT	13	9	117	0	1	1	0
CTT	13	9	128	0	1	1	0
CTT	13	9	139	0	1	1	0
CTT	13	10	117	0	1	1	0
CTT	13	10	128	0	1	1	0
CTT	13	10	139	0	1	1	0
CTT	15	8	117	0	1	1	0
CTT	15	8	128	0	1	1	0
CTT	15	8	139	0	1	1	0
CTT	15	9	117	0	1	1	0
CTT	15	9	128	0	1	1	0
CTT	15	9	139	0	1	1	0
CTT	15	10	117	0	1	1	0
CTT	15	10	128	0	1	1	0
CTT	15	10	139	0	1	1	0
CTT	20	8	117	0	1	1	0
CTT	20	8	128	0	1	1	0
CTT	20	8	139	0	1	1	0
CTT	20	9	117	0	1	1	0
CTT	20	9	128	0	1	1	0
CTT	20	9	139	0	1	1	0
CTT	20	10	117	0	1	1	0

CTT	20	10	128	0	1	1	0
CTT	20	10	139	0	1	1	0
CTT	2	8	117	0	1	1	1
CTT	2	8	128	0	1	1	1
CTT	2	8	139	0	1	1	1
CTT	2	9	117	0	1	1	1
CTT	2	9	128	0	1	1	1
CTT	2	9	139	0	1	1	1
CTT	2	10	117	0	1	1	1
CTT	2	10	128	0	1	1	1
CTT	2	10	139	0	1	1	1
CTT	3	8	117	0	1	1	1
CTT	3	8	128	0	1	1	1
CTT	3	8	139	0	1	1	1
CTT	3	9	117	0	1	1	1
CTT	3	9	128	0	1	1	1
CTT	3	9	139	0	1	1	1
CTT	3	10	117	0	1	1	1
CTT	3	10	128	0	1	1	1
CTT	3	10	139	0	1	1	1
CTT	4	8	117	0	1	1	1
CTT	4	8	128	0	1	1	1
CTT	4	8	139	0	1	1	1
CTT	4	9	117	0	1	1	1
CTT	4	9	128	0	1	1	1
CTT	4	9	139	0	1	1	1
CTT	4	10	117	0	1	1	1
CTT	4	10	128	0	1	1	1
CTT	4	10	139	0	1	1	1
CTT	5	8	117	0	1	1	1
CTT	5	8	128	0	1	1	1
CTT	5	8	139	0	1	1	1
CTT	5	9	117	0	1	1	1
CTT	5	9	128	0	1	1	1
CTT	5	9	139	0	1	1	1
CTT	5	10	117	0	1	1	1
CTT	5	10	128	0	1	1	1

CTT	5	10	139	0	1	1	1
CTT	6	8	117	0	1	1	1
CTT	6	8	128	0	1	1	1
CTT	6	8	139	0	1	1	1
CTT	6	9	117	0	1	1	1
CTT	6	9	128	0	1	1	1
CTT	6	9	139	0	1	1	1
CTT	6	10	117	0	1	1	1
CTT	6	10	128	0	1	1	1
CTT	6	10	139	0	1	1	1
CTT	7	8	117	0	1	1	1
CTT	7	8	128	0	1	1	1
CTT	7	8	139	0	1	1	1
CTT	7	9	117	0	1	1	1
CTT	7	9	128	0	1	1	1
CTT	7	9	139	0	1	1	1
CTT	7	10	117	0	1	1	1
CTT	7	10	128	0	1	1	1
CTT	7	10	139	0	1	1	1
CTT	8.7	8	117	0	1	1	1
CTT	8.7	8	128	0	1	1	1
CTT	8.7	8	139	0	1	1	1
CTT	8.7	9	117	0	1	1	1
CTT	8.7	9	128	0	1	1	1
CTT	8.7	9	139	0	1	1	1
CTT	8.7	10	117	0	1	1	1
CTT	8.7	10	128	0	1	1	1
CTT	8.7	10	139	0	1	1	1
CTT	9.5	8	117	0	1	1	1
CTT	9.5	8	128	0	1	1	1
CTT	9.5	8	139	0	1	1	1
CTT	9.5	9	117	0	1	1	1
CTT	9.5	9	128	0	1	1	1
CTT	9.5	9	139	0	1	1	1
CTT	9.5	10	117	0	1	1	1
CTT	9.5	10	128	0	1	1	1
CTT	9.5	10	139	0	1	1	1

CTT	10.3	8	117	0	1	1	1
CTT	10.3	8	128	0	1	1	1
CTT	10.3	8	139	0	1	1	1
CTT	10.3	9	117	0	1	1	1
CTT	10.3	9	128	0	1	1	1
CTT	10.3	9	139	0	1	1	1
CTT	10.3	10	117	0	1	1	1
CTT	10.3	10	128	0	1	1	1
CTT	10.3	10	139	0	1	1	1
CTT	13	8	117	0	1	1	1
CTT	13	8	128	0	1	1	1
CTT	13	8	139	0	1	1	1
CTT	13	9	117	0	1	1	1
CTT	13	9	128	0	1	1	1
CTT	13	9	139	0	1	1	1
CTT	13	10	117	0	1	1	1
CTT	13	10	128	0	1	1	1
CTT	13	10	139	0	1	1	1
CTT	15	8	117	0	1	1	1
CTT	15	8	128	0	1	1	1
CTT	15	8	139	0	1	1	1
CTT	15	9	117	0	1	1	1
CTT	15	9	128	0	1	1	1
CTT	15	9	139	0	1	1	1
CTT	15	10	117	0	1	1	1
CTT	15	10	128	0	1	1	1
CTT	15	10	139	0	1	1	1
CTT	20	8	117	0	1	1	1
CTT	20	8	128	0	1	1	1
CTT	20	8	139	0	1	1	1
CTT	20	9	117	0	1	1	1
CTT	20	9	128	0	1	1	1
CTT	20	9	139	0	1	1	1
CTT	20	10	117	0	1	1	1
CTT	20	10	128	0	1	1	1
CTT	20	10	139	0	1	1	1

Subset 4 of complex trains

Stimulus	CF [kHz]	NPH	IPI [ms]	BW [kHz]	NHR	RPA	AMC
CTT	8.7	8	117	30	1	1	1
CTT	8.7	8	117	65	1	1	1
CTT	8.7	8	117	125	1	1	1
CTT	8.7	8	117	250	1	1	1
CTT	8.7	8	117	500	1	1	1
CTT	8.7	8	117	1000	1	1	1
CTT	8.7	8	117	2000	1	1	1
CTT	8.7	8	117	3000	1	1	1
CTT	8.7	8	117	4000	1	1	1
CTT	8.7	8	117	5000	1	1	1
CTT	8.7	8	117	6942	1	1	1
CTT	8.7	8	128	30	1	1	1
CTT	8.7	8	128	65	1	1	1
CTT	8.7	8	128	125	1	1	1
CTT	8.7	8	128	250	1	1	1
CTT	8.7	8	128	500	1	1	1
CTT	8.7	8	128	1000	1	1	1
CTT	8.7	8	128	2000	1	1	1
CTT	8.7	8	128	3000	1	1	1
CTT	8.7	8	128	4000	1	1	1
CTT	8.7	8	128	5000	1	1	1
CTT	8.7	8	128	6942	1	1	1
CTT	8.7	8	139	30	1	1	1
CTT	8.7	8	139	65	1	1	1
CTT	8.7	8	139	125	1	1	1
CTT	8.7	8	139	250	1	1	1
CTT	8.7	8	139	500	1	1	1
CTT	8.7	8	139	1000	1	1	1
CTT	8.7	8	139	2000	1	1	1
CTT	8.7	8	139	3000	1	1	1
CTT	8.7	8	139	4000	1	1	1
CTT	8.7	8	139	5000	1	1	1
CTT	8.7	8	139	6942	1	1	1
CTT	8.7	9	117	30	1	1	1

CTT	8.7	9	117	65	1	1	1
CTT	8.7	9	117	125	1	1	1
CTT	8.7	9	117	250	1	1	1
CTT	8.7	9	117	500	1	1	1
CTT	8.7	9	117	1000	1	1	1
CTT	8.7	9	117	2000	1	1	1
CTT	8.7	9	117	3000	1	1	1
CTT	8.7	9	117	4000	1	1	1
CTT	8.7	9	117	5000	1	1	1
CTT	8.7	9	117	6942	1	1	1
CTT	8.7	9	128	30	1	1	1
CTT	8.7	9	128	65	1	1	1
CTT	8.7	9	128	125	1	1	1
CTT	8.7	9	128	250	1	1	1
CTT	8.7	9	128	500	1	1	1
CTT	8.7	9	128	1000	1	1	1
CTT	8.7	9	128	2000	1	1	1
CTT	8.7	9	128	3000	1	1	1
CTT	8.7	9	128	4000	1	1	1
CTT	8.7	9	128	5000	1	1	1
CTT	8.7	9	128	6942	1	1	1
CTT	8.7	9	139	30	1	1	1
CTT	8.7	9	139	65	1	1	1
CTT	8.7	9	139	125	1	1	1
CTT	8.7	9	139	250	1	1	1
CTT	8.7	9	139	500	1	1	1
CTT	8.7	9	139	1000	1	1	1
CTT	8.7	9	139	2000	1	1	1
CTT	8.7	9	139	3000	1	1	1
CTT	8.7	9	139	4000	1	1	1
CTT	8.7	9	139	5000	1	1	1
CTT	8.7	9	139	6942	1	1	1
CTT	8.7	10	117	30	1	1	1
CTT	8.7	10	117	65	1	1	1
CTT	8.7	10	117	125	1	1	1
CTT	8.7	10	117	250	1	1	1
CTT	8.7	10	117	500	1	1	1

CTT	8.7	10	117	1000	1	1	1
CTT	8.7	10	117	2000	1	1	1
CTT	8.7	10	117	3000	1	1	1
CTT	8.7	10	117	4000	1	1	1
CTT	8.7	10	117	5000	1	1	1
CTT	8.7	10	117	6942	1	1	1
CTT	8.7	10	128	30	1	1	1
CTT	8.7	10	128	65	1	1	1
CTT	8.7	10	128	125	1	1	1
CTT	8.7	10	128	250	1	1	1
CTT	8.7	10	128	500	1	1	1
CTT	8.7	10	128	1000	1	1	1
CTT	8.7	10	128	2000	1	1	1
CTT	8.7	10	128	3000	1	1	1
CTT	8.7	10	128	4000	1	1	1
CTT	8.7	10	128	5000	1	1	1
CTT	8.7	10	128	6942	1	1	1
CTT	8.7	10	139	30	1	1	1
CTT	8.7	10	139	65	1	1	1
CTT	8.7	10	139	125	1	1	1
CTT	8.7	10	139	250	1	1	1
CTT	8.7	10	139	500	1	1	1
CTT	8.7	10	139	1000	1	1	1
CTT	8.7	10	139	2000	1	1	1
CTT	8.7	10	139	3000	1	1	1
CTT	8.7	10	139	4000	1	1	1
CTT	8.7	10	139	5000	1	1	1
CTT	8.7	10	139	6942	1	1	1
CTT	9.5	8	117	30	1	1	1
CTT	9.5	8	117	65	1	1	1
CTT	9.5	8	117	125	1	1	1
CTT	9.5	8	117	250	1	1	1
CTT	9.5	8	117	500	1	1	1
CTT	9.5	8	117	1000	1	1	1
CTT	9.5	8	117	2000	1	1	1
CTT	9.5	8	117	3000	1	1	1
CTT	9.5	8	117	4000	1	1	1

CTT	9.5	8	117	5000	1	1	1
CTT	9.5	8	117	6942	1	1	1
CTT	9.5	8	128	30	1	1	1
CTT	9.5	8	128	65	1	1	1
CTT	9.5	8	128	125	1	1	1
CTT	9.5	8	128	250	1	1	1
CTT	9.5	8	128	500	1	1	1
CTT	9.5	8	128	1000	1	1	1
CTT	9.5	8	128	2000	1	1	1
CTT	9.5	8	128	3000	1	1	1
CTT	9.5	8	128	4000	1	1	1
CTT	9.5	8	128	5000	1	1	1
CTT	9.5	8	128	6942	1	1	1
CTT	9.5	8	139	30	1	1	1
CTT	9.5	8	139	65	1	1	1
CTT	9.5	8	139	125	1	1	1
CTT	9.5	8	139	250	1	1	1
CTT	9.5	8	139	500	1	1	1
CTT	9.5	8	139	1000	1	1	1
CTT	9.5	8	139	2000	1	1	1
CTT	9.5	8	139	3000	1	1	1
CTT	9.5	8	139	4000	1	1	1
CTT	9.5	8	139	5000	1	1	1
CTT	9.5	8	139	6942	1	1	1
CTT	9.5	9	117	30	1	1	1
CTT	9.5	9	117	65	1	1	1
CTT	9.5	9	117	125	1	1	1
CTT	9.5	9	117	250	1	1	1
CTT	9.5	9	117	500	1	1	1
CTT	9.5	9	117	1000	1	1	1
CTT	9.5	9	117	2000	1	1	1
CTT	9.5	9	117	3000	1	1	1
CTT	9.5	9	117	4000	1	1	1
CTT	9.5	9	117	5000	1	1	1
CTT	9.5	9	117	6942	1	1	1
CTT	9.5	9	128	30	1	1	1
CTT	9.5	9	128	65	1	1	1

CTT	9.5	9	128	125	1	1	1
CTT	9.5	9	128	250	1	1	1
CTT	9.5	9	128	500	1	1	1
CTT	9.5	9	128	1000	1	1	1
CTT	9.5	9	128	2000	1	1	1
CTT	9.5	9	128	3000	1	1	1
CTT	9.5	9	128	4000	1	1	1
CTT	9.5	9	128	5000	1	1	1
CTT	9.5	9	128	6942	1	1	1
CTT	9.5	9	139	30	1	1	1
CTT	9.5	9	139	65	1	1	1
CTT	9.5	9	139	125	1	1	1
CTT	9.5	9	139	250	1	1	1
CTT	9.5	9	139	500	1	1	1
CTT	9.5	9	139	1000	1	1	1
CTT	9.5	9	139	2000	1	1	1
CTT	9.5	9	139	3000	1	1	1
CTT	9.5	9	139	4000	1	1	1
CTT	9.5	9	139	5000	1	1	1
CTT	9.5	9	139	6942	1	1	1
CTT	9.5	10	117	30	1	1	1
CTT	9.5	10	117	65	1	1	1
CTT	9.5	10	117	125	1	1	1
CTT	9.5	10	117	250	1	1	1
CTT	9.5	10	117	500	1	1	1
CTT	9.5	10	117	1000	1	1	1
CTT	9.5	10	117	2000	1	1	1
CTT	9.5	10	117	3000	1	1	1
CTT	9.5	10	117	4000	1	1	1
CTT	9.5	10	117	5000	1	1	1
CTT	9.5	10	117	6942	1	1	1
CTT	9.5	10	128	30	1	1	1
CTT	9.5	10	128	65	1	1	1
CTT	9.5	10	128	125	1	1	1
CTT	9.5	10	128	250	1	1	1
CTT	9.5	10	128	500	1	1	1
CTT	9.5	10	128	1000	1	1	1

CTT	9.5	10	128	2000	1	1	1
CTT	9.5	10	128	3000	1	1	1
CTT	9.5	10	128	4000	1	1	1
CTT	9.5	10	128	5000	1	1	1
CTT	9.5	10	128	6942	1	1	1
CTT	9.5	10	139	30	1	1	1
CTT	9.5	10	139	65	1	1	1
CTT	9.5	10	139	125	1	1	1
CTT	9.5	10	139	250	1	1	1
CTT	9.5	10	139	500	1	1	1
CTT	9.5	10	139	1000	1	1	1
CTT	9.5	10	139	2000	1	1	1
CTT	9.5	10	139	3000	1	1	1
CTT	9.5	10	139	4000	1	1	1
CTT	9.5	10	139	5000	1	1	1
CTT	9.5	10	139	6942	1	1	1
CTT	10.3	8	117	30	1	1	1
CTT	10.3	8	117	65	1	1	1
CTT	10.3	8	117	125	1	1	1
CTT	10.3	8	117	250	1	1	1
CTT	10.3	8	117	500	1	1	1
CTT	10.3	8	117	1000	1	1	1
CTT	10.3	8	117	2000	1	1	1
CTT	10.3	8	117	3000	1	1	1
CTT	10.3	8	117	4000	1	1	1
CTT	10.3	8	117	5000	1	1	1
CTT	10.3	8	117	6942	1	1	1
CTT	10.3	8	128	30	1	1	1
CTT	10.3	8	128	65	1	1	1
CTT	10.3	8	128	125	1	1	1
CTT	10.3	8	128	250	1	1	1
CTT	10.3	8	128	500	1	1	1
CTT	10.3	8	128	1000	1	1	1
CTT	10.3	8	128	2000	1	1	1
CTT	10.3	8	128	3000	1	1	1
CTT	10.3	8	128	4000	1	1	1
CTT	10.3	8	128	5000	1	1	1

CTT	10.3	8	128	6942	1	1	1
CTT	10.3	8	139	30	1	1	1
CTT	10.3	8	139	65	1	1	1
CTT	10.3	8	139	125	1	1	1
CTT	10.3	8	139	250	1	1	1
CTT	10.3	8	139	500	1	1	1
CTT	10.3	8	139	1000	1	1	1
CTT	10.3	8	139	2000	1	1	1
CTT	10.3	8	139	3000	1	1	1
CTT	10.3	8	139	4000	1	1	1
CTT	10.3	8	139	5000	1	1	1
CTT	10.3	8	139	6942	1	1	1
CTT	10.3	9	117	30	1	1	1
CTT	10.3	9	117	65	1	1	1
CTT	10.3	9	117	125	1	1	1
CTT	10.3	9	117	250	1	1	1
CTT	10.3	9	117	500	1	1	1
CTT	10.3	9	117	1000	1	1	1
CTT	10.3	9	117	2000	1	1	1
CTT	10.3	9	117	3000	1	1	1
CTT	10.3	9	117	4000	1	1	1
CTT	10.3	9	117	5000	1	1	1
CTT	10.3	9	117	6942	1	1	1
CTT	10.3	9	128	30	1	1	1
CTT	10.3	9	128	65	1	1	1
CTT	10.3	9	128	125	1	1	1
CTT	10.3	9	128	250	1	1	1
CTT	10.3	9	128	500	1	1	1
CTT	10.3	9	128	1000	1	1	1
CTT	10.3	9	128	2000	1	1	1
CTT	10.3	9	128	3000	1	1	1
CTT	10.3	9	128	4000	1	1	1
CTT	10.3	9	128	5000	1	1	1
CTT	10.3	9	128	6942	1	1	1
CTT	10.3	9	139	30	1	1	1
CTT	10.3	9	139	65	1	1	1
CTT	10.3	9	139	125	1	1	1

CTT	10.3	9	139	250	1	1	1
CTT	10.3	9	139	500	1	1	1
CTT	10.3	9	139	1000	1	1	1
CTT	10.3	9	139	2000	1	1	1
CTT	10.3	9	139	3000	1	1	1
CTT	10.3	9	139	4000	1	1	1
CTT	10.3	9	139	5000	1	1	1
CTT	10.3	9	139	6942	1	1	1
CTT	10.3	10	117	30	1	1	1
CTT	10.3	10	117	65	1	1	1
CTT	10.3	10	117	125	1	1	1
CTT	10.3	10	117	250	1	1	1
CTT	10.3	10	117	500	1	1	1
CTT	10.3	10	117	1000	1	1	1
CTT	10.3	10	117	2000	1	1	1
CTT	10.3	10	117	3000	1	1	1
CTT	10.3	10	117	4000	1	1	1
CTT	10.3	10	117	5000	1	1	1
CTT	10.3	10	117	6942	1	1	1
CTT	10.3	10	128	30	1	1	1
CTT	10.3	10	128	65	1	1	1
CTT	10.3	10	128	125	1	1	1
CTT	10.3	10	128	250	1	1	1
CTT	10.3	10	128	500	1	1	1
CTT	10.3	10	128	1000	1	1	1
CTT	10.3	10	128	2000	1	1	1
CTT	10.3	10	128	3000	1	1	1
CTT	10.3	10	128	4000	1	1	1
CTT	10.3	10	128	5000	1	1	1
CTT	10.3	10	128	6942	1	1	1
CTT	10.3	10	139	30	1	1	1
CTT	10.3	10	139	65	1	1	1
CTT	10.3	10	139	125	1	1	1
CTT	10.3	10	139	250	1	1	1
CTT	10.3	10	139	500	1	1	1
CTT	10.3	10	139	1000	1	1	1
CTT	10.3	10	139	2000	1	1	1

CTT	10.3	10	139	3000	1	1	1
CTT	10.3	10	139	4000	1	1	1
CTT	10.3	10	139	5000	1	1	1
CTT	10.3	10	139	6942	1	1	1

Subset 5 of complex trains

Stimulus	CF [kHz]	NPH	IPI [ms]	BW [kHz]	NHR	RPA	AMC
CTT	8.7	8	117	30	2	1	1
CTT	8.7	8	117	65	2	1	1
CTT	8.7	8	117	125	2	1	1
CTT	8.7	8	117	250	2	1	1
CTT	8.7	8	117	500	2	1	1
CTT	8.7	8	117	1000	2	1	1
CTT	8.7	8	117	2000	2	1	1
CTT	8.7	8	117	3000	2	1	1
CTT	8.7	8	117	4000	2	1	1
CTT	8.7	8	117	5000	2	1	1
CTT	8.7	8	117	6942	2	1	1
CTT	8.7	8	128	30	2	1	1
CTT	8.7	8	128	65	2	1	1
CTT	8.7	8	128	125	2	1	1
CTT	8.7	8	128	250	2	1	1
CTT	8.7	8	128	500	2	1	1
CTT	8.7	8	128	1000	2	1	1
CTT	8.7	8	128	2000	2	1	1
CTT	8.7	8	128	3000	2	1	1
CTT	8.7	8	128	4000	2	1	1
CTT	8.7	8	128	5000	2	1	1
CTT	8.7	8	128	6942	2	1	1
CTT	8.7	8	139	30	2	1	1
CTT	8.7	8	139	65	2	1	1
CTT	8.7	8	139	125	2	1	1
CTT	8.7	8	139	250	2	1	1
CTT	8.7	8	139	500	2	1	1
CTT	8.7	8	139	1000	2	1	1

CTT	8.7	8	139	2000	2	1	1
CTT	8.7	8	139	3000	2	1	1
CTT	8.7	8	139	4000	2	1	1
CTT	8.7	8	139	5000	2	1	1
CTT	8.7	8	139	6942	2	1	1
CTT	8.7	9	117	30	2	1	1
CTT	8.7	9	117	65	2	1	1
CTT	8.7	9	117	125	2	1	1
CTT	8.7	9	117	250	2	1	1
CTT	8.7	9	117	500	2	1	1
CTT	8.7	9	117	1000	2	1	1
CTT	8.7	9	117	2000	2	1	1
CTT	8.7	9	117	3000	2	1	1
CTT	8.7	9	117	4000	2	1	1
CTT	8.7	9	117	5000	2	1	1
CTT	8.7	9	117	6942	2	1	1
CTT	8.7	9	128	30	2	1	1
CTT	8.7	9	128	65	2	1	1
CTT	8.7	9	128	125	2	1	1
CTT	8.7	9	128	250	2	1	1
CTT	8.7	9	128	500	2	1	1
CTT	8.7	9	128	1000	2	1	1
CTT	8.7	9	128	2000	2	1	1
CTT	8.7	9	128	3000	2	1	1
CTT	8.7	9	128	4000	2	1	1
CTT	8.7	9	128	5000	2	1	1
CTT	8.7	9	128	6942	2	1	1
CTT	8.7	9	139	30	2	1	1
CTT	8.7	9	139	65	2	1	1
CTT	8.7	9	139	125	2	1	1
CTT	8.7	9	139	250	2	1	1
CTT	8.7	9	139	500	2	1	1
CTT	8.7	9	139	1000	2	1	1
CTT	8.7	9	139	2000	2	1	1
CTT	8.7	9	139	3000	2	1	1
CTT	8.7	9	139	4000	2	1	1
CTT	8.7	9	139	5000	2	1	1

CTT	8.7	9	139	6942	2	1	1
CTT	8.7	10	117	30	2	1	1
CTT	8.7	10	117	65	2	1	1
CTT	8.7	10	117	125	2	1	1
CTT	8.7	10	117	250	2	1	1
CTT	8.7	10	117	500	2	1	1
CTT	8.7	10	117	1000	2	1	1
CTT	8.7	10	117	2000	2	1	1
CTT	8.7	10	117	3000	2	1	1
CTT	8.7	10	117	4000	2	1	1
CTT	8.7	10	117	5000	2	1	1
CTT	8.7	10	117	6942	2	1	1
CTT	8.7	10	128	30	2	1	1
CTT	8.7	10	128	65	2	1	1
CTT	8.7	10	128	125	2	1	1
CTT	8.7	10	128	250	2	1	1
CTT	8.7	10	128	500	2	1	1
CTT	8.7	10	128	1000	2	1	1
CTT	8.7	10	128	2000	2	1	1
CTT	8.7	10	128	3000	2	1	1
CTT	8.7	10	128	4000	2	1	1
CTT	8.7	10	128	5000	2	1	1
CTT	8.7	10	128	6942	2	1	1
CTT	8.7	10	139	30	2	1	1
CTT	8.7	10	139	65	2	1	1
CTT	8.7	10	139	125	2	1	1
CTT	8.7	10	139	250	2	1	1
CTT	8.7	10	139	500	2	1	1
CTT	8.7	10	139	1000	2	1	1
CTT	8.7	10	139	2000	2	1	1
CTT	8.7	10	139	3000	2	1	1
CTT	8.7	10	139	4000	2	1	1
CTT	8.7	10	139	5000	2	1	1
CTT	8.7	10	139	6942	2	1	1
CTT	9.5	8	117	30	2	1	1
CTT	9.5	8	117	65	2	1	1
CTT	9.5	8	117	125	2	1	1

CTT	9.5	8	117	250	2	1	1
CTT	9.5	8	117	500	2	1	1
CTT	9.5	8	117	1000	2	1	1
CTT	9.5	8	117	2000	2	1	1
CTT	9.5	8	117	3000	2	1	1
CTT	9.5	8	117	4000	2	1	1
CTT	9.5	8	117	5000	2	1	1
CTT	9.5	8	117	6942	2	1	1
CTT	9.5	8	128	30	2	1	1
CTT	9.5	8	128	65	2	1	1
CTT	9.5	8	128	125	2	1	1
CTT	9.5	8	128	250	2	1	1
CTT	9.5	8	128	500	2	1	1
CTT	9.5	8	128	1000	2	1	1
CTT	9.5	8	128	2000	2	1	1
CTT	9.5	8	128	3000	2	1	1
CTT	9.5	8	128	4000	2	1	1
CTT	9.5	8	128	5000	2	1	1
CTT	9.5	8	128	6942	2	1	1
CTT	9.5	8	139	30	2	1	1
CTT	9.5	8	139	65	2	1	1
CTT	9.5	8	139	125	2	1	1
CTT	9.5	8	139	250	2	1	1
CTT	9.5	8	139	500	2	1	1
CTT	9.5	8	139	1000	2	1	1
CTT	9.5	8	139	2000	2	1	1
CTT	9.5	8	139	3000	2	1	1
CTT	9.5	8	139	4000	2	1	1
CTT	9.5	8	139	5000	2	1	1
CTT	9.5	8	139	6942	2	1	1
CTT	9.5	9	117	30	2	1	1
CTT	9.5	9	117	65	2	1	1
CTT	9.5	9	117	125	2	1	1
CTT	9.5	9	117	250	2	1	1
CTT	9.5	9	117	500	2	1	1
CTT	9.5	9	117	1000	2	1	1
CTT	9.5	9	117	2000	2	1	1

CTT	9.5	9	117	3000	2	1	1
CTT	9.5	9	117	4000	2	1	1
CTT	9.5	9	117	5000	2	1	1
CTT	9.5	9	117	6942	2	1	1
CTT	9.5	9	128	30	2	1	1
CTT	9.5	9	128	65	2	1	1
CTT	9.5	9	128	125	2	1	1
CTT	9.5	9	128	250	2	1	1
CTT	9.5	9	128	500	2	1	1
CTT	9.5	9	128	1000	2	1	1
CTT	9.5	9	128	2000	2	1	1
CTT	9.5	9	128	3000	2	1	1
CTT	9.5	9	128	4000	2	1	1
CTT	9.5	9	128	5000	2	1	1
CTT	9.5	9	128	6942	2	1	1
CTT	9.5	9	139	30	2	1	1
CTT	9.5	9	139	65	2	1	1
CTT	9.5	9	139	125	2	1	1
CTT	9.5	9	139	250	2	1	1
CTT	9.5	9	139	500	2	1	1
CTT	9.5	9	139	1000	2	1	1
CTT	9.5	9	139	2000	2	1	1
CTT	9.5	9	139	3000	2	1	1
CTT	9.5	9	139	4000	2	1	1
CTT	9.5	9	139	5000	2	1	1
CTT	9.5	9	139	6942	2	1	1
CTT	9.5	10	117	30	2	1	1
CTT	9.5	10	117	65	2	1	1
CTT	9.5	10	117	125	2	1	1
CTT	9.5	10	117	250	2	1	1
CTT	9.5	10	117	500	2	1	1
CTT	9.5	10	117	1000	2	1	1
CTT	9.5	10	117	2000	2	1	1
CTT	9.5	10	117	3000	2	1	1
CTT	9.5	10	117	4000	2	1	1
CTT	9.5	10	117	5000	2	1	1
CTT	9.5	10	117	6942	2	1	1

CTT	9.5	10	128	30	2	1	1
CTT	9.5	10	128	65	2	1	1
CTT	9.5	10	128	125	2	1	1
CTT	9.5	10	128	250	2	1	1
CTT	9.5	10	128	500	2	1	1
CTT	9.5	10	128	1000	2	1	1
CTT	9.5	10	128	2000	2	1	1
CTT	9.5	10	128	3000	2	1	1
CTT	9.5	10	128	4000	2	1	1
CTT	9.5	10	128	5000	2	1	1
CTT	9.5	10	128	6942	2	1	1
CTT	9.5	10	139	30	2	1	1
CTT	9.5	10	139	65	2	1	1
CTT	9.5	10	139	125	2	1	1
CTT	9.5	10	139	250	2	1	1
CTT	9.5	10	139	500	2	1	1
CTT	9.5	10	139	1000	2	1	1
CTT	9.5	10	139	2000	2	1	1
CTT	9.5	10	139	3000	2	1	1
CTT	9.5	10	139	4000	2	1	1
CTT	9.5	10	139	5000	2	1	1
CTT	9.5	10	139	6942	2	1	1
CTT	10.3	8	117	30	2	1	1
CTT	10.3	8	117	65	2	1	1
CTT	10.3	8	117	125	2	1	1
CTT	10.3	8	117	250	2	1	1
CTT	10.3	8	117	500	2	1	1
CTT	10.3	8	117	1000	2	1	1
CTT	10.3	8	117	2000	2	1	1
CTT	10.3	8	117	3000	2	1	1
CTT	10.3	8	117	4000	2	1	1
CTT	10.3	8	117	5000	2	1	1
CTT	10.3	8	117	6942	2	1	1
CTT	10.3	8	128	30	2	1	1
CTT	10.3	8	128	65	2	1	1
CTT	10.3	8	128	125	2	1	1
CTT	10.3	8	128	250	2	1	1

CTT	10.3	8	128	500	2	1	1
CTT	10.3	8	128	1000	2	1	1
CTT	10.3	8	128	2000	2	1	1
CTT	10.3	8	128	3000	2	1	1
CTT	10.3	8	128	4000	2	1	1
CTT	10.3	8	128	5000	2	1	1
CTT	10.3	8	128	6942	2	1	1
CTT	10.3	8	139	30	2	1	1
CTT	10.3	8	139	65	2	1	1
CTT	10.3	8	139	125	2	1	1
CTT	10.3	8	139	250	2	1	1
CTT	10.3	8	139	500	2	1	1
CTT	10.3	8	139	1000	2	1	1
CTT	10.3	8	139	2000	2	1	1
CTT	10.3	8	139	3000	2	1	1
CTT	10.3	8	139	4000	2	1	1
CTT	10.3	8	139	5000	2	1	1
CTT	10.3	8	139	6942	2	1	1
CTT	10.3	9	117	30	2	1	1
CTT	10.3	9	117	65	2	1	1
CTT	10.3	9	117	125	2	1	1
CTT	10.3	9	117	250	2	1	1
CTT	10.3	9	117	500	2	1	1
CTT	10.3	9	117	1000	2	1	1
CTT	10.3	9	117	2000	2	1	1
CTT	10.3	9	117	3000	2	1	1
CTT	10.3	9	117	4000	2	1	1
CTT	10.3	9	117	5000	2	1	1
CTT	10.3	9	117	6942	2	1	1
CTT	10.3	9	128	30	2	1	1
CTT	10.3	9	128	65	2	1	1
CTT	10.3	9	128	125	2	1	1
CTT	10.3	9	128	250	2	1	1
CTT	10.3	9	128	500	2	1	1
CTT	10.3	9	128	1000	2	1	1
CTT	10.3	9	128	2000	2	1	1
CTT	10.3	9	128	3000	2	1	1

CTT	10.3	9	128	4000	2	1	1
CTT	10.3	9	128	5000	2	1	1
CTT	10.3	9	128	6942	2	1	1
CTT	10.3	9	139	30	2	1	1
CTT	10.3	9	139	65	2	1	1
CTT	10.3	9	139	125	2	1	1
CTT	10.3	9	139	250	2	1	1
CTT	10.3	9	139	500	2	1	1
CTT	10.3	9	139	1000	2	1	1
CTT	10.3	9	139	2000	2	1	1
CTT	10.3	9	139	3000	2	1	1
CTT	10.3	9	139	4000	2	1	1
CTT	10.3	9	139	5000	2	1	1
CTT	10.3	9	139	6942	2	1	1
CTT	10.3	10	117	30	2	1	1
CTT	10.3	10	117	65	2	1	1
CTT	10.3	10	117	125	2	1	1
CTT	10.3	10	117	250	2	1	1
CTT	10.3	10	117	500	2	1	1
CTT	10.3	10	117	1000	2	1	1
CTT	10.3	10	117	2000	2	1	1
CTT	10.3	10	117	3000	2	1	1
CTT	10.3	10	117	4000	2	1	1
CTT	10.3	10	117	5000	2	1	1
CTT	10.3	10	117	6942	2	1	1
CTT	10.3	10	128	30	2	1	1
CTT	10.3	10	128	65	2	1	1
CTT	10.3	10	128	125	2	1	1
CTT	10.3	10	128	250	2	1	1
CTT	10.3	10	128	500	2	1	1
CTT	10.3	10	128	1000	2	1	1
CTT	10.3	10	128	2000	2	1	1
CTT	10.3	10	128	3000	2	1	1
CTT	10.3	10	128	4000	2	1	1
CTT	10.3	10	128	5000	2	1	1
CTT	10.3	10	128	6942	2	1	1
CTT	10.3	10	139	30	2	1	1

CTT	10.3	10	139	65	2	1	1
CTT	10.3	10	139	125	2	1	1
CTT	10.3	10	139	250	2	1	1
CTT	10.3	10	139	500	2	1	1
CTT	10.3	10	139	1000	2	1	1
CTT	10.3	10	139	2000	2	1	1
CTT	10.3	10	139	3000	2	1	1
CTT	10.3	10	139	4000	2	1	1
CTT	10.3	10	139	5000	2	1	1
CTT	10.3	10	139	6942	2	1	1

***Table S3.1 Spectro-temporal characteristics of complex train stimuli**. Ech row describes the spectro temporal characteristics of a stimulus across all testing phases. Center frequency of first phrase (CF), Number of phrases per stimulus (NPH), Inter-phrase interval (IPI), Frequency bandwidth of middle phrase (BW), Number of harmonics (NHR), relative-phrase amplitude (RPA), and amplitude modulated contour (AMC).*

4.8 References

Ackermann, H., Hage, S.R., Ziegler, W., 2014. Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. Behav Brain Sci 37, 529–546. https://doi.org/10.1017/S0140525X13003099

Agamaite, J.A., Chang, C.-J., Osmanski, M.S., Wang, X., 2015. A quantitative acoustic analysis of the vocal repertoire of the common marmoset (*Callithrix jacchus*). The Journal of the Acoustical Society of America 138, 2906–2928. https://doi.org/10.1121/1.4934268

Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B.G., Witzel, T., Belliveau, J.W., 2006. Task-modulated “what” and “where” pathways in human auditory cortex. Proc. Natl. Acad. Sci. U.S.A. 103, 14608–14613. https://doi.org/10.1073/pnas.0510480103

Andrews, M.W., Rosenblum, L.A., 1994. Automated recording of individual performance and hand preference during joystick-task acquisition in group-living bonnet macaques (Macaca radiata). Journal of Comparative Psychology 108, 358–362. https://doi.org/10.1037/0735-7036.108.4.358

Baddeley, A., 2003. Working memory: looking back and looking forward. Nat Rev Neurosci 4, 829–839. https://doi.org/10.1038/nrn1201

Belin, P., 2006. Voice processing in human and non-human primates. Phil. Trans. R. Soc. B 361, 2091–2107. https://doi.org/10.1098/rstb.2006.1933

Belmonte, J.C.I., Callaway, E.M., Caddick, S.J., Churchland, P., Feng, G., Homanics, G.E., Lee, K.-F., Leopold, D.A., Miller, C.T., Mitchell, J.F., Mitalipov, S., Moutri, A.R., Movshon, J.A., Okano, H., Reynolds, J.H., Ringach, D.L., Sejnowski, T.J., Silva, A.C., Strick, P.L., Wu, J., Zhang, F., 2015. Brains, Genes, and Primates. Neuron 86, 617–631. https://doi.org/10.1016/j.neuron.2015.03.021

Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165. https://doi.org/10.1038/nature03867

Bisley, J.W., Goldberg, M.E., 2003. Neuronal Activity in the Lateral Intraparietal Area and Spatial Attention. Science 299, 81–86. https://doi.org/10.1126/science.1077395

Bohlen, M., Hayes, E.R., Bohlen, B., Bailoo, J.D., Crabbe, J.C., Wahlsten, D., 2014. Experimenter effects on behavioral test scores of eight inbred mouse strains under the influence of ethanol. Behavioural Brain Research 272, 46–54. https://doi.org/10.1016/j.bbr.2014.06.017

Bonini, L., 2019. Refinement techniques in non-human primate neuroscientific research. Annali dell’Istituto Superiore di Sanità 55, 409–412.

Brewer, A.A., Barton, B., 2016. Maps of the Auditory Cortex. Annu. Rev. Neurosci. 39, 385–407. https://doi.org/10.1146/annurev-neuro-070815-014045

Bugnyar, T., Huber, L., 1997. Push or pull: an experimental study on imitation in marmosets. Animal Behaviour 54, 817–831. https://doi.org/10.1006/anbe.1996.0497

Burkart, J.M., Strasser, A., Foglia, M., 2009. Trade-offs between social learning and individual innovativeness in common marmosets, Callithrix jacchus. Animal Behaviour 77, 1291–1301. https://doi.org/10.1016/j.anbehav.2009.02.006

Calapai, A., Cabrera-Moreno, J., Moser, T., Jeschke, M., 2022. Flexible auditory training, psychophysics, and enrichment of common marmosets with an automated, touchscreen-based system. Nat Commun 13, 1648. https://doi.org/10.1038/s41467-022-29185-9

Cappe, C., Murray, M.M., Barone, P., Rouiller, E.M., 2010. Multisensory facilitation of behavior in monkeys: effects of stimulus intensity. Journal of Cognitive Neuroscience 22, 2850–2863.

Chandrasekaran, C., Lemus, L., Trubanova, A., Gondan, M., Ghazanfar, A.A., 2011. Monkeys and Humans Share a Common Computation for Face/Voice Integration. PLoS Comput Biol 7, e1002165. https://doi.org/10.1371/journal.pcbi.1002165

Choudhury, N., Gorman, K.S., 2000. The relationship between sustained attention and cognitive performance in 17-24-month old toddlers. Inf. Child Develop. 9, 127–146. https://doi.org/10.1002/1522-7219(200009)9:3<127::AID-ICD225>3.0.CO;2-5

Cibulski, L., Wascher, C.A.F., Weiß, B.M., Kotrschal, K., 2014. Familiarity with the experimenter influences the performance of Common ravens (Corvus corax) and Carrion crows (Corvus corone corone) in cognitive tasks. Behavioural Processes 103, 129–137. https://doi.org/10.1016/j.beproc.2013.11.013

Clark, F., 2017. Cognitive enrichment and welfare: Current approaches and future directions. AB&C 4, 52–71. https://doi.org/10.12966/abc.05.02.2017

Day, R.L., Coe, R.L., Kendal, J.R., Laland, K.N., 2003. Neophilia, innovation and social learning: a study of intergeneric differences in callitrichid monkeys. Animal Behaviour 65, 559–571. https://doi.org/10.1006/anbe.2003.2074

Dieter, A., Duque-Afonso, C.J., Rankovic, V., Jeschke, M., Moser, T., 2019. Near physiological spectral selectivity of cochlear optogenetics. Nat Commun 10, 1962. https://doi.org/10.1038/s41467-019-09980-7

DiMattina, C., Wang, X., 2006. Virtual Vocalization Stimuli for Investigating Neural Representations of Species-Specific Vocalizations. Journal of Neurophysiology 95, 1244–1262. https://doi.org/10.1152/jn.00818.2005

Dombrowski, T., Rankovic, V., Moser, T., 2019. Toward the Optical Cochlear Implant. Cold Spring Harb Perspect Med 9, a033225. https://doi.org/10.1101/cshperspect.a033225

Egan, J.P., Hake, H.W., 1950. On the Masking Pattern of a Simple Auditory Stimulus. The Journal of the Acoustical Society of America 22, 622–630. https://doi.org/10.1121/1.1906661

Egeth, H.E., Yantis, S., 1997. VISUAL ATTENTION: Control, Representation, and Time Course. Annu. Rev. Psychol. 48, 269–297. https://doi.org/10.1146/annurev.psych.48.1.269

Epple, G., 1968. Comparative Studies on Vocalization in Marmoset Monkeys (Hapalidae). IJFP 8, 1–40. https://doi.org/10.1159/000155129

Fagot, J., Gullstrand, J., Kemp, C., Defilles, C., Mekaouche, M., 2014. Effects of freely accessible computerized test systems on the spontaneous behaviors and stress level of Guinea baboons (*Papio papio*): Computerized Testing and Welfare in Baboons. Am. J. Primatol. 76, 56–64. https://doi.org/10.1002/ajp.22193

Fagot, J., Paleressompoulle, D., 2009. Automatic testing of cognitive performance in baboons maintained in social groups. Behavior Research Methods 41, 396–404. https://doi.org/10.3758/BRM.41.2.396

Fan-Gang Zeng, Rebscher, S., Harrison, W., Xiaoan Sun, Haihong Feng, 2008. Cochlear Implants: System Design, Integration, and Evaluation. IEEE Rev. Biomed. Eng. 1, 115–142. https://doi.org/10.1109/RBME.2008.2008250

Fayed, E.A., Zaghloul, H.S., Morgan, A.E., 2020. Electrode impedance changes over time in MED El cochlear implant children recipients: Relation to stimulation levels and behavioral measures. Cochlear Implants International 21, 192–197. https://doi.org/10.1080/14670100.2020.1716464

Fischer, J., 1998. Barbary macaques categorize shrill barks into two call types. Animal Behaviour 55, 799–807. https://doi.org/10.1006/anbe.1997.0663

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007. Auditory attention—focusing the searchlight on sound. Current Opinion in Neurobiology 17, 437–455. https://doi.org/10.1016/j.conb.2007.07.011

Furuyama, T., Kobayasi, K.I., Riquimaroux, H., 2017. Acoustic characteristics used by Japanese macaques for individual discrimination. Journal of Experimental Biology jeb.154765. https://doi.org/10.1242/jeb.154765

Gazes, R.P., Brown, E.K., Basile, B.M., Hampton, R.R., 2013. Automated cognitive testing of monkeys in social groups yields results comparable to individual laboratory-based testing. Anim Cogn 16, 445–458. https://doi.org/10.1007/s10071-012-0585-8

Ghazanfar, A.A., Rendall, D., 2008. Evolution of human vocal production. Current Biology 18, R457–R460. https://doi.org/10.1016/j.cub.2008.03.030

Gomes, H., 2000. The development of auditory attention in children. Front Biosci 5, d108. https://doi.org/10.2741/Gomes

Hernandez, V.H., Gehrt, A., Reuter, K., Jing, Z., Jeschke, M., Mendoza Schulz, A., Hoch, G., Bartels, M., Vogt, G., Garnham, C.W., Yawo, H., Fukazawa, Y., Augustine, G.J., Bamberg, E., Kügler, S., Salditt, T., de Hoz, L., Strenzke, N., Moser, T., 2014. Optogenetic stimulation of the auditory pathway. J. Clin. Invest. 124, 1114–1129. https://doi.org/10.1172/JCI69050

Hook-costigan, M.A., Rogers, L.J., 1998. Lateralized use of the mouth in production of vocalizations by marmosets. Neuropsychologia 36, 1265–1273. https://doi.org/10.1016/S0028-3932(98)00037-2

Hosoya, M., Fujioka, M., Murayama, A.Y., Ozawa, H., Okano, H., Ogawa, K., 2021. Neuronal development in the cochlea of a nonhuman primate model, the common marmoset. Develop. Neurobiol. 81, 905–938. https://doi.org/10.1002/dneu.22850

Hosoya, M., Fujioka, M., Ogawa, K., Okano, H., 2016. Distinct Expression Patterns Of Causative Genes Responsible For Hereditary Progressive Hearing Loss In Non-Human Primate Cochlea. Sci Rep 6, 22250. https://doi.org/10.1038/srep22250

Jablonski, L., Harczos, T., Wolf, B., Hoch, G., Dieter, A., Hessler, R., Ayub, S., Ruther, P., Moser, T., 2020. Hearing restoration by a low-weight power-efficient multichannel optogenetic cochlear implant system (preprint). Neuroscience. https://doi.org/10.1101/2020.05.25.114868

Jeschke, M., Moser, T., 2015. Considering optogenetic stimulation for cochlear implants. Hearing Research 322, 224–234. https://doi.org/10.1016/j.heares.2015.01.005

Johnson, J.A., Zatorre, R.J., 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. NeuroImage 31, 1673–1681. https://doi.org/10.1016/j.neuroimage.2006.02.026

Johnson, L.A., Della Santina, C.C., Wang, X., 2016. Selective Neuronal Activation by Cochlear Implant Stimulation in Auditory Cortex of Awake Primate. J. Neurosci. 36, 12468–12484. https://doi.org/10.1523/JNEUROSCI.1699-16.2016

Johnson, L.A., Della Santina, C.C., Wang, X., 2012. Temporal bone characterization and cochlear implant feasibility in the common marmoset (Callithrix jacchus). Hearing Research 290, 37–44. https://doi.org/10.1016/j.heares.2012.05.002

Kaiser, T., Feng, G., 2015. Modeling psychiatric disorders for developing effective treatments. Nat Med 21, 979–988. https://doi.org/10.1038/nm.3935

Kaplan, J.N., Winship-Ball, A., Sim, L., 1978. Maternal discrimination of infant vocalizations in squirrel monkeys. Primates 19, 187–193. https://doi.org/10.1007/BF02373235

Keppeler, D., Kampshoff, C.A., Thirumalai, A., Duque-Afonso, C.J., Schaeper, J.J., Quilitz, T., Töpperwien, M., Vogl, C., Hessler, R., Meyer, A., Salditt, T., Moser, T., 2021. Multiscale photonic imaging of the native and implanted cochlea. Proc. Natl. Acad. Sci. U.S.A. 118, e2014472118. https://doi.org/10.1073/pnas.2014472118

Keppeler, D., Merino, R.M., Lopez de la Morena, D., Bali, B., Huet, A.T., Gehrt, A., Wrobel, C., Subramanian, S., Dombrowski, T., Wolf, F., Rankovic, V., Neef, A., Moser, T., 2018. Ultrafast optogenetic stimulation of the auditory pathway by targeting-optimized Chronos. EMBO J 37. https://doi.org/10.15252/embj.201899649

King, R.D., Rowland, J., Oliver, S.G., Young, M., Aubrey, W., Byrne, E., Liakata, M., Markham, M., Pir, P., Soldatova, L.N., Sparkes, A., Whelan, K.E., Clare, A., 2009. The Automation of Science. Science 324, 85–89. https://doi.org/10.1126/science.1165620

Kishi, N., Sato, K., Sasaki, E., Okano, H., 2014. Common marmoset as a new model animal for neuroscience research and genome editing technology. Development, Growth & Differentiation 56, 53–62. https://doi.org/10.1111/dgd.12109

Knudsen, E.I., 2007. Fundamental Components of Attention. Annu. Rev. Neurosci. 30, 57–78. https://doi.org/10.1146/annurev.neuro.30.051606.094256

Lieberman, P., Blumstein, S.E., 1988. Speech Physiology, Speech Perception, and Acoustic Phonetics, 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9781139165952

Lin, D.P.-Y., Chen, J.K.-C., Tung, T.-H., Li, L.P.-H., 2019. Differences in the impedance of cochlear implant devices within 24 hours of their implantation. PLoS ONE 14, e0222711. https://doi.org/10.1371/journal.pone.0222711

Mansfield, K., 2003. Marmoset models commonly used in biomedical research. Comp. Med. 53, 383–392.

Mason, S., Premereur, E., Pelekanos, V., Emberton, A., Honess, P., Mitchell, A.S., 2019. Effective chair training methods for neuroscience research involving rhesus macaques (Macaca mulatta). Journal of Neuroscience Methods 317, 82–93.

May, B., Moody, D.B., Stebbins, W.C., 1989. Categorical perception of conspecific communication sounds by Japanese macaques, *Macacafuscata*. The Journal of the Acoustical Society of America 85, 837–847. https://doi.org/10.1121/1.397555

McKinney, W., 2010. Data Structures for Statistical Computing in Python. Presented at the Python in Science Conference, Austin, Texas, pp. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Mesgarani, N., Cheung, C., Johnson, K., Chang, E.F., 2014. Phonetic Feature Encoding in Human Superior Temporal Gyrus. Science 343, 1006–1010. https://doi.org/10.1126/science.1245994

Miller, B.T., D'Esposito, M., 2005. Searching for "the Top" in Top-Down Control. Neuron 48, 535–538. https://doi.org/10.1016/j.neuron.2005.11.002

Miller, C.T., Freiwald, W.A., Leopold, D.A., Mitchell, J.F., Silva, A.C., Wang, X., 2016. Marmosets: A Neuroscientific Model of Human Social Behavior. Neuron 90, 219–233. https://doi.org/10.1016/j.neuron.2016.03.018

Miller, J., Ulrich, R., Lamarre, Y., 2001. Locus of the redundant-signals effect in bimodal divided attention: a neurophysiological analysis. Perception & Psychophysics 63, 555–562.

Millman, K.J., Aivazis, M., 2011. Python for Scientists and Engineers. Comput. Sci. Eng. 13, 9–12. https://doi.org/10.1109/MCSE.2011.36

Moore, B.C.J., 2013. An introduction to the psychology of hearing, Sixth edition. ed. Brill, Leiden.

Moser, T., 2015. Optogenetic stimulation of the auditory pathway for research and future prosthetics. Current Opinion in Neurobiology 34, 29–36. https://doi.org/10.1016/j.conb.2015.01.004

Neuburger, J., Lenarz, T., Lesinski-Schiedat, A., Büchner, A., 2009. Spontaneous increases in impedance following cochlear implantation: Suspected causes and management. International Journal of Audiology 48, 233–239. https://doi.org/10.1080/14992020802600808

Newman, J.D., Wollberg, Z., 1973. Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. Brain Research 54, 287–304. https://doi.org/10.1016/0006-8993(73)90050-4

Oliphant, T.E., 2007. Python for Scientific Computing. Comput. Sci. Eng. 9, 10–20. https://doi.org/10.1109/MCSE.2007.58

Oliphant, T.E., 2006. A guide to Numpy. USA: Trelgol Publishing.

Ortiz-Rios, M., KuÅ›mierek, P., DeWitt, I., Archakov, D., Azevedo, F.A.C., Sams, M., JÃ¤Ã¤skelÃ¤inen, I.P., Keliris, G.A., Rauschecker, J.P., 2015. Functional MRI of the vocalization-processing network in the macaque brain. Front. Neurosci. 9. https://doi.org/10.3389/fnins.2015.00113

Osmanski, M.S., Song, X., Guo, Y., Wang, X., 2016. Frequency discrimination in the common marmoset (Callithrix jacchus). Hearing Research 341, 1–8. https://doi.org/10.1016/j.heares.2016.07.006

Osmanski, M.S., Wang, X., 2011. Measurement of absolute auditory thresholds in the common marmoset (Callithrix jacchus). Hearing Research 277, 127–133. https://doi.org/10.1016/j.heares.2011.02.001

Owren, M.J., Seyfarth, R.M., Cheney, D.L., 1997. The acoustic features of vowel-like *grunt* calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions. The Journal of the

Acoustical Society of America 101, 2951–2963. https://doi.org/10.1121/1.418523

Petkov, C.I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., Logothetis, N.K., 2008. A voice region in the monkey brain. Nat Neurosci 11, 367–374. https://doi.org/10.1038/nn2043

Pfingst, B.E., Laycock, J., Flammino, F., Lonsbury-Martin, B., Martin, G., 1978. Pure tone thresholds for the rhesus monkey. Hearing Research 1, 43–47. https://doi.org/10.1016/0378-5955(78)90008-4

Rauschecker, J.P., Tian, B., Hauser, M., 1995. Processing of Complex Sounds in the Macaque Nonprimary Auditory Cortex. Science 268, 111–114. https://doi.org/10.1126/science.7701330

Remez, R.E., Rubin, P.E., Pisoni, D.B., Carrell, T.D., 1981. Speech Perception Without Traditional Speech Cues. Science 212, 947–950. https://doi.org/10.1126/science.7233191

Remington, R.W., Johnston, J.C., Yantis, S., 1992. Involuntary attentional capture by abrupt onsets. Perception & Psychophysics 51, 279–290. https://doi.org/10.3758/BF03212254

Rendall, D., Owren, M.J., Weerts, E., Hienz, R.D., 2004. Sex differences in the acoustic structure of vowel-like grunt vocalizations in baboons and their perceptual discrimination by baboon listeners. The Journal of the Acoustical Society of America 115, 411–421. https://doi.org/10.1121/1.1635838

Repp, B.H., Liberman, A.M., Harnad, S., 1984. Phonetic category boundaries are flexible. Haskins Laboratories Status Report on Speech Research, SR-77 78, 31–53.

Risueno-Segovia, C., Hage, S.R., 2020. Theta Synchronization of Phonatory and Articulatory Systems in Marmoset Monkey Vocal Production. Current Biology 30, 4276-4283.e3. https://doi.org/10.1016/j.cub.2020.08.019

Rivalan, M., Munawar, H., Fuchs, A., Winter, Y., 2017. Correction: An Automated, Experimenter-Free Method for the Standardised, Operant Cognitive Testing of Rats. PLoS ONE 12, e0176807. https://doi.org/10.1371/journal.pone.0176807

Rocchi, F., Dylla, M.E., Bohlen, P.A., Ramachandran, R., 2017. Spatial and temporal disparity in signals and maskers affects signal detection in non-human primates. Hearing Research 344, 1–12. https://doi.org/10.1016/j.heares.2016.10.013

Romanski, L.M., Averbeck, B.B., Diltz, M., 2005. Neural Representation of Vocalizations in the Primate Ventrolateral Prefrontal Cortex. Journal of Neurophysiology 93, 734–747. https://doi.org/10.1152/jn.00675.2004

Rosenblum, L.A. (Ed.), 1975. Primate behavior. Volume 4: developments in field and laboratory research. Academic Press, New York.

Schmitt, V., Schloegl, C., Fischer, J., 2014. Seeing the Experimenter Influences the Response to Pointing Cues in Long-Tailed Macaques. PLoS ONE 9, e91348. https://doi.org/10.1371/journal.pone.0091348

Schubiger, M.N., Wüstholz, F.L., Wunder, A., Burkart, J.M., 2015. High emotional reactivity toward an experimenter affects participation, but not performance, in cognitive tests with common marmosets (Callithrix jacchus). Anim Cogn 18, 701–712. https://doi.org/10.1007/s10071-015-0837-5

Schütt, H.H., Harmeling, S., Macke, J.H., Wichmann, F.A., 2016. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. Vision Research 122, 105–123. https://doi.org/10.1016/j.visres.2016.02.002

Šlipogor, V., Graf, C., Massen, J.J.M., Bugnyar, T., 2022. Personality and social environment predict cognitive performance in common marmosets (Callithrix jacchus). Sci Rep 12, 6702. https://doi.org/10.1038/s41598-022-10296-8

Song, X., Osmanski, M.S., Guo, Y., Wang, X., 2016. Complex pitch perception mechanisms are shared by humans and a New World monkey. Proc. Natl. Acad. Sci. U.S.A. 113, 781–786. https://doi.org/10.1073/pnas.1516120113

Sorge, R.E., Martin, L.J., Isbester, K.A., Sotocinal, S.G., Rosen, S., Tuttle, A.H., Wieskopf, J.S., Acland, E.L., Dokova, A., Kadoura, B., Leger, P., Mapplebeck, J.C.S., McPhail, M., Delaney, A., Wigerblad, G., Schumann, A.P., Quinn, T., Frasnelli, J., Svensson, C.I., Sternberg, W.F., Mogil, J.S., 2014a. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat Methods 11, 629–632. https://doi.org/10.1038/nmeth.2935

Sorge, R.E., Martin, L.J., Isbester, K.A., Sotocinal, S.G., Rosen, S., Tuttle, A.H., Wieskopf, J.S., Acland, E.L., Dokova, A., Kadoura, B., Leger, P., Mapplebeck, J.C.S., McPhail, M., Delaney, A., Wigerblad, G., Schumann, A.P., Quinn, T., Frasnelli, J., Svensson, C.I., Sternberg, W.F., Mogil, J.S., 2014b. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat Methods 11, 629–632. https://doi.org/10.1038/nmeth.2935

Stebbins, W.C., 1973. Hearing of old world monkeys (Cercopithecinae). Am. J. Phys. Anthropol. 38, 357–364. https://doi.org/10.1002/ajpa.1330380233

Stevens, K.N., 1983. ACOUSTIC PROPERTIES USED FOR THE IDENTIFICATION OF SPEECH SOUNDS. Ann NY Acad Sci 405, 2–17. https://doi.org/10.1111/j.1749-6632.1983.tb31613.x

Stevenson, M.F., Poole, T.B., 1976. An ethogram of the common marmoset (Calithrix jacchus jacchus): General behavioural repertoire. Animal Behaviour 24, 428–451. https://doi.org/10.1016/S0003-3472(76)80053-X

Takahashi, D.Y., Narayanan, D.Z., Ghazanfar, A.A., 2013. Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. Current Biology 23, 2162–2168. https://doi.org/10.1016/j.cub.2013.09.005

Voelkl, B., Huber, L., 2000. True imitation in marmosets. Animal Behaviour 60, 195–202. https://doi.org/10.1006/anbe.2000.1457

Washburn, D.A., Hopkins, W.D., Rumbaugh, D.M., 1989. Video-task assessment of learning and memory in macaques (Macaca mulatta): effects of stimulus movement on performance. Journal of Experimental Psychology: Animal Behavior Processes 15, 393–400.

Wei, J.J.-F., Tung, T.-H., Li, L.P.-H., 2021. Evolution of impedance values in cochlear implant patients after early switch-on. PLoS ONE 16, e0246545. https://doi.org/10.1371/journal.pone.0246545

Wilk, M., Hessler, R., Mugridge, K., Jolly, C., Fehr, M., Lenarz, T., Scheper, V., 2016. Impedance Changes and Fibrous Tissue Growth after Cochlear Implantation Are Correlated and Can Be Reduced Using a Dexamethasone Eluting Electrode. PLoS ONE 11, e0147552. https://doi.org/10.1371/journal.pone.0147552

Wrobel, C., Dieter, A., Huet, A., Keppeler, D., Duque-Afonso, C.J., Vogl, C., Hoch, G., Jeschke, M., Moser, T., 2018. Optogenetic stimulation of cochlear neurons activates the auditory pathway and restores auditory-driven behavior in deaf adult gerbils. Sci. Transl. Med. 10, eaao0540. https://doi.org/10.1126/scitranslmed.aao0540

Chapter 4

Group-based, autonomous, individualized training and testing of long-tailed macaques (*Macaca fascicularis*) in their home enclosure to a visuo-acoustic discrimination task

Cabrera-Moreno, J., Jeanson, L., Jeschke, M. Calapai, A.

Frontiers in Psychology 13, 1047242 (2022)
DOI: 10.3389/fpsyg.2022.1047242

Chapter 4 describes the results of six long-tailed macaques trained to solve a visuo-acoustic discrimination task using an automated unsupervised training protocol. Where contrary to the marmoset results reported in Chapter 2, long-tailed macaques failed to associate acoustic stimuli with visual targets. Chapter 4 also describes the implementation of a neural network to automate animal identification based on pictures.

Author contributions:

AC, MJ conceived the study. JCM, MJ, AC designed and developed hardware and software. LJ performed the experiments with the help of AC. JCM curated and analyzed the data, and generated the figures. JCM, LJ, MJ AC discussed and interpreted the data. MJ, AC supervised the study. JCM wrote the paper with input from AC, MJ.

Abstract

In recent years the utility and efficiency of automated procedures for cognitive assessment in psychology and neuroscience have been demonstrated in non-human primates (NHP). This approach mimics conventional shaping principles of breaking down a final desired behavior into smaller components that can be trained in a staircase manner. When combined with home-cage based approaches, this could lead to a reduction in human workload, enhancement in data quality, and improvement in animal welfare. However, to our knowledge, there are no reported attempts to develop automated training and testing protocols for long-tailed macaques (*Macaca fascicularis*), a ubiquitous NHP model in neuroscience and pharmaceutical research. In the current work, we present the results from 6 long-tailed macaques that were trained using an automated unsupervised training (AUT) protocol for introducing the animals to the basics of a two-alternative choice (2AC) task where they had to discriminate a conspecific vocalization from a pure tone relying on images presented on a touchscreen to report their response. We found that animals 1) consistently engaged with the device across several months; 2) interacted in bouts of high engagement; 3) alternated peacefully to interact with the device; and 4) smoothly ascended from step to step in the visually-guided section of the procedure, in line with previous results from other NHPs. However, we also found 5) that animals' performance remained at chance level as soon as the acoustically-guided steps were reached; and 6) that the engagement level decreased significantly with decreasing performance during the transition from visual to acoustic guided sections. We conclude that with an autonomous approach, it is possible to train long-tailed macaques in their social group using computer vision techniques and without dietary restriction to solve a visually-guided discrimination task but not an acoustically-guided task. We provide suggestions on what future attempts could take into consideration to instruct acoustically-guided discrimination tasks successfully.

4.1 Introduction

Training non-human primates (NHP) in various husbandry and veterinary procedures is essential to animal behavior management in most captive settings. Positive reinforcement training (PRT) (Skinner, 1938) is the most efficient and ethical technique to train a wide variety of behaviors as it rewards the animals for desired behaviors while ignoring unwanted ones (Westlund, 2015). The standard procedure in PRT training is to break down a desired final behavior into small pieces that can be gradually and sequentially taught to the animal. However, training behaviors required to perform typical experimental tasks in sensory-motor systems research and cognitive neuroscience represents a more significant challenge for classical PRT training. First, most PRT protocols need human trainers to start and end each session and, in some cases, each trial (manual shaping). Besides the time cost - namely that a human trainer can only handle a single animal at a time - there is an unavoidable diversity of training strategies that trainers apply for different animals, ultimately making comparisons across animals and replicability of results challenging (Berger et al., 2018). Finally, in neuroscientific laboratories, NHPs are usually taken from the home cages to insulated experimental setups where they are trained in isolation, potentially reducing the training time and the natural species-specific behavioral repertoire that an animal can express.

Therefore, we would like to argue that the optimization of training protocols has the potential to enhance animal welfare while increasing the standardization of training and ultimately broadening the scope of scientific research. Towards such aims, several studies have already reported various optimization of behavioral training (Berger et al., 2018; Butler and Kennerley, 2019; Calapai et al., 2022, 2017; Sacchetti et al., 2021; Walker et al., 2019) across two important NHP models used in neuroscience, rhesus macaques (*Macaca mulatta*) and common marmosets (*Callithrix jacchus*). However, to the best of our knowledge, there is a lack of reported attempts to develop automated training and testing protocols for long-tailed macaques (*Macaca fascicularis*), a ubiquitous NHP model in neuroscience and - in particular - pharmaceutical research. Long-tailed macaques are 38–55 cm large cercopithecine primates native to Southeast Asia. Animals of this species live in complex social groups - multi-male/multi-female, 6 to 40 individuals - with a dominance hierarchy among females that can be passed through generations of matrilines (van Noordwijk and van Schaik, 1999; Van Noordwijk and Van Schaik, 1985). Due to their close physiological proximity to humans, long-tailed macaques represent a valuable model for biomedical research, especially for basic research studies in disease pathology and treatment, vaccine development, immunology, and neuroscience. Hence, the refinement of protocols to evaluate cognition and behaviors in long-tailed macaques is highly important for phenotyping in treatment development and understanding cognition, affection, and social processes.

In this study, we describe a computerized, automated protocol for training and testing captive long-tailed macaques in their social group. Our approach achieves self-paced, step-wise, individualized training employing picture-based animal identification at the beginning of each trial, which is instrumental in adjusting the training based on the animals' trial-by-trial proficiency. With this approach, no human interaction with the animals is needed, and only minimal maintenance and supervision are required, with presumed positive repercussions on the data quality and the results' replicability. Furthermore, we also argue that removing physical constraints while also keeping the animals in their housing environment with their social group opens the possibility of investigating a broader range of more complex behaviors, including social interactions. Home-cage training also enables the opportunity to record neural activity for extended periods by using wireless recording technologies (Borton et al., 2013; Chestek et al., 2009; Zhou et al., 2019).

Here we report the results from 6 long-tailed macaques navigating an Automated Unsupervised Training (AUT) procedure to reach a visuo-acoustic two-alternative choice (2AC) task. We show that our animals can successfully navigate an AUT procedure to learn a visually guided 2AC on a touchscreen but fail to do the same based on acoustic information.

4.2 Results

In this study 6 female long-tailed macaques housed in two groups, see Table 1, were given access to a touchscreen device attached to their home cage. At the same time, solid food and fluid were provided *ad libitum*. All animals had previous exposure to a similar device during a separate experiment a year prior to this study and were already familiar with the basics of touchscreen interaction. Sessions were mainly autonomously conducted with sporadic human supervision (except for animals R and F trained by an experimenter in 4 and 2 shaping sessions, respectively; see below). Upon the initiation of each trial, throughout the autonomous and unsupervised training procedure, a machine learning algorithm identified the animals from a picture taken by a camera placed on top of the screen (Figure 4.1a). In this way, animals could progress in step-wise training between and within sessions (see methods: Automated unsupervised training (AUT)).

General engagement across sessions

Animals' engagement varied within and between sessions, with a median of 32 trials (IQR=Q3-Q1=90) per session across 50 and 30 sessions per Group 1 and Group 2, respectively (Figure 4.1b). Animals R and F underwent individualized shaping sessions to improve touching accuracy (for 4 and 2 sessions, respectively). The total number of sessions is the number of times the device was offered to the group, regardless of the number of interactions. Except for animal R, which did not perform a single trial across all the sessions, the mean number of sessions with 0 trials per animal is 6 (Figure 4.1c). The session duration ranged from 1.3 to 7 hours with a median of 2 hours and 45 minutes (starting and ending at 10:34 hours and 13:25 hours, respectively - Figure 4.1d). To describe potential habituation effects, we statistically evaluated whether the number of trials per animal varied as a function of session duration or session number and whether the number of trials per hour varied across consecutive sessions. Initial sessions during which solely pictures (see methods) were taken were excluded from this analysis as they were designed to be longer in duration and easier to solve by the animals. We found a significant positive correlation between the number of trials each animal performed and the session duration (partial Pearson correlation, n = 135, r = 0.213, CI95% = 0.05, 0.37, p-value = 0.01) (Figure 4.1f), suggesting that longer sessions lead to more trials. We found no significant correlation between the number of trials performed and the session number (partial Pearson correlation, n = 135, r = 0.00008, CI95% = -0.17, 0.17, p-value = 0.99); as well as between the session duration and the sessions number (partial Pearson correlation, n = 59, r = 0.0088, CI95% = -0.25, 0.27, p-value = 0.94); suggesting that animals did not lose interest in the experiment across consecutive sessions while access to the device remained consistent. Finally, by looking at the distribution of trials across all sessions and all animals, we found that animals mostly engaged during the first 2 hours of the sessions, performing 50% of the trials within the first 53 minutes (Figure 4.1e).

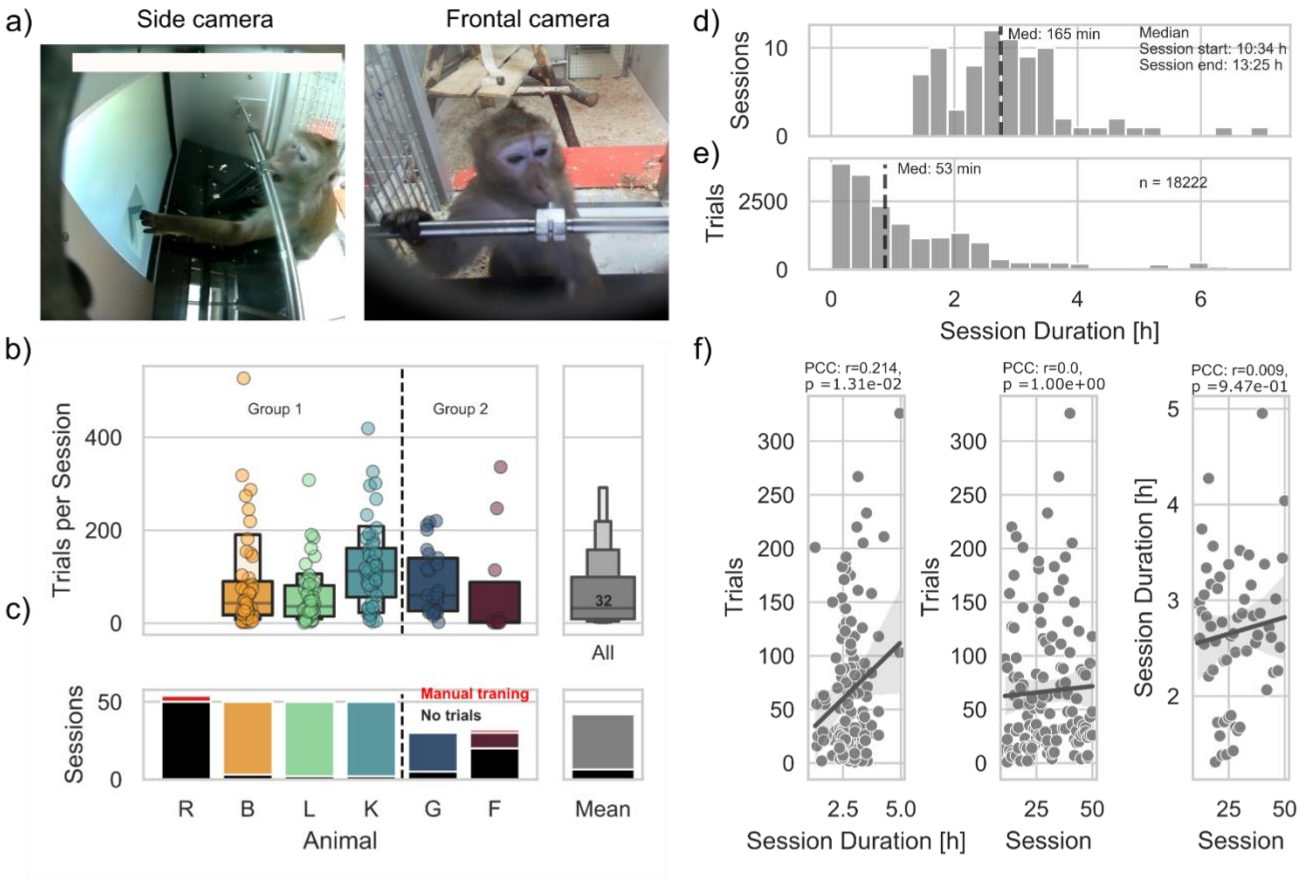


Figure 4.1. General engagement across sessions. a) *Pictures of animal L interacting with the LXBI device. The left picture shows the view from the side camera used for surveillance during sessions. The right picture shows the view from the frontal camera used for animal identification.* ***b)*** *Left panel shows the number of trials per session across animals. The right panel shows the distribution across all animals, with a median of 32 trials per session (IQR = 90 trials).* ***c)*** *Left panel shows the number of sessions across animals. Red indicates the amount of manual training sessions conducted in separation from the rest of the group. Black indicates the amount of sessions with no trials. The right panel shows the mean across animals.* ***d)*** *Distribution of all session durations. The dashed line indicates the median of the distribution.* ***e)*** *Distribution of trial initiation across session duration. The dashed line indicates the median of the distribution.* ***f)*** *From right to left. Distribution of number of trials per animal as a function of session duration, shows a significant positive correlation (partial Pearson's correlation, n = 135, r = 0.213, CI95% = [0.05, 0.37], p = 0.01). Distribution of number of trials per animal as a function of session number, shows non-significant correlation (partial Pearson's correlation, n = 135, r = 0.00008, CI95% = [−0.17, 0.17], p = 0.99). Distribution of session duration as a function of session number shows no significant correlation (partial Pearson's correlation, n = 59, r = 0.0088, CI95% = [−0.25, 0.27], p = 0.94).*

Visuo-acoustic automated unsupervised training (AUT)

In this study, we adapted a visuo-acoustic automated unsupervised training protocol (AUT) we previously used to train marmoset monkeys (Calapai et al., 2022). Here, 5 long-tailed macaques belonging to 2 groups underwent an AUT comprised of 49 training steps. The AUT protocol was designed to 1) improve touch precision (milestone *size*), 2) spatial touch precision and tolerance to acoustic stimuli (milestone *location-sound*), and 3) train a 2 alternative audio-visual association (milestone *distractor*). Training data for animal R are not available as the animal never interacted with the device.

An algorithm that monitored the animals' hit rate within a sliding window of 10 trials loaded the subsequent step when 8 out of 10 trials were correct or the previous step when 2 out of 10 trials were correct, modulating the task difficulty as a function of the animal's performance. Although the design of the AUT aimed to individualize and smoothen the transition between steps according to the animals' learning progress, certain milestones required more trials to be acquired. Therefore, different hit rates can be observed across AUT steps and milestones (Figure 4.2a). An important feature to note is the consistent decrease in performance starting with the last milestone, during which a visual distractor was added. Except for animal F, which did not overcome the milestone size (with 250 trials and 54 sessions), 4 out of 5 animals reached the *distractor* milestone (B, L, K, G) and successfully acquired the visual part of the last milestone. In contrast, none successfully acquired the acoustic part. To visualize the learning progress through the milestones of the AUT across animals with potentially different engagement levels, we quantified the number of trials as a function of the total trials performed (Figure 4.2b). The animals needed an average of 200, 304, and 1141 trials; and 9, 4, and 25 sessions to overcome the *size*, *location-sound*, and *distractor* milestones, respectively (Figure 4.2c). This suggested that two of these milestones (*size* and *position-sound*) were easier to solve than the final milestone (*distractor*), which might have needed a smoother training set of steps than the one used in the current study. To assess whether individual animals' performance influenced subsequent task engagement, we analyzed the likelihood of initiating a trial after a correct or wrong response within the first 30 seconds following a response. We found that the likelihood of starting a trial after a correct response remained stable during the first two milestones (*size* and *location-sound*). In contrast, it consistently changed in the milestone *distractor*, decreasing from 90% to 55%. This pattern is mirrored by the likelihood of initiating another trial after a wrong trial, going from 25% in the initial milestones to 40% in the distractor (Figure 4.2d). The same was observed when controlling for the non-uniform number of trials across steps by recomputing the likelihood based on an equal number but randomly selected sample of trials belonging to all steps (see methods). We found a significant positive correlation between the hit rate and the likelihood of initiating a trial (Figure 4.2e), suggesting that the animals' engagement is heavily dependent on short-term performance as lower hit rates over time tend to promote similar trial initiation for correct and wrong trials.

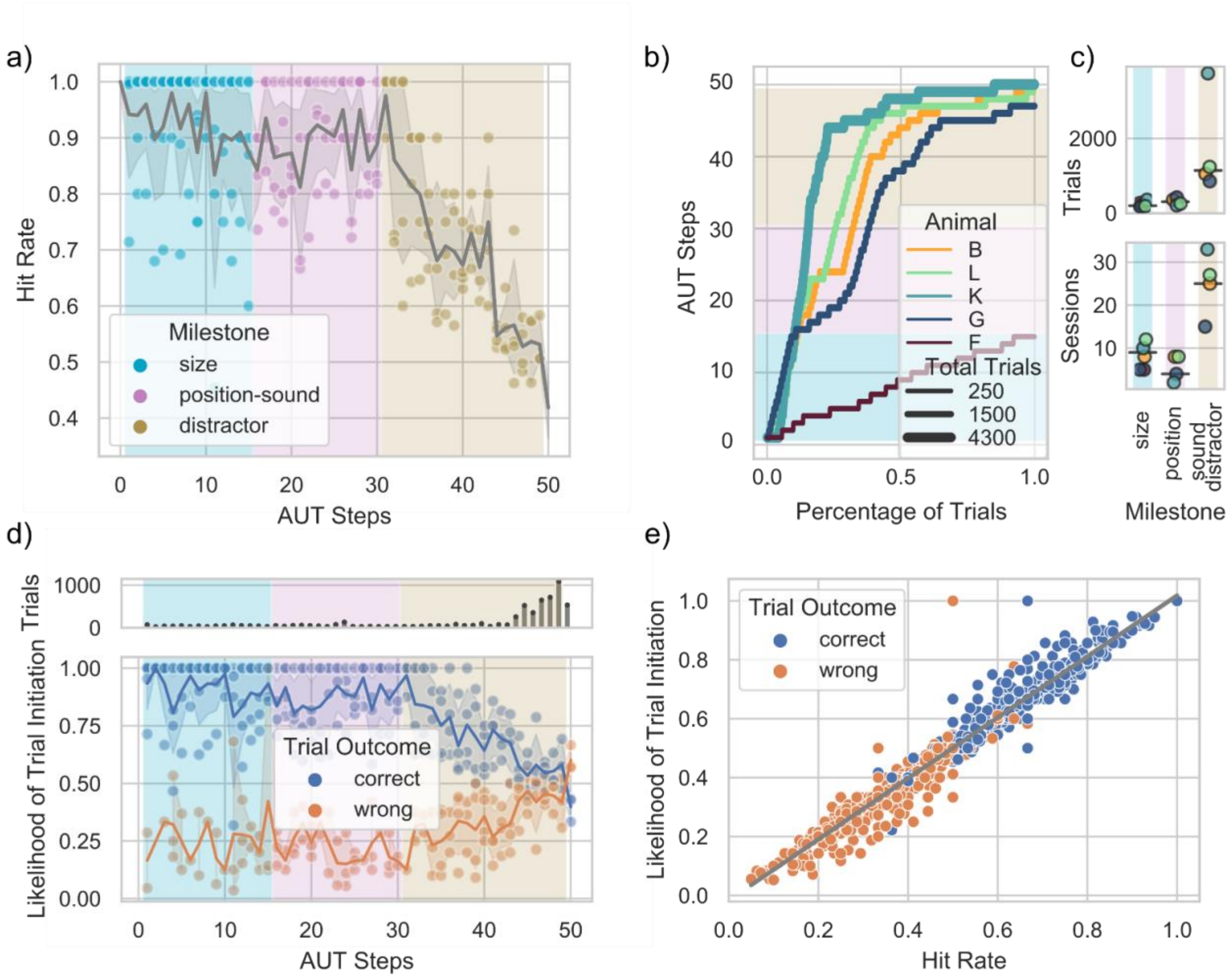


Figure 4.2. Performance through the Amutomated Unsupervised Training (AUT) protocol. a) *Hit rate as a function of AUT steps per animal. Gray shade represents 95% confidence interval of the mean across animals.* ***b)*** *Animal progress through the steps of the AUT protocol. Background colors indicate the milestones.* ***c)*** *From top to bottom, number of trials and number of sessions as a function of milestones across animals.* ***d)*** *Distribution of the likelihood of trial initiation as a function of hit rate in blocks of 100 randomly selected trials across animals. The upper panel shows the number of trials per step.* ***e)*** *Highly significant positive correlation between the hit rate and the likelihood of initiating a trial when controlling for the non-uniform number of trials across steps (partial Pearson's correlation, n = 840, r = 0.98, CI95% = [−0.98, 0.98], p =5e-18).*

2 alternative visuo-acoustic discrimination.

From steps 31 to 49, the AUT protocol attempted to train the animals to discriminate between a target and distractor simultaneously presented on the screen (Figure 4.3a) based on two cues: a visual cue, the difference in the size of the visual stimuli; and an acoustic cue, the specific sound played throughout the trial. Animals could use either cue to determine the target of a given trial. However, from step 50 onward, only the acoustic cue was present as the target and distractor had the same size. While 4 out of 5 animals reached step 50, none had a performance above chance at this stage of the training. This suggests that animals did not use the acoustic cue to identify the target of a given trial but relied exclusively on the stimuli's size difference. A psychometric estimation based on the proportion of correct trials across steps 31 to 50 revealed that the minimum detectable size differences between the target and distractor are: 22.94 cm^2, 25.79 cm^2, 40.87 cm^2, and 37.08 cm^2 for animals B, L, K, and G, respectively (Figure 4.3b). In addition, animals showed a stable hit rate (around the chance level) once the difference between the target and the distractor was around 0.8 cm^2 (step 44) (Figure 4.2c). After step 44, animals B, and G, showed a bias for the vocalization and animal K for the simple train (Figure 4.3d). Also, no significant difference in the response latency between the two stimuli was found (Kruskar-Wallis, Bonferroni corrected (B p-val=0.19, G p-val=0.17, K p-val=0.18, L p-val=0.09).

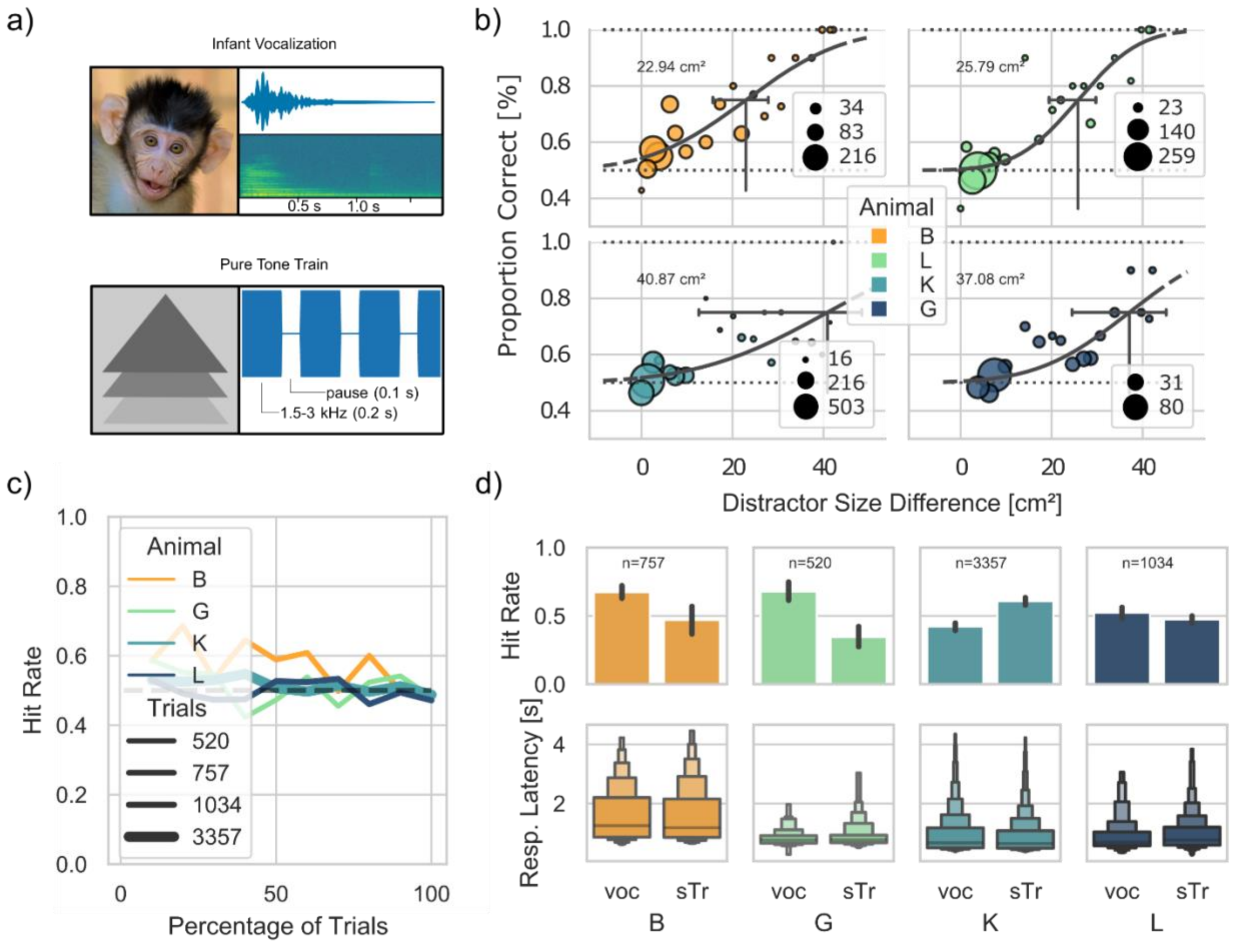


Figure 4.3. Visuo-acoustic discrimination task. a) *Visual and acoustic stimuli used across the AUT (milestones position-sound and distractor) and in step 50. Visual stimuli are shown on the left, and the spectro-temporal information of the acoustic stimuli is shown on the right. The pure tone train was 4 kHz.* ***b)*** *Psychometric curves for the minimum size difference between distractor and target, calculated as the proportions of correct trials across steps of the AUT. 95% confidence intervals (CI) are indicated with black horizontal lines (Animal B: threshold 22.94 cm^2; CI between 15.7 and 26.7; L: 25.79 cm^2, CI between 19.4 and 28.9; K: 40.87 cm^2, CI between 12.6 and 46.4; G: 37.08 cm^2, CI between 24.4 and 43.3).* ***c)*** *Hit rate as a function of the percentage of trials performed by each animal (after step 44 where all animals mostly performed below 60% hit rate), grouped into bins of 10%. The thickness of the lines represents the number of trials. The dashed line at 0.5 represents the chance level.* ***d)*** *Letter-value plots show the reaction times for each stimulus across animals after step 44. The central box represents the 1st quartile, 2nd quartile and 3rd quartile. No statistical difference was found between the response latencies between stimuli at a Bonferroni post-hoc-corrected Kruskal–Wallis Test (B p = 0.19, G p = 0.17, K p = 0.18, L p = 0.09).*

Face identification performance

In order to individualize the automatized training protocol for each animal, we trained a convolutional neural network with a structure optimized for object categorization (LeCun et al., 2015), to identify the animals at the start of each trial. We manually labeled all pictures offline to assess the neuronal network's animal identification performance. We observed stable animal identification performance of the network across consecutive sessions for both groups (Figure 4.4a). The network was retrained after sessions: 5, 29, and 35 for Group 1; and after session 4 for Group 2 (indicated with stars in Figure 4.4a) to expand the training set and potentially prevent drops in identification accuracy. Session 36 of Group 1 was removed from the analysis due to a technical problem with the training of the network (the animals' labels were swapped inadvertently). Figure 4.4b shows that individual animal identification accuracies for Group 1 were around 90%, while for Group 2, animal G held an identification accuracy of almost 100% and animal F of 70%. The accuracy for a given animal was calculated as the number of times labels from the network matched manual labels divided by the total number of network labels for that animal. Furthermore, we computed a more general measure of accuracy for each animal by dividing the number of times labels from the network-matched manual labels by the total amount of manual labels for that given animal. We found this general accuracy above 90% in Group 1 and between 77% and 89% in Group 2 (Figure 4.4c). Finally, to avoid that wrongly identified animal’s influenced a given animals’ progress within the AUT, we took and fed to the network a second picture at the end of each trial before computing the AUT progression. This allowed online identification of trials with different labels from start and end to prevent potential problems with the AUT progression. In addition, this prevented 319 wrong assignments of the trial outcome out of a total of 8784 trials.

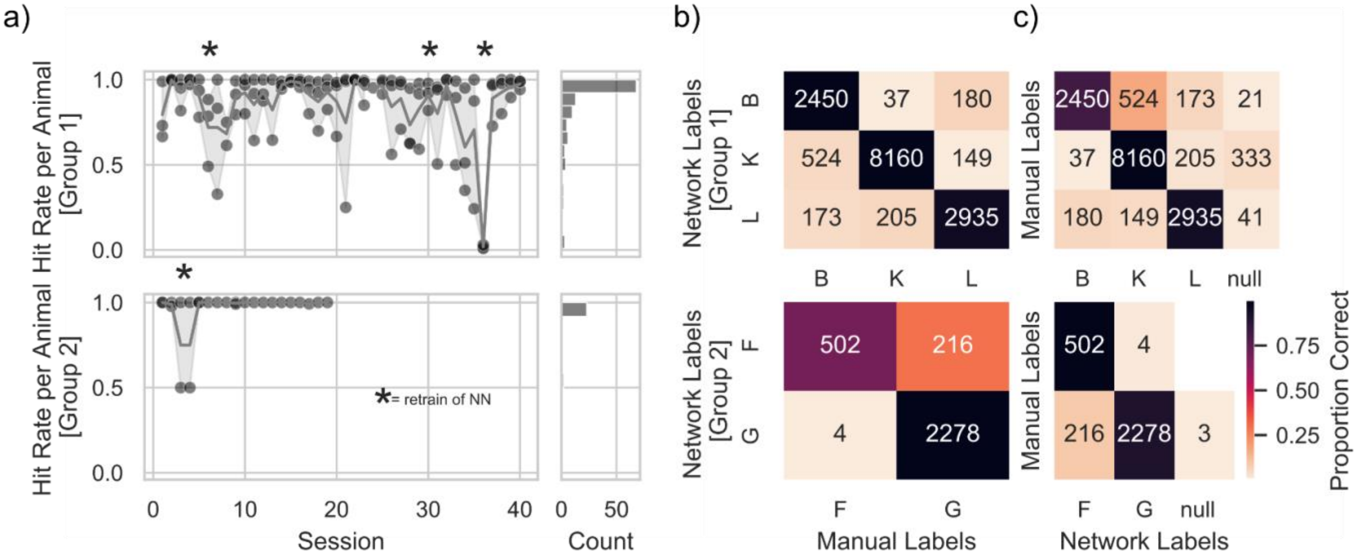


Figure 4.4. Animal identification accuracy. a) *Animal identification accuracy across sessions for Group 1 on the upper right panel and Group 2 on the lower right panel. The right panels show a count histogram for both groups.* ***b)*** *Individual animal identification accuracies were calculated as the number of times labels from the network-matched manual labels, divided by the total number of network labels for that animal. Accuracies for Group 1 were around 90%. At the same time, for Group 2, animal G held an identification accuracy of almost 100% and animal F of 70%.* ***c)*** *General measure of accuracy for each animal, calculated by dividing the number of times labels from the network-matched manual labels by the total amount of manual labels for that given animal. Again, accuracies were above 90% in Group 1 and between 77 and 89% in Group 2. An additional animal label, called null, was assigned to those pictures where the animal's identity was unclear (animals triggering a trial by accident, e.g., with their back). Numbers inside the heatmap represent the number of trials from which the hit rate was calculated.*

Animal turn taking

The online animal identification algorithm, allowing for individualized training and testing of our animals living in social groups, allowed for assessment of the animal-device interaction from a group-level perspective. First, we observed that the level of engagement with the device (taken as the number of interactions as a function of time within a session) is consistently higher at the beginning and lower towards the end (Figures 4.5a). Specifically, in Group 1, we found that within each session, animal B was often the first to interact with the device, followed by animal L, and later by animal K. Moreover, within and across sessions, we observed 463 total transitions from a given animal to a different animal, with a median interval of 101.53 seconds (Figure 4.5b). We found that transitions from L to K occured the most (112), while L to B the least (44), in contrast with other transitions that occurred relatively evenly (B to L and B to K with 63 and 68 transitions, respectively; K to B and K to L with 87 and 88 transitions, respectively). To graphically describe the transition probability among animals we calculated a Markov transition matrix for Group 1 (Figure 4.5c) and statistically assessed whether transitions were the due to random transitions between animals. Towards this, we quantified the probability of obtaining similar results with shuffled data (1000 repetitions) while keeping the same amount of interactions as in the original data. Except for transitions of animal K to L (two-sided permutation test; $p = 0.123$) and K to B (two-sided permutation test; $p = 0.123$), none of the transitions can be explained by chance alone (two-sided permutation test; B to K $p = 0.055$, B to L $p = 0.055$, L to B $p = 0.001$, and L to K $p = 0.001$). These results suggest a preferred turn-taking order with which animals interacted with the LXBI. Such structure cannot be explained by chance, and is likely the product of complex social dynamic interactions within the group. It remains unclear whether the LXBI created such structure or whether the structure existed before and the animals used it as it would naturally happen in naturalistic foraging situations.

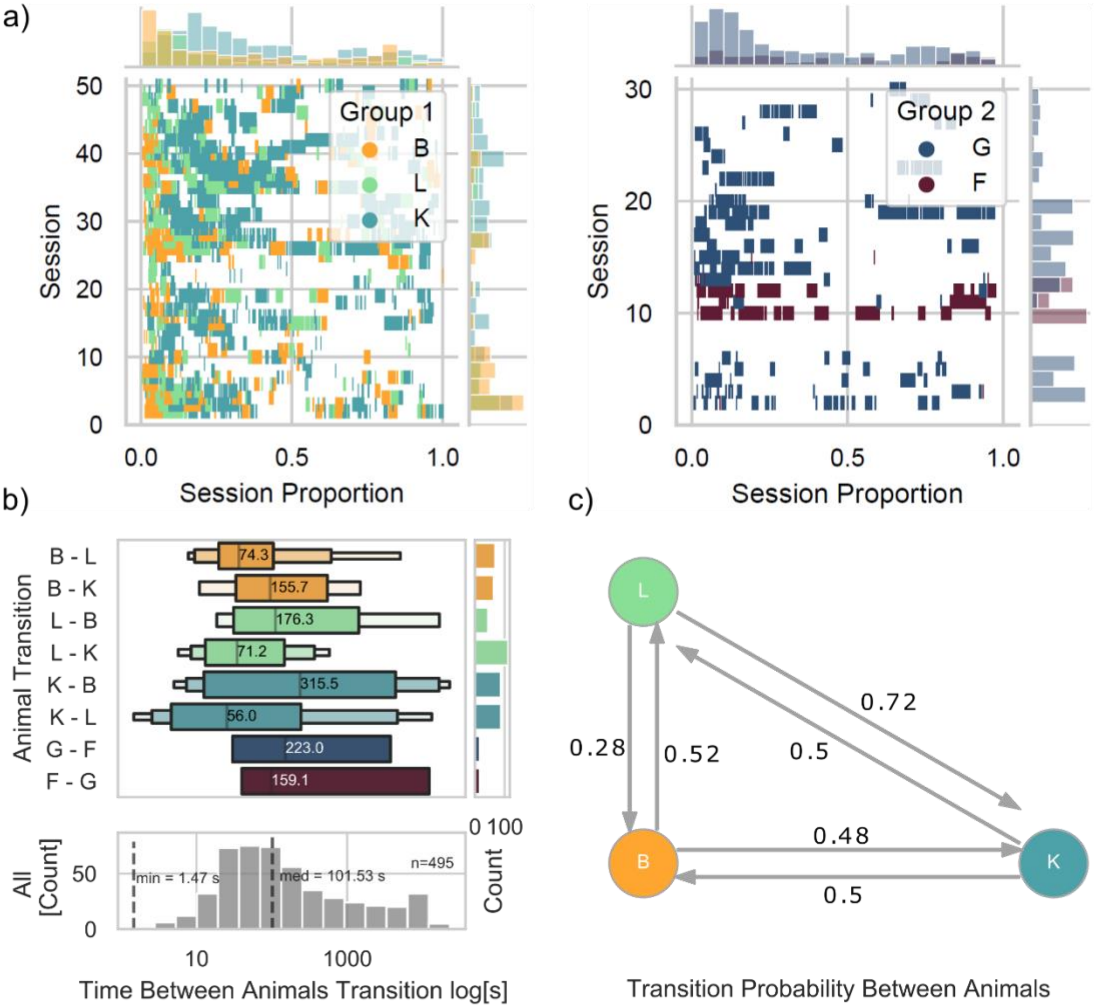


Figure 4.5. Animal turn-taking. a) *Event plot showing each animal's individual trial initiation as a function of session proportion. Left panel for Group 1 and right panel for Group 2. Marginal plots show the density histograms of trial initiation instances across sessions on the ordinate and within sessions on the abscissa.* ***b)*** *The upper panel shows the time distribution between animal transitions across all animals. The lower panel shows the distribution when merging all animals. The marginal plot is a count histogram for the number of transitions.* ***c)*** *Markov transition matrix, showing the probability of transitions among animals.*

4.3 Discussion

Non-human primates (NHP) play an essential role in biomedical research due to their physiological, psychological, and cognitive proximity to humans. However, the requirement to manually train NHPs to understand complex rules and perform complex behaviors bears several caveats. Among them are the inter-experimenter variability of training, the difficulty in generalizing the results, the time and personnel needed, as well as ethical considerations related to the animals' well-being. In an effort to address these issues, we designed a touchscreen-based, autonomous, individualized experimental protocol to train and test long-tailed macaques directly in their home enclosure without fluid/food control or social separation that integrates trial-by-trial animal identification employing a convolutional neural network. Six female long-tailed macaques, across two separate captive groups, underwent daily training sessions (Monday to Friday) for around 3 hours on a touchscreen device attached to their home cage. Our results suggest that: 1) captive long-tailed macaques successfully learn a visually guided discrimination task with autonomous protocols, but demonstrated that more sophisticated approaches than the gradual implicit sound-to-stimulus association we employed are needed for acoustically guided discrimination; 2) animals engage with the device without the necessity of food/fluid control, but such engagement strongly correlates with success rate, as interactions decrease as the task becomes more difficult; 3) picture-base animal identification through machine learning was stable across several months and animals, making it a reliable and non-invasive technique for animal tagging to achieve individualized training without social separation; 4) it is possible to assess group-level dynamics (such as turn-taking) in socially housed non-human primates.

Visuo-acoustic automated training

Our home-cage, automated training protocol was designed based on similar experimental protocols developed for NHPs across the last two decades (Bala et al., 2020; Bullock and Myers, 2009; Butler and Kennerley, 2019; Mandell and Sackett, 2008; Sacchetti et al., 2021; Tulip et al., 2017; Walker et al., 2019; Wither et al., 2020). Specifically, it shared the structural design of the apparatus described for Rhesus macaques (Berger *et al.*, 2018) while replacing costly hardware (Apple computers) with an open-source system (Raspberry Pi computers), allowing for more straightforward modification and expansion of the system by others. It additionally used a visuo-acoustic protocol developed for common marmosets (Calapai et al., 2022). Furthermore, our protocol relies on computer vision technology for the identification of subjects on a trial basis, which could, in principle, allow for testing subjects in natural settings. Our results suggest that long-tailed macaques can be trained in an automated manner to perform basic visually guided tasks using a touchscreen system but failed to generalize to an acoustically guided 2AC task. These findings are in line with previous reports that showed that long-tailed macaques could perform a stimulus-directed touch behavior using a touchscreen system engaging consistently over several sessions (Bullock and Myers, 2009; Rice et al., 2017). However, previous reports have shown that macaques (*Macaca fascicularis*, *Macaca fuscata*, and *Macaca nemistrina*) are indeed able to solve acoustic discrimination tasks (Brosch et al., 2004; Furuyama et al., 2017; Kuhl and Padden, 1983; Petersen et al., 1984). However, differences in stimulus type (human vocalizations versus conspecific vocalizations), setup conditions (attenuated sound chamber versus animal colony, lever versus touchscreen), and testing paradigms (Go No-Go versus 2AC) might account for differences in performance (Waskom et al., 2019), preventing direct comparison across studies. When comparing our results to those reported for common marmosets using a similar system (Calapai et al., 2022) where 9 out of 11 marmosets learned to discriminate conspecific vocalizations from pure tone trains using a 2AC or 3AC paradigm, we found a substantial difference in the engagement of the animals when low hit rates are observed. Our analysis showed that even though the number of trials performed per animal remained relatively constant over sessions (engagement), the likelihood of performing more trials in a row depended on the performance. We argue that this change in engagement dynamics might have contributed to the failure to acquire the visuo-acoustic 2AC from the long-tail macaques because it hindered the necessary exposure time required to learn the discrimination. Regardless of this change, our aim was to elicit an implicit audio-visual association during the later steps of an automated training protocol. Instead, animals ignored or discounted the acoustic information presented and focused exclusively on the visual information (i.e. the difference in stimulus size)**.** Finally, from the necessity to train one long-tailed macaque in the current study and personal communication with the authors that previously exposed the same group of long-tailed macaques to a similar touchscreen device in the context of a different study (Cassidy et al., 2021), we would argue that naïve long-tailed macaques could be trained to interact with our device.

Level of engagement with automatized training protocols

Five out of six animals interacted consistently with the device across several months, in sessions of 3 hour duration during which fluid and food were available *ad libitum*. This was presumably due to the sugary fluid reward delivered by the device in combination with the provision of an activity that provided a form of cognitive enrichment (Calapai et al., 2022; Clark, 2017, 2022; Murphy et al., 2003; Tarou and Bashaw, 2007). We found that engagement strongly depended on short-term performance levels as the likelihood of initiating a trial decreased with an increase in difficulty and throughout the training section in which the task gradually moved from visuo-acoustic to acoustic only. This dependency should be considered for future experiments, especially when generalizations across sensory modalities are needed for experimental purposes. Interestingly, a similar dependence was observed for some individuals at the same stage of the AUT protocol in a previously published marmoset study (Calapai et al., 2022). Finally, while our animals were aged between 10 and 22 years old, considered already "aged" animals (Veenema et al., 2001, 1997), the reported marmosets were significantly younger (2 to 7 years old), marmosets are often referred to as "aged" at 8 years of age (Abbott et al., 2003). Because the cognitive decline in aging NHPs is well demonstrated and particularly relevant for translational neuroscientific research (Gray and Barnes, 2019; Herndon et al., 1997; Lacreuse et al., 2020; Nagahara et al., 2010; Sadoun et al., 2019; Smith, 2004), our approach could be helpful to assess and describe aspects of cognitive decline in captive NHPs in a standardized way.

Animal recognition with machine learning and computer vision

Reliable identification of individuals in socially housed settings and operating device for automated training and cognitive assessment represents a necessity to establish successful high-throughput pipelines (as argued before (Calapai et al., 2022)) and still a significant challenge. A common approach is to employ tracking devices for animals, such as colored jackets, collars, and a combination of video monitoring or electronic devices such as RFID chips; to allow identification (Andrews and Rosenblum, 1994; Ballesta et al., 2014; Calapai et al., 2022; Fagot and Bonté, 2010; Gazes et al., 2013; Maddali et al., 2014; Rose et al., 2012; Tulip et al., 2017). Due to a combination of physiological and technical issues related to implanting and reading RFID chips in large animals such as macaque monkeys (Fagot and Bonté, 2010) we opted for a picture-based identification algorithm that employed a convolutional neuronal network (Butler and Kennerley, 2019; Jacob et al., 2021; Schofield et al., 2019; Witham, 2018). With our network, the classification accuracy for individual animals was in line with the reported accuracy achieved for rhesus macaques using similar methods (Butler and Kennerley, 2019; Witham, 2018) to allow individualized autonomous training. However, we found that running the recognition algorithm twice (at the beginning of each trial) only marginally improved the network performance compared to running the algorithm once per trial. With trials longer than a few seconds (in contrast to our experiment) this strategy could more significantly improve recognition accuracy. We finally suggest that taking a picture from multiple vantage points would improve recognition significantly. Overall, this technique revealed to be reliable in efficiency and easy to implement in a python-based task control. Nonetheless, we believe that further optimizations are needed to establish for example: 1) an unsupervised and automatic updating of the network as well as 2) an internal quality control system to evaluate tagging accuracy. Based on the rapid advancements in machine learning this technique will continue to improve to be suited for non-invasive real time animal classification in social groups.

Insights into turn-taking and social dynamics

It is essential to note that due to the low number of animals and the low engagement of animal F in Group 2, the following analysis will focus mainly on Group 1, and it is intended to be taken as a proof of concept. We observed no fighting or substantial behavioral alteration in our animals throughout the experiment. All animals who interacted with the device across several sessions (5 out of 6) could do so by taking turns. In the early steps of our automated protocol, there were strong differences in the level of engagement across animals, presumably as a result of social dynamics present in a small captive group of primates. Previous reports have shown that the social rank of animals affects their access to resources (Barton and Whiten, 1993; Boogert et al., 2006), with lower-ranked individuals having the least access. Since a trainer (or training device) may be seen as a resource by the animals, engagement in training might be influenced by the social rank of the animals (Prescott and Buchanan-Smith, 2003). However, it has also been suggested that low-ranking individuals performed better at cognitive tasks than higher-ranked individuals when isolated from the rest of the members (Bunnell and Perkins, 1980; Drea and Wallen, 1999; Reader and Laland, 2001), indicating that a failure to learn a specific task in low ranked individuals might be a consequence of personality rather than social ranking (Wergård et al., 2016). While a comparison between individually trained animals when separated from their social group and our group-based training would have helped us to elucidate the difference in performance relative to the social context, we decided not to focus on such comparison as temporary social isolation could have negatively impacted the welfare of the animals. Finally, as a detailed ethological assessment of group hierarchy was not available for our groups, an in-depth comparison with previous studies is not possible. Our analysis revealed a specific non-random structure in the animal turn-taking that was stable across several months. This proof-of-concept represents an encouraging step forward toward the development of efficient and standardized techniques to assess NHPs' social states and dynamics.

In summary, we described a study with 6 captive long-tailed macaques (across two groups) who were given access to a touchscreen device equipped with a step-wise automated training protocol and picture-based, real-time animal identification. Across 3 months of daily 3 hours sessions (Monday to Friday, 10:00 to 13:00), animals successfully learned the basics of a visually guided discrimination task. Still, they failed to generalize to an acoustic-only discrimination task. Furthermore, in structured turns, animals interacted with the device in a self-paced manner, without fluid/food control nor social separation, with the likelihood of initiating a trial getting independent from the trial outcome as the performance drops to chance.

4.4 Methods

All animal procedures of this study were approved by the responsible regional government office [Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES),protocol number: 33.19-42502-04-16/2278] and were in accordance with all applicable German and European regulations on husbandry procedures and conditions.

Animals

Six female long-tailed macaque monkeys housed in two groups were involved in this study (Group 1 with four animals: B, K, L and R and Group 2 with two animals: G and F; see Table 1 for more details about the animals). The animals were group-housed in the facilities of the German Primate Center (DPZ) in Goettingen, Germany, equipped with an enriched environment including a multitude of toys and wooden structures, natural as well as artificial light and exceeding the size requirements of the European regulations, including access to outdoor space. The animals' psychological and veterinary welfare was monitored by the DPZ's staff veterinarians, the animal facility staff, and the lab's scientists, all specialized on working with non-human primates. During the testing sessions animals were fed their regular diet and water *ad libitum*. Training sessions took place mostly in the morning before the feeding time, with a single session taking place in the afternoon. The regular duration of a session was around 2 to 3 hours, where the system was attached to the cage for animals to interact with at their own pace. Animal R (4 sessions) and F (2 sessions) were separated for individual training, while all remaining sessions were conducted with all animals having access to the device as a group.

Apparatus

Data was collected with a custom-made, autonomous, touchscreen device tailored towards macaque monkeys (Calapai et al., 2017) and based on two python-based computers (Raspberry Pi; adapted from (Calapai et al., 2022)).

The device was modified to deliver acoustic stimulation via two speakers located at the upper left and right corners of the device. The Long-tailed eXperimental Behavioral Instrument, in short LXBI (50 x 57 x 30 cm - HxWxD) operates as an unsupervised, standalone, waterproof device that can be attached directly to the home enclosure of the animals via a custom-made frame (Figure 4.1). The device comprises: two Raspberry Pi single board I/O computers (Raspberry Pi 3B+, raspberry.org) to control the experiment and provide real-time video monitoring; a camera module attached to the task controller for animal identification (Raspberry Pi wide-angle camera module RB-Camera-WW Joy-IT); a capacitive touchscreen (15-inches touchscreen, ELO 1537L Securetouch); two peristaltic pumps (Verderflex OEM-Schlauchpumpe M025 DC, 10-30V, 6,5 W) and a custom-made reward tube (placed at 25 cm distance from the screen); and two speakers (Visaton FR58, 8 Ω, 120–20,000 Hz). All components operated at low voltage - between 5 and 12v - at a maximum of 2.5 ampere (touchscreen).

Picture-based animal identification

During AUT experimental sessions, when an animal triggered the start stimulus, a picture was taken from the front camera (left panel of Figure 4.1a); downsampled to a 300x300 pixel, converted to gray values and fed into a custom-made, convolutional neural network optimized for object categorization (inspired by LeCun et al., 2015), to label the picture with one of the animals' identities. A second picture was taken (in later sessions) to increase the robustness of the identification of a given animal. This second picture followed the same processing of the first picture described above.

Structure of the network

The network was designed, trained, and used during the experiment through the *tensorflow* module (Abadi et al., 2016), version 2.0; under Python 3.7. The network consisted of 9 layers in total, from input to output: an Average Pooling input layer (6x3 pooling size); 3 convolutional layers (3x3 kernel, 'relu' activation function, with 64, 16, 32 neurons, respectively); 3 pooling layers (MaxPool 2x2; Dropout; Flatten); 1 Dense layer (with a 'relu' activation function); and a final Dense output layer (with a 'softmax' activation function). The network was compiled with an 'adam' optimizer, a sparse categorical 'crossentropy' function, and 'accuracy' as metrics. The fitting was done in 10 epochs and with a batch size of 32. The output layer, representing the animals in each group, contained an additional neuron, here called *null*, that was trained on pictures triggered by the animals by accident (e.g. with their back). Parameters include the size of the average pooling kernel; the number of neurons in the three convolutional hidden layers; and the number of neurons in the hidden dense layer; they were all bootstrapped beforehand on the platform google colab (https://colab.research.google.com/). Here, with a test dataset of 3000 pictures of two male macaque monkeys taken with the same device and in the same facility, we trained and tested 46 combinations of the parameters mentioned before. Finally, we compared the performances of the 46 resulting networks and handpicked the combination of parameters of the network with the highest accuracy (98.7%). This combination was used as the final configuration for the network used during the experiment.

Training and maintenance of the network

The initial training set was collected in 2 weeks (10 sessions) during the experiment's first phase and consisted of 300 pictures per animal. The network was retrained again after 5, 29, and 32 sessions for Group 1; and 4 sessions for Group 2, to account for possible changes in environmental factors from day to day. Every picture collected in both phases of the experiment was labeled by one of the experimenters, that was already very familiar with the animals, with a custom-made python interface. Labels were used to train and assess the network throughout the study.

Procedure

The following training procedure is an adaptation of a protocol already described for common marmosets (Calapai et al., 2022). However, substantial changes regarding the dimensions and the identity of the stimulus were made. Therefore, the following description aims at highlighting the differences from the marmoset study. In order to run a session, a LXBI device was first attached to the animal's cage and then turned on; leading to automatic starts of a local camera server for remote monitoring and video recording; the mount of a local network server for recursive data logging; and the loading of custom-made graphical user interface that allowed the experimenter to set up the parameters of the session (if needed) and launch the experiment. During this time, the reward (Pineapple, banana, or grape juice at 25% dilution with water) was loaded on the bottles of the reward system and manually pumped along the tubes that led to the mouthpiece (with a custom circuit operated by a momentary switch). Once the device was ready, the experimenter granted access to the device to the animals by removing a panel that divided the LXB from the group's cage. The LXBI was left in the cage while remote surveillance took place every 15 – 20 minutes. At the end of the session, the panel was placed back, allowing the experimenter to open the device from one side (via dedicated hinges) and clean it thoroughly. Pumps were left to run for 30 minutes with hot water to clean debris, and if needed, the device was removed from the animal's cage and stored for the next session.

Sessions

Most of the sessions took place in the morning from 10:00 to 14:30 with two exceptions that extended until the afternoon (10:00 to 17:00). Food was provided at 14:00 by facilities' caretakers, and water was available ad libitum throughout the session. For most sessions, videos of the animals working on the LXBI were recorded.

Manual training sessions

Even though all animals had prior interaction with touchscreen devices, animals R and F underwent individual training sessions after we noticed that they did not adequately interact with the touchscreen. These animals were separated from the group for 4 (animal R) and 2 (animal F) sessions, during which, through PRT shaping techniques, they were manually trained to touch the screen to receive the reward.

Experimental paradigm

Across and during all the sessions, animals never left their home cage. Except for animals R and F, which underwent 4 and 2 manual training sessions, the remaining 4 animals did not require manual training to operate the device. All animals underwent a series of picture-taking sessions (10 sessions) necessary to collect training pictures for the identification network. After this phase, all animals underwent an autonomous, unsupervised training protocol (AUT) comprised of preconfigured training steps to acquire basic proficiency in a standard 2AC discrimination task. It is essential to highlight that all sessions for both groups were gathered with a single LXBI device, and the data collection occurred in successive order, first for Group 1 followed by Group 2.

Picture taking

A total of 10 sessions per group were designed to collect as many pictures from the animals as possible to train the picture-based identification algorithm. A simple behavioral task was built for this purpose. The animals were presented with a white screen, and every time a touch was registered, a picture was taken, labeled, and stored, while the reward pump delivered 1.5 ml of juice during the first session to attract the attention of the animals and then reduced to 0.5ml. The animals had to wait for 2 to 3 seconds between one trial and the next.

Automated unsupervised training (AUT)

After the picture-taking sessions, all animals started an automated training procedure comprised of 50 steps. During the first 49 steps, an algorithm would gradually guide the animals according to their performance (Berger et al., 2018), while in step 50 no change in the task occurred (Calapai et al., 2022). In the AUT section (steps 1-49), animals had to learn the basic rules of a visuo-acoustic 2AC, in which the proficiency of an animal was assessed at step 50. Specific parameters, such as size, location, and time delays, were adjusted during the AUT. Animals could step up when 8 or more out of 10 trials were correct and step down when 2 or less out of 10 were correct. During and across sessions, the progress of every animal was stored and retrieved every time an animal started to perform a trial. In this way, animals could individually navigate the total series of steps resuming after pauses or end of sessions at the same step they left in the last interaction. As mentioned above, the present training protocol is an adaptation from an AUT developed for marmosets (Calapai et al., 2022), with the main differences found in the stimuli's identity and stimuli configuration. The AUT was comprised of 49 steps, grouped into 3 milestones that aim at training long-tailed macaques on the basics of touchscreen interactions within the context of a visuo-acoustic 2AC task. 1) *Size* milestone (steps 2 to 15) aimed to train animals to execute precise touches by decreasing the size of the visual stimulus that triggered the reward. A white circle embedded in a blue rectangle (called *trigger*) placed in the center of the screen had to be touched to obtain the reward (0.5 ml). Throughout the steps, the trigger gradually decreased from 20 x 20 cm to 6 x 6 cm. Touching outside the trigger resulted in a 5 to 7-second long inter-trial interval signaled by a grey screen, during which screen touches were ignored and resulted in a reset of the inter-trial interval. In contrast, touching inside the trigger would allow a new trial initiation after a randomized inter-trial interval of 2 to 4 seconds. 2) *Location-sound* milestone (steps 16 to 30). Here the AUT attempted to train animals to associate a sound with a visual target while also improving the spatial precision touch behavior. Throughout the steps, the trigger appeared at the center of the screen, and upon touch, an acoustic stimulus (either a repeated infant long-tailed macaque vocalization; or a pure tone train of 4 kHz at variable loudness – see below) was presented for 1 to 1.5 s before a visual target appeared, at step 16 (the first of this milestone) the visual target appeared in the center of the screen, but gradually moved away, to either side of the screen, by 1 cm on each step until reaching the edge. In contrast, the side randomly changed from left to right on a trial-by-trial basis. The visual targets consisted of a picture of an infant long-tailed macaque face (6.5 × 6.5 cm), or an abstract geometric pattern (6.5 × 6.5 cm) (Figure 4.2a). The vocalization was matched with the long-tailed macaque face while the pure tone train with the geometric pattern. Along the steps of the *location-sound* milestone, the intensity of the sound was gradually increasing (in steps of 10 dB) from 30 ± 2 dB SPL on step 16 to a final loudness of 72 ± 2 dB SPL on step 22. 3) *Distractor* milestone (steps 31 to 49). Here the AUT trained the animals to discriminate the two visual targets by introducing a second visual target as a distractor with a smaller size than the target. A second visual target (distractor) was displayed together with the first target but on the opposite side of the screen, at the same eccentricity. In the case of a 'vocalization' trial, the visual distractor was the geometric pattern and vice versa. The distractor was gradually increased in size on each step from 0.3 × 0.3 cm in step 31 up to 6.5 × 6.5 cm in step 50 (step:size – 31:0.9 cm^2, 32:1.8 cm^2, 33: 2.56 cm^2, 34: 4.84 cm^2, 35: 8.41 cm^2, 36: 11.55 cm^2, 37:

13.69 cm^2, 38: 15.21 cm^2, 39: 17.64 cm^2, 40: 20.25 cm^2, 41: 22.09 cm^2, 42: 25 cm^2, 43: 28.09 cm^2, 44: 32.49 cm^2, 45: 34.81 cm^2, 46: 36 cm^2, 47: 38.44 cm^2, 48: 39.69 cm^2, 49: 40.96 $cm^{2,}$ 50:42.25 cm^2) at which point it reached the same size as the target. Throughout the protocol, the lack of further interaction within 8 seconds after trial initiation resulted in an aborted trial, and the trial outcome was labeled as 'ignored.' The AUT aimed to instruct the animals in a visuo-acoustic discrimination experiment. They had to distinguish two different sounds and select the corresponding visual stimulus to indicate their choice.

Acoustic-only discrimination task

After having completed the AUT protocol, and therefore having reached step 50, animals were presented with an acoustic-only 2AC task in which they had to discriminate a vocalization from a pure tone train and report their choice by touching the correspondent visual target on the screen. As mentioned earlier, the vocalization was associated with the picture of an infant long-tailed macaque, whereas the pure tone train with a geometric pattern. This association was instructed during the AUT protocol (steps 1 to 49). In contrast to the AUT protocol, in step 50, animals were required to rely solely on acoustic cues to identify the target of a given trial and thus obtain the reward above chance. A trial was counted as correct when an animal could respond to the sound with the correct visual target on the screen and rewarded with 0.5 ml of juice, followed by a 1 to 2 seconds timeout. When the animal chose the wrong visual target, the screen turned grey for 8 seconds, during which interactions with the touchscreen were ignored. Throughout this task, the lack of further interaction within 8 seconds after trial initiation resulted in an aborted trial, and the trial outcome was labeled as 'ignored.'

4.5 Data treatment and Statistics

Data acquisition, formatting, and analysis were performed using Python 3.5.3 and 3.7.7. All figures, means, and medians were calculated using the Python libraries *numpy*, *pandas*, *seaborn*, and *matplotlib*. Double-sided Pearson correlations were calculated using the module *pingouin*. Psychometric functions for obtaining thresholds in size difference were calculated using the module *psignifit* (Schütt et al., 2016) and setting the fit to cumulative normal sigmoid function, with all parameters free and with 95% confidence intervals. This resulted in the following function:

$$\psi(x; m, w, \lambda, \gamma) = \gamma + (1 - \lambda - \gamma)S(x; m, w)$$

Where *m* represents the threshold (the level at 0.5), *w* represents the width (difference between levels 0.5 and 0.95), λ and γ represent the lower and upper asymptote, respectively.

4.6 Acknowledgments

We would like to thank Prof. Susann Boretius of the Functional Imaging Department of the German Primate Center for kindly providing access to the groups of long-tailed macaques tested in this work and Dr. Roger Mundry for providing the shuffling procedure between animal transitions. Further, we want to thank the Cognitive Neuroscience Laboratory (CNL) of the German Primate center for providing a device to be modified in this work.

Funding

This work was partially funded by an Audacity grant of the Leibniz Science Campus' Primate Cognition' to A.C. and M.J.

4.6 References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. https://doi.org/10.48550/ARXIV.1603.04467

Abbott, D.H., Barnett, D.K., Colman, R.J., Yamamoto, M.E., Schultz-Darken, N.J., 2003. Aspects of common marmoset basic biology and life history important for biomedical research. Comp Med 53, 339–350.

Andrews, M.W., Rosenblum, L.A., 1994. Automated recording of individual performance and hand preference during joystick-task acquisition in group-living bonnet macaques (Macaca radiata). Journal of Comparative Psychology 108, 358–362. https://doi.org/10.1037/0735-7036.108.4.358

Bala, P.C., Eisenreich, B.R., Yoo, S.B.M., Hayden, B.Y., Park, H.S., Zimmermann, J., 2020. Automated markerless pose estimation in freely moving macaques with OpenMonkeyStudio. Nat Commun 11, 4560. https://doi.org/10.1038/s41467-020-18441-5

Ballesta, S., Reymond, G., Pozzobon, M., Duhamel, J.-R., 2014. A real-time 3D video tracking system for monitoring primate groups. Journal of Neuroscience Methods 234, 147–152. https://doi.org/10.1016/j.jneumeth.2014.05.022

Barton, R.A., Whiten, A., 1993. Feeding competition among female olive baboons, Papio anubis. Animal Behaviour 46, 777–789. https://doi.org/10.1006/anbe.1993.1255

Berger, M., Calapai, A., Stephan, V., Niessing, M., Burchardt, L., Gail, A., Treue, S., 2018. Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. Journal of Neurophysiology 119, 796–807. https://doi.org/10.1152/jn.00614.2017

Boogert, N.J., Reader, S.M., Laland, K.N., 2006. The relation between social rank, neophobia and individual learning in starlings. Animal Behaviour 72, 1229–1239. https://doi.org/10.1016/j.anbehav.2006.02.021

Borton, D.A., Yin, M., Aceros, J., Nurmikko, A., 2013. An implantable wireless neural interface for recording cortical circuit dynamics in moving primates. J. Neural Eng. 10, 026010. https://doi.org/10.1088/1741-2560/10/2/026010

Brosch, M., Selezneva, E., Bucks, C., Scheich, H., 2004. Macaque monkeys discriminate pitch relationships. Cognition 91, 259–272. https://doi.org/10.1016/j.cognition.2003.09.005

Bullock, C.E., Myers, T.M., 2009. *STIMULUS-FOOD PAIRINGS PRODUCE STIMULUS-DIRECTED TOUCH-SCREEN RESPONDING IN CYNOMOLGUS MONKEYS* (MACACA FASCICULARIS) *WITH OR WITHOUT A POSITIVE RESPONSE CONTINGENCY*. Journal of the Experimental Analysis of Behavior 92, 41–55. https://doi.org/10.1901/jeab.2009.92-41

Bunnell, B.N., Perkins, M.N., 1980. Performance correlates of social behavior and organization: Social rank and complex problem solving in crab-eating macaques (M. fascicularis). Primates 21, 515–523. https://doi.org/10.1007/BF02373840

Butler, J.L., Kennerley, S.W., 2019. Mymou: A low-cost, wireless touchscreen system for automated training of nonhuman primates. Behav Res 51, 2559–2572. https://doi.org/10.3758/s13428-018-1109-5

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., Gail, A., 2017. A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. Behav Res 49, 35–45. https://doi.org/10.3758/s13428-016-0707-3

Calapai, A., Cabrera-Moreno, J., Moser, T., Jeschke, M., 2022. Flexible auditory training, psychophysics, and enrichment of common marmosets with an automated, touchscreen-based system. Nat Commun 13, 1648. https://doi.org/10.1038/s41467-022-29185-9

Cassidy, L.C., Bethell, E.J., Brockhausen, R.R., Boretius, S., Treue, S., Pfefferle, D., 2021. The Dot-Probe Attention Bias Task as a Method to Assess Psychological Well-Being after Anesthesia: A Study with Adult Female Long-Tailed Macaques (***Macaca fascicularis***). Eur Surg Res. https://doi.org/10.1159/000521440

Chestek, C.A., Gilja, V., Nuyujukian, P., Kier, R.J., Solzbacher, F., Ryu, S.I., Harrison, R.R., Shenoy, K.V., 2009. HermesC: Low-Power Wireless Neural Recording System for Freely Moving Primates. IEEE Trans. Neural Syst. Rehabil. Eng. 17, 330–338. https://doi.org/10.1109/TNSRE.2009.2023293

Clark, F., 2017. Cognitive enrichment and welfare: Current approaches and future directions. AB&C 4, 52–71. https://doi.org/10.12966/abc.05.02.2017

Clark, F.E., 2022. Bridging pure cognitive research and cognitive enrichment. Anim Cogn. https://doi.org/10.1007/s10071-022-01636-2

Drea, C.M., Wallen, K., 1999. Low-status monkeys “play dumb” when learning in mixed social groups. Proc. Natl. Acad. Sci. U.S.A. 96, 12965–12969. https://doi.org/10.1073/pnas.96.22.12965

Fagot, J., Bonté, E., 2010. Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio papio). Behavior Research Methods 42, 507–516. https://doi.org/10.3758/BRM.42.2.507

Furuyama, T., Kobayasi, K.I., Riquimaroux, H., 2017. Acoustic characteristics used by Japanese macaques for individual discrimination. Journal of Experimental Biology jeb.154765. https://doi.org/10.1242/jeb.154765

Gazes, R.P., Brown, E.K., Basile, B.M., Hampton, R.R., 2013. Automated cognitive testing of monkeys in social groups yields results comparable to individual laboratory-based testing. Anim Cogn 16, 445–458. https://doi.org/10.1007/s10071-012-0585-8

Gray, D.T., Barnes, C.A., 2019. Experiments in macaque monkeys provide critical insights into age-associated changes in cognitive and sensory function. Proc. Natl. Acad. Sci. U.S.A. 116, 26247–26254. https://doi.org/10.1073/pnas.1902279116

Herndon, J.G., Moss, M.B., Rosene, D.L., Killiany, R.J., 1997. Patterns of cognitive decline in aged rhesus monkeys. Behavioural Brain Research 87, 25–34. https://doi.org/10.1016/S0166-4328(96)02256-5

Jacob, G., Katti, H., Cherian, T., Das, J., Zhivago, K., Arun, S., 2021. A naturalistic environment to study visual cognition in unrestrained monkeys. eLife 10, e63816. https://doi.org/10.7554/eLife.63816

Kuhl, P.K., Padden, D.M., 1983. Enhanced discriminability at the phonetic boundaries for the place feature in macaques. The Journal of the Acoustical Society of America 73, 1003–1010. https://doi.org/10.1121/1.389148

Lacreuse, A., Raz, N., Schmidtke, D., Hopkins, W.D., Herndon, J.G., 2020. Age-related decline in executive function as a hallmark of cognitive ageing in primates: an overview of cognitive and neurobiological studies. Phil. Trans. R. Soc. B 375, 20190618. https://doi.org/10.1098/rstb.2019.0618

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Maddali, H.T., Novitzky, M., Hrolenok, B., Walker, D., Balch, T., Wallen, K., 2014. Inferring Social Structure and Dominance Relationships Between Rhesus macaques using RFID Tracking Data. https://doi.org/10.48550/ARXIV.1407.0330

Mandell, D.J., Sackett, G.P., 2008. A computer touch screen system and training procedure for use with primate infants: Results from pigtail monkeys (Macaca nemestrina). Dev. Psychobiol. 50, 160–170. https://doi.org/10.1002/dev.20251

Murphy, E.S., McSweeney, F.K., Smith, R.G., McComas, J.J., 2003. DYNAMIC CHANGES IN REINFORCER EFFECTIVENESS: THEORETICAL, METHODOLOGICAL, AND PRACTICAL IMPLICATIONS FOR APPLIED RESEARCH. Journal of Applied Behavior Analysis 36, 421–438. https://doi.org/10.1901/jaba.2003.36-421

Nagahara, A.H., Bernot, T., Tuszynski, M.H., 2010. Age-related cognitive deficits in rhesus monkeys mirror human deficits on an automated test battery. Neurobiology of Aging 31, 1020–1031. https://doi.org/10.1016/j.neurobiolaging.2008.07.007

Petersen, M.R., Beecher, M.D., Zoloth, S.R., Green, S., Marler, P.R., Moody, D.B., Stebbins, W.C., 1984. Neural lateralization of vocalizations by Japanese macaques: Communicative significance is more important than acoustic structure. Behavioral Neuroscience 98, 779–790. https://doi.org/10.1037/0735-7044.98.5.779

Prescott, M.J., Buchanan-Smith, H.M., 2003. Training Nonhuman Primates Using Positive Reinforcement Techniques. Journal of Applied Animal Welfare Science 6, 157–161. https://doi.org/10.1207/S15327604JAWS0603_01

Reader, S.M., Laland, K.N., 2001. Primate Innovation: Sex, Age ans SOcial Rank Differences. International Journal of Primatology 22, 787–805. https://doi.org/10.1023/A:1012069500899

Rice, N.C., Makar, J.R., Myers, T.M., 2017. Sex and the stimulus-movement effect: Differences in acquisition of autoshaped responding in cynomolgus monkeys. Physiology & Behavior 171, 40–49. https://doi.org/10.1016/j.physbeh.2016.12.028

Rose, C., de Heer, R.C., Korte, S., van der Harst, J.E., Weinbauer, G.F., Spruijt, B.M., 2012. Quantified tracking and monitoring of diazepam treated socially housed cynomolgus monkeys. Regulatory Toxicology and Pharmacology 62, 292–301. https://doi.org/10.1016/j.yrtph.2011.10.008

Sacchetti, S., Ceccarelli, F., Ferrucci, L., Benozzo, D., Brunamonti, E., Nougaret, S., Genovesio, A., 2021. Macaque monkeys learn and perform a non-match-to-goal task using an automated home cage training procedure. Sci Rep 11, 2700. https://doi.org/10.1038/s41598-021-82021-w

Sadoun, A., Rosito, M., Fonta, C., Girard, P., 2019. Key periods of cognitive decline in a nonhuman primate model of cognitive aging, the common marmoset (Callithrix jacchus). Neurobiology of Aging 74, 1–14. https://doi.org/10.1016/j.neurobiolaging.2018.10.003

Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., Carvalho, S., 2019. Chimpanzee face recognition from videos in the wild using deep learning. Sci. Adv. 5, eaaw0736. https://doi.org/10.1126/sciadv.aaw0736

Schütt, H.H., Harmeling, S., Macke, J.H., Wichmann, F.A., 2016. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. Vision Research 122, 105–123. https://doi.org/10.1016/j.visres.2016.02.002

Skinner, B.F., 1938. The behavior of organisms: an experimental analysis, The behavior of organisms: an experimental analysis. Appleton-Century, Oxford, England.

Smith, D.E., 2004. Memory Impairment in Aged Primates Is Associated with Focal Death of Cortical Neurons and Atrophy of Subcortical Neurons. Journal of Neuroscience 24, 4373–4381. https://doi.org/10.1523/JNEUROSCI.4289-03.2004

Tarou, L.R., Bashaw, M.J., 2007. Maximizing the effectiveness of environmental enrichment: Suggestions from the experimental analysis of behavior. Applied Animal Behaviour Science 102, 189–204. https://doi.org/10.1016/j.applanim.2006.05.026

Tulip, J., Zimmermann, J.B., Farningham, D., Jackson, A., 2017. An automated system for positive reinforcement training of group-housed macaque monkeys at breeding and research facilities. Journal of Neuroscience Methods 285, 6–18. https://doi.org/10.1016/j.jneumeth.2017.04.015

van Noordwijk, M.A., van Schaik, C.P., 1999. The effects of dominance rank and group size on female lifetime reproductive success in wild long-tailed macaques,Macaca fascicularis. Primates 40, 105–130. https://doi.org/10.1007/BF02557705

Van Noordwijk, M.A., Van Schaik, C.P., 1985. Male migration and rank acquisition in wild long-tailed macaques (Macaca fascicularis). Animal Behaviour 33, 849–861. https://doi.org/10.1016/S0003-3472(85)80019-1

Veenema, H.C., Spruijt, B.M., Gispen, W.H., van Hooff, J.A.R.A.M., 1997. Aging, Dominance History, and Social Behavior in Java-Monkeys (Macaca fascicularis). Neurobiology of Aging 18, 509–515. https://doi.org/10.1016/S0197-4580(97)00107-3

Veenema, H.C., van Hooff, J.A.R.A.M., Gispen, W.H., Spruijt, B.M., 2001. Increased rigidity with age in social behavior of Java-monkeys (Macaca fascicularis). Neurobiology of Aging 22, 273–281. https://doi.org/10.1016/S0197-4580(00)00204-9

Walker, J.D., Pirschel, F., Gidmark, N., MacLean, J.N., Hatsopoulos, N.G., 2019. A platform for semi-automated voluntary training of common marmosets for behavioral neuroscience: Voluntary training of common marmosets (preprint). Neuroscience. https://doi.org/10.1101/635334

Waskom, M.L., Okazawa, G., Kiani, R., 2019. Designing and Interpreting Psychophysical Investigations of Cognition. Neuron 104, 100–112. https://doi.org/10.1016/j.neuron.2019.09.016

Wergård, E.-M., Westlund, K., Spångberg, M., Fredlund, H., Forkman, B., 2016. Training success in group-housed long-tailed macaques (Macaca fascicularis) is better explained by personality than by social rank. Applied Animal Behaviour Science 177, 52–58. https://doi.org/10.1016/j.applanim.2016.01.017

Westlund, K., 2015. Training laboratory primates &ndash; benefits and techniques. Primate Biol. 2, 119–132. https://doi.org/10.5194/pb-2-119-2015

Witham, C.L., 2018. Automated face recognition of rhesus macaques. Journal of Neuroscience Methods 300, 157–165. https://doi.org/10.1016/j.jneumeth.2017.07.020

Wither, R.G., Boehnke, S.E., Lablans, A., Armitage-Brown, B., Munoz, D.P., 2020. Behavioral shaping of rhesus macaques using the Cambridge neuropsychological automated testing battery. Journal of Neuroscience Methods 342, 108803. https://doi.org/10.1016/j.jneumeth.2020.108803

Zhou, A., Santacruz, S.R., Johnson, B.C., Alexandrov, G., Moin, A., Burghardt, F.L., Rabaey, J.M., Carmena, J.M., Muller, R., 2019. A wireless and artefact-free 128-channel neuromodulation device for closed-loop stimulation and recording in non-human primates. Nat Biomed Eng 3, 15–26. https://doi.org/10.1038/s41551-018-0323-x

Chapter 5

General Discussion

5.1 Summary and interpretation

The underlying aim of this dissertation was to contribute to the development of more efficient methods of assessing auditory cognition in non-human primates (NHPs).

Chapter 2 describes a novel protocol that demonstrates the feasibility of automatically training and testing common marmosets (*Callithrix jacchus*) in auditory tasks directly in their homecages without water and food control or social separation. In its first section, Chapter 2 details the development of the novel instrument named MXBI, a standalone device mostly comprised of off-the-shelf or 3d printed components, entirely programmed in Python, and based on Raspberry Pi computers, for maximum flexibility of use, accessibility, and to allow for easy adaptation by others without comprising the systems' performance. In a later section, Chapter 2 describes an automated unsupervised training protocol that ultimately instructed marmosets on a visuo-acoustic discrimination task. In this task, marmosets learned to group pairs of visual and acoustic stimuli (an infant marmoset vocalization with an infant picture of a marmoset, a pure tone train with a geometric pattern) to then discriminate between the pairs guided by acoustic signals. Finally, Chapter 2 describes a series of experiments demonstrating that marmosets flexibly discriminate novel sets of visual and acoustic stimuli and that hearing thresholds can be gathered in potentially distracting auditory conditions.

Nonetheless, Chapter 2 most significant caveat was the failure of training marmosets in acoustic tasks without relying on conspecific vocalizations. Since traditional assessments in psychoacoustics rely on simpler acoustic stimuli like pure tones, devising tasks to tackle this limitation was of paramount importance.

Chapter 3 details a series of experiments that aimed to expand the systems' flexibility of use by developing novel tasks to assess different aspects of auditory cognition. The first section describes an automated training protocol that successfully instructed common marmosets into a pure tone detection task later used for collecting audiograms in five marmosets. Thresholds across the tested frequencies exhibited similarities to previously reported data (Osmanski and Wang, 2011), with a slight drop in sensitivity of 10 dB SPL across tested frequencies, presumably a consequence of background noise. Chapter 3 later describes the implementation of wireless control of cochlear implants in freely moving marmosets. Here, two cochlear-implanted marmosets flexibly used electrical stimulation to solve a detection task that later helped to collect electrical stimulation thresholds (data still under collection). Finally, the last section of Chapter 3 details the development and implementation of a vocalization discrimination task, where four marmosets learned to discriminate, generalize, and sort in two groups over 900 synthesized acoustic stimuli with several spectro-temporal variations. Showing that manipulations of the center frequency and harmonic content have the greatest disruption in the stimulus discrimination responses.

Finally, Chapter 4 describes the adaptation and implementation of the previously described training protocol used to instruct marmosets into visuo-acoustic discrimination in a novel species, the long-tailed macaque (*Macaca fascicularis*). However, despite the success in marmosets, none of the tested long-tailed macaques learned to solve the visuo-acoustic discrimination task, remaining at a chance level as soon as the acoustically guided steps were reached. Interestingly, although the animals' engagement remained stable across sessions, their trial initiation dynamics changed over time. Chapter 4 also describes the implementation of a picture-based animal identification software that enabled our system to automatically identify individual animals with high accuracy (over 90% hit rate) to provide personalized training and track interaction dynamics in a social context.

The technological advances throughout the history of behavioral sciences have been fundamental for the optimization of experimental procedures. The automating of the behavioral analysis translates to more animal data with substantially greater reproducibility. Over the last three decades, a new wave of technological development has strengthened the bridging of behavioral sciences research. Psychology, ethology, and neurobiology are currently facing the promise of a new framework for studying behavior and its mechanisms more holistically. However, the methodological advances achieved by this new wave of technological influx are still under current development thus, limited in application to a few fields.

Auditory cognition in NHPs still faces substantial limitations in its experimental methods. Although the current behavioral methodology has barely changed over the last decades (e.g., the use of primate chair in auditory tasks can be traced back 50 years (Stebbins, 1973), and continues to be a common method (Mackey et al., 2022)), they have been crucial for understanding essential mechanisms of primate auditory cognition. These mechanisms include sound perception, frequency resolution, pitch perception, sound localization, stream segregation, vocal perception, control and development, and attention (Brosch et al., 2004, 2004; Brown et al., 1982; Geissmann, 2002; Hage et al., 2016, 2013; Kayser et al., 2005; Knyazeva et al., 2018; Pisanski et al., 2016; Pomberger et al., 2019; Remington et al., 2022; Roy et al., 2016; Song et al., 2016; Takahashi et al., 2015, 2013; Wakita, 2020, 2019).

The findings described in this dissertation represent a substantial contribution to the optimization of experimental protocols for studying auditory cognition in NHPs in more naturalistic environments by demonstrating that:

1. Auditory training and testing can be automated for common marmosets.
2. Marmosets can flexibly learn a broad variety of auditory tasks relying on a wide range of acoustic stimuli, including electric stimulation through a cochlear implant.
3. Marmosets can perform complex auditory tasks in a naturalistic environment with inherent acoustic and social distractors.
4. The auditory data collected for marmosets under the aforementioned circumstances is comparable to those already published.
5. Marmosets and long-tailed macaques show stable, voluntary, and sufficient engagement for evaluating cognitive capabilities free of coercion.
6. The training device shows flexible adaptability between experimental species and testing conditions without losing performance.

7. With high accuracy, computer vision technology enables non-invasive, markerless individual animal identification of socially housed long-tailed macaques.
8. Social structures in the colony might underlie turn-taking dynamics using our device.
9. The engagement of the animals in the training protocols might represent a form of cognitive enrichment improving animal welfare.

The common marmosets' ability to flexibly learn and utilize a diverse range of acoustic stimuli, including pure tones, complex synthetic sounds, conspecific vocalizations, and white noise, has been demonstrated. Additionally, they can employ electric stimulation patterns through a cochlear implant to successfully navigate Go-NoGo and 2 or 3-alternative choice tasks. This proficiency, combined with their use of visual cues on a touchscreen, provides a foundation for the advancement of automated cognitive testing protocols for non-human primates. One immediate possibility is the automation of already existing psychoacoustic tests such as discrimination of artificial sounds needed for stream segregation, frequency and intensity resolution, auditory scene analysis, vocal perception, masking effects, or sound localization that have largely remained manually operated (Brown et al., 1982; Izumi, 1999; Osmanski et al., 2016; Remington et al., 2022; Wienicke et al., 2001). Considering that cognitive testing in social settings provides a closer estimate of real-world cognitive capabilities, a second possibility is the development of novel protocols to assess different aspects of physical and social cognition. Auditory attention mechanisms are suitable for assessment under more naturalistic conditions, considering that attention can be voluntarily directed when performing a task (top-down) or captured by salient stimuli in the background (bottom-up) (Ahveninen et al., 2006; Fritz et al., 2007; Johnson and Zatorre, 2006; Knudsen, 2007). An evaluation of the influence of background sounds on the animals' performance might provide relevant information regarding selective hearing during voluntary directed attention, not yet evaluated in detail in NHPs.

The demonstration of feasible wireless control over cochlear implants in freely moving marmosets enables the further investigation of behavioral responses to more complex electric stimulation patterns to evaluate the behavioral and physiological responses to specific components of the electrical stimulation, which in turn may help to develop better sound coding strategies. Additionally, the relatively long duration of the testing sessions could provide an opportunity to test for plasticity when cochlear implants are used for longer periods, providing insights into neuronal plasticity along the auditory pathway and the animals' learning and adaptation processes to electric signals. Furthermore, the development of behavioral methods to assess auditory cognition under cochlear implant circumstances lays the ground for future behavioral comparison of hearing restoration by means of a novel cochlear implant technology based on light stimulation of genetically modified spiral ganglion neurons (Jeschke and Moser, 2015).

Multisensory integration is crucial to adaptive behavior because it enables the coherent perception of objects, allowing individuals to increase salience created by multisensory redundancy (Lewkowicz and Ghazanfar, 2009; Lewkowicz and Kraebel, 2004). A common area of study is the interaction between sensory modalities and their reciprocal processing modulation (Kuang and Zhang, 2014). Demonstrating that marmosets successfully learned to use specific audio-visual cues (with presumed biological significance) to solve a discrimination task but failed to solve the same task relying on a different set of stimuli (synthetic stimuli) offers the possibility to investigate the effects of visuo-acoustic stimuli characteristics interactions on discrimination success.

Across all experiments described in this dissertation, psychophysical approaches allowed for characterizing perceptual and cognitive processes by quantitatively associating the physical characteristics of acoustic stimuli with their perceptual effects through the animals' more frequent choices (Fechner, 1860; Kepecs et al., 2008; Sutton and Barto, 2018). Although this approach offers the possibility to guide, isolate and manipulate behavior to systematically investigate its mechanisms, it also restricts the behavioral assessment to discrete units, ignoring concomitant behaviors. In recent years, machine learning algorithms have enabled the automatic detection, identification, and tracking of individual subjects, specific body part movements, or multiple individual interactions through video recordings (Graving et al., 2019; Lauer et al., 2022; Mathis et al., 2018), they have also enabled the automate the identification and classification of NHPs vocalizations (Desai et al., 2018; Mielke and Zuberbühler, 2013; Phaniraj et al., 2022; Romero-Mujalli et al., 2021; Turesson et al., 2016; Zhang et al., 2018). Implementing this technology with our approaches promises the expansion in the dimensionality of the behavioral assay, opening novel opportunities to put animals' behavior during cognitive tasks into context.

Decision-making mechanisms likely depend on continuously evolving perceptual or cognitive variables. Access to video recordings of continuously expressed behavior during task performance could provide information about the time course of unfolding decisions, such as the micro-movements of the body (Gold and Shadlen, 2001; Gouvêa et al., 2014; Selen et al., 2012).

NHP are capable of recognizing the caller's identity and relative relatedness (Biben, 1993; Miller et al., 2010, 2004, 2001a, 2001b; Miller and Wang, 2006). The potential of automatically identifying animals' identities through their vocalizations offers the opportunity to evaluate in detail the influence of the relativeness of background calls during task performance. This is particularly relevant for marmosets, a highly vocal species that remains loquacious in captivity, which has been demonstrated to show behavioral preference in vocalization exchange towards relatives (Miller and Wang, 2006).

Tracking animal interactions in group settings can potentially mitigate the logistical challenges of social testing by evaluating the animals' performance with respect to the social context. Additionally, learning mechanisms influenced by conspecifics (social learning), cooperation, deception, and prosociality studies offer a novel dimension of study, and broaden the scope of the results (Canteloup et al., 2020; de Waal and Suchak, 2010).

Although machine learning algorithms have been deployed under several experimental circumstances to automate the analysis of visual and acoustic data, challenges in high-resolution camera technology, multi-animal tracking in 3D settings, online analysis, signal-to-noise ratio robustness with acoustic data, and a broad cover of species vocalization limit the application of this technology to particular settings (Desai et al., 2018; Lauer et al., 2022; Phaniraj et al., 2022; Turesson et al., 2016)

The implementation of neural recordings in NHPs while performing cognitive tasks is a common practice in cognitive neuroscience to investigate the neural substrates of cognitive processes. Although neuronal coding mechanisms governing some auditory processess such as frequency tunning, sound intensity, pitch perception, sound location, vocal processing and sensory motor integration has been widely studied in NHPs (Aitkin et al., 1986; Bendor et al., 2012; Bendor and Wang, 2005; Brosch et al., 2004; Cheung et al., 2005; Eliades and Wang, 2008a, 2008b; Lui et al., 2015; Miller et al., 2015; Remington et al., 2022; Remington and Wang, 2019; Sadagopan and Wang, 2008; Watkins and Barbour, 2011), the possibility of recording neural activity in voluntarily motivated animals while performing auditory cognitive tasks under more naturalistic environments could provide additional data to perform context-dependent comparisons that might help to differentiate indirect from direct neural-behavioral correlations, specially since evidence shows that the physiological properties of neural circuits are context dependent, suggesting the activation of alternative neural substrates for the similar actions (Balleine, 1992; Carmena et al., 2003; Mante et al., 2013). Furthermore, testing under social context might be particularly relevant for investigating auditory processes that involve higher-level, top-down cortical influences, such as attention mechanisms (Fishman et al., 2017; Knyazeva et al., 2018; Schneider et al., 2014). Wireless technology in recent years has allowed researchers to record neural activity in freely moving and behaving animals (Harrison et al., 2011; Lee et al., 2013; Rizk et al., 2009; Sato et al., 2009; Szuts et al., 2011), including NHPs (Borton et al., 2013; Foster et al., 2014; Jackson et al., 2007; Konoike et al., 2022; Miranda et al., 2010; Mohseni et al., 2005; Roy et al., 2016; Roy and Wang, 2012; Schwarz et al., 2014; Walker et al., 2021; Yin et al., 2014). However, the amount and distance of data transmitted, the size and weight of the head stage, the implant longevity, and battery life are some of the challenges that currtent limit the implementation of wireless technology (Marx, 2021).

The MXBI system was designed and built to provide affordability, flexibility of use, and easy adaptation by others. Furthermore, key features such as wireless connectivity for multi-directional data transfer among several MXBIs and a central server node, battery-based power supply, automatic individual animal identification, and homecage design enable the implementation of high-throughput behavioral assessment in NHPs. Common marmosets and long-tailed macaques are essential models in biomedical research. The development of transgenic primate models for various human diseases (Harris, 2021; Okano and Kishi, 2018; Sasaki et al., 2009; Tomioka et al., 2017) will require the behavioral characterization of a large number of animals. The automation of the behavioral assay will undoubtedly improve the development of novel protocols to characterize new disease therapies in a more standardized fashion, similar to rodent phenotyping pipelines (Aoki et al., 2017; Bohlen et al., 2014; Winter and Schaefers, 2011).

Finally, in recent years ethical concerns have arisen regarding animal experimentation, especially in the context of NHP. In Europe, the framework that regulates animal experimentation (European Union Directive 2010/63/EU on the Protection of Animals used for Scientific Purposes) is based on the so-called 3-R principles: Replacement – substituting animals when possible; Reduction – limiting the number of animals to the minimum; Refinement – decreasing the severity or incidence applied to animals (Russell and Burch, 1992). The interpretation of these principles is a matter of constant debate, which recently focused on the autonomy and self-determination of animals (Arnason, 2020). Within this framework, attempts to improve animal welfare rely on the design of experimental protocols that offer voluntary training and testing to the animals directly in their home cages with minimum human interaction to promote a sense of agency (Berger et al., 2018; Calapai et al., 2017; Evans et al., 2008; Fagot et al., 2014; Richardson et al., 1990; Washburn and Rumbaugh, 1992). Developing cognitive testing systems as means of enrichment enables access to captive colonies of several NHP species that otherwise would be difficult or even impossible to test, such as breeding colonies, sanctuaries, and zoos.

5.2 Limitations and future directions

Voluntary testing in a distracting environment

Traditionally, experimenters in the laboratory rely on controlling individual animals' external and internal factors (e.g., water and food availability, social context) to influence their motivation to cooperate in experimental procedures (Sedaghat-Nejad et al., 2019), achieving a relatively homogeneous engagement across the animals tested. Additionally, conventional auditory psychophysics experiments take place under well-controlled acoustic environments, namely sound-attenuated chambers (Coleman, 2009; Leinwand et al., 2020; Remington et al., 2012), to avoid the interference of external sounds during the performance. Contrary to this, the training protocols presented in this dissertation took place directly at the animals' home cage within a NHP colony, where different sensory stimuli and social interactions act as distractors. Furthermore, no water or food control was implemented to motivate the animals to participate. Instead, animals freely engaged with our device.

Carrying experiments in a dynamic environment, such as a NHP colony, without relying on additional coercion to motivate the animals to engage in the experimental procedures have inherent challenges when interpreting the results, namely engagement dynamics and performance success or failure. Internal individual factors like the degree of curiosity, boldness, and reward appeal at a given time (thirst or hunger) have been shown to play an important role in the willingness to engage and cognitive performance (Forss et al., 2017; Freeman and Gosling, 2010; Koski et al., 2017; Koski and Burkart, 2015; Šlipogor et al., 2022). Additionally, it has been reported that the social context in which animals have been tested influences the performance in the task (Bunnell and Perkins, 1980; Drea and Wallen, 1999; Hopper et al., 2007; Reader and Laland, 2001; Santos et al., 2002). Furthermore, since the training device can be seen as a source of reward (Barton and Whiten, 1993; Boogert et al., 2006) the access to the training device (engagement) might be influenced by social rank (Prescott and Buchanan-Smith, 2003), limiting the interaction time for those animals with a lower rank. Moreover, the dynamic acoustic background added another level of interference. Continuous and intermittent sounds, the product of the ventilation system, animals' vocalizations, and staff activities might have interfered with the animals' performance during testing sessions by redirecting animals' attention to more salient sounds (Fritz et al., 2007; Kayser et al., 2005) or making the testing stimuli less intelligible due to spectro-temporal masking effects (Greenwood, 1961).

For these reasons, it still remains unclear and difficult to disentangle whether animals' engagement dynamics and the failure to learn some tasks presented in this dissertation reflect individual cognitive constraints or just a satisfactory interaction strategy for the animals, given a complex combination of social and individual states. Moreover, statistically speaking, testing entire social groups may constrain the sample size since all subjects within a social group might not be considered independent from each other (Burkart et al., 2014; Cronin et al., 2014). This disadvantage is increased because groups are not usually uniform in size, composition, and physical environment (Cronin et al., 2017).

While auditory cognitive testing in social contexts opens the possibility of assessing behavior at a new level of complexity, the influences of this new level must be taken into strict consideration when designing future projects to reduce or avoid confined data interpretations. Current efforts are trying to assess the impact of intermittent background sound (vocalizations) on the animals' cognitive performance, implementing offline analysis of background recordings snippets during task engagement.

Further remarks

Individualized multi-dimensional closed-loop experiments

All experimental procedures presented in this dissertation relied on tasks structured as a staircase. This was devised so that animals could be gradually challenged with particular problems across a fixed number of pre-programmed steps that ultimately led the animals to show the desired behavior. To automate the navigation throughout the staircases, we implemented an algorithm that monitored individual animal performance adjusting the difficulty accordingly, presumably to keep animals at a comfortable level to avoid frustration (Berger et al., 2018). Although the implementation of this close-loop feedback eventually trained most of the tested animals, the unidimensionality of the feedback based on the "difficulty" with fixed parameters across the staircase made the design, deployment, testing, and adaption of tasks a time-consuming process. Proof of that process is the number of task versions tested in Chapter 2 before finding the ideal features that most animals needed to learn the visio-acoustic discrimination task, including the implementation of a 3-alternative choice task that improved the performance of those marmosets that performed at a chance level using a 2-alternative choice task.

Task design is crucial when testing for cognitive mechanisms because, as illustrated earlier, failure to learn a given task might underlie inappropriate task structure rather than cognitive limitations to solve the problem. The implementation of additional close-loops to keep track of parallel features in the task (e.g., target delay presentation, target position in the screen) in relationship with individual animal interactions (e.g., target response delay, area of touch in the screen) could contribute to a faster task design by identifying response patterns across animals, but also to a greater engagement and performance by further personalizing training protocols where physical and personality features are taken into account.

Testing beyond the laboratory

Cognitive testing with touchscreen devices have been implemented in several scenarios apart from laboratories (e.g., zoos, sanctuaries, and in the field) testing for several cognitive capabilities (e.g., memory, visual perception) (Allritz et al., 2016; Gazes et al., 2017; Jacobson et al., 2019; Macdonald and Whiten, 2011; Perdue et al., 2012; Ross, 2009; Schmitt, 2016; Shumaker, 2018; Tanaka, 2016; Wagner et al., 2016; Whitehouse et al., 2013). However, auditory testing is scarce in the literature. A common observation is that NHPs are notoriously difficult to train in auditory tasks, generally displaying a bias toward vision (Hirst et al., 2018; Schmitt, 2019; Van Essen et al., 2019). The possibility of expanding the auditory testing procedures to different species, rarely available in experimental primate colonies, presents an opportunity for novel phylogenetic comparisons to understand the evolution of specific cognitive traits better (Egelkamp and Ross, 2019; MacDonald and Ritvo, 2016).

5.3 Concluding remarks

A recurrent point throughout all chapters of this dissertation has been the optimization of experimental procedures in auditory cognition through automation. Historically, automation in behavioral sciences has been deployed at 3 levels. First, it aims to reduce the constraint of human time resources to the greatest extent. Second, it intends to scale up data collection. Third, it seeks the standardization of protocols to achieve reproducibility. Finally, once automation is achieved, the possibility to expand the detail and dimensionality of observations and measurements relying on state-of-the-art technology seems unavoidable.

Technological advances in the last 15 years have unprecedentedly enabled a considerable amount of behavioral data collection (Marx, 2013). Powerful computer vision algorithms have been used to track an individual's pose, position, and interactions with conspecifics (Lauer et al., 2022; Mathis et al., 2018; Nath et al., 2019). Miniaturized probes allow for high-density chronic recordings of neural activity (Steinmetz et al., 2021, 2019). The advent of the so-called "big data" promises the possibility to: scale up the behavioral data through the automation of the behavioral assay (Branson et al., 2009; Winter and Schaefers, 2011), go from discrete to continuous measurements of behavior (Gold and Shadlen, 2001; Selen et al., 2012), find behavioral relevance instead of nuisance in unrestrained animals (Gouvêa et al., 2014; Kawagoe et al., 1998; Kepecs et al., 2008; Luce, 1986), develop computational models to describe behavior (Gold and Shadlen, 2001; Kepecs et al., 2008; Sugrue et al., 2004), find simplicity in higher levels of behavior, namely group behavior (Bialek et al., 2014; Jin and Kozhevnikov, 2011; Jordan et al., 2013; Nathan et al., 2022), contrast behavioral responses across different contexts (Kane and Zamani, 2014; Portugues et al., 2014). However, how exactly this data will be collected (e.g., the conceptual framework of study) and then handled (e.g., data openness) are among the biggest challenges to overcome (Gomez-Marin et al., 2014; Jazayeri and Afraz, 2017; Krakauer et al., 2017).

It is clear that behavior as a product of evolution is a unifying organismal process where genes, neural function, anatomy, and environment converge and interrelate to warrant the individual's survival and successful reproduction. As such, its study and its mechanisms will always require an intricate collaboration among several study fields. Experimental psychology, ethology, and neuroscience face an exciting future, and I look forward to being part of it.

5.4 References

Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B.G., Witzel, T., Belliveau, J.W., 2006. Task-modulated "what" and "where" pathways in human auditory cortex. Proc. Natl. Acad. Sci. U.S.A. 103, 14608–14613. https://doi.org/10.1073/pnas.0510480103

Aitkin, L.M., Merzenich, M.M., Irvine, D.R., Clarey, J.C., Nelson, J.E., 1986. Frequency representation in auditory cortex of the common marmoset (Callithrix jacchus jacchus). Journal of Comparative Neurology 252, 175–185.

Allritz, M., Call, J., Borkenau, P., 2016. How chimpanzees (Pan troglodytes) perform in a modified emotional Stroop task. Anim Cogn 19, 435–449. https://doi.org/10.1007/s10071-015-0944-3

Aoki, R., Tsubota, T., Goya, Y., Benucci, A., 2017. An automated platform for high-throughput mouse behavior and physiology with voluntary head-fixation. Nat Commun 8, 1196. https://doi.org/10.1038/s41467-017-01371-0

Arnason, G., 2020. The Emergence and Development of Animal Research Ethics: A Review with a Focus on Nonhuman Primates. Sci Eng Ethics 26, 2277–2293. https://doi.org/10.1007/s11948-020-00219-z

Balleine, B., 1992. Instrumental performance following a shift in primary motivation depends on incentive learning. Journal of Experimental Psychology: Animal Behavior Processes 18, 236–250. https://doi.org/10.1037/0097-7403.18.3.236

Barton, R.A., Whiten, A., 1993. Feeding competition among female olive baboons, Papio anubis. Animal Behaviour 46, 777–789. https://doi.org/10.1006/anbe.1993.1255

Bendor, D., Osmanski, M.S., Wang, X., 2012. Dual-pitch processing mechanisms in primate auditory cortex. Journal of Neuroscience 32, 16149–16161.

Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165. https://doi.org/10.1038/nature03867

Berger, M., Calapai, A., Stephan, V., Niessing, M., Burchardt, L., Gail, A., Treue, S., 2018. Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. Journal of Neurophysiology 119, 796–807. https://doi.org/10.1152/jn.00614.2017

Bialek, W., Cavagna, A., Giardina, I., Mora, T., Pohl, O., Silvestri, E., Viale, M., Walczak, A.M., 2014. Social interactions dominate speed control in poising natural flocks near criticality. Proceedings of the National Academy of Sciences 111, 7212–7217.

Biben, M., 1993. Recognition of order effects in squirrel monkey antiphonal call sequences. American Journal of Primatology 29, 109–124.

Bohlen, M., Hayes, E.R., Bohlen, B., Bailoo, J.D., Crabbe, J.C., Wahlsten, D., 2014. Experimenter effects on behavioral test scores of eight inbred mouse strains under the influence of ethanol. Behavioural Brain Research 272, 46–54. https://doi.org/10.1016/j.bbr.2014.06.017

Boogert, N.J., Reader, S.M., Laland, K.N., 2006. The relation between social rank, neophobia and individual learning in starlings. Animal Behaviour 72, 1229–1239. https://doi.org/10.1016/j.anbehav.2006.02.021

Borton, D.A., Yin, M., Aceros, J., Nurmikko, A., 2013. An implantable wireless neural interface for recording cortical circuit dynamics in moving primates. J. Neural Eng. 10, 026010. https://doi.org/10.1088/1741-2560/10/2/026010

Branson, K., Robie, A.A., Bender, J., Perona, P., Dickinson, M.H., 2009. High-throughput ethomics in large groups of Drosophila. Nature methods 6, 451–457.

Brosch, M., Selezneva, E., Bucks, C., Scheich, H., 2004. Macaque monkeys discriminate pitch relationships. Cognition 91, 259–272. https://doi.org/10.1016/j.cognition.2003.09.005

Brown, C.H., Schessler, T., Moody, D., Stebbins, W., 1982. Vertical and horizontal sound localization in primates. The Journal of the Acoustical Society of America 72, 1804–1811. https://doi.org/10.1121/1.388653

Bunnell, B.N., Perkins, M.N., 1980. Performance correlates of social behavior and organization: Social rank and complex problem solving in crab-eating macaques (M. fascicularis). Primates 21, 515–523. https://doi.org/10.1007/BF02373840

Burkart, J.M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., Huber, J., Isler, K., Kosonen, Z.K., Martins, E., 2014. The evolutionary origin of human hyper-cooperation. Nature communications 5, 1–9.

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., Gail, A., 2017. A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. Behav Res 49, 35–45. https://doi.org/10.3758/s13428-016-0707-3

Canteloup, C., Hoppitt, W., van de Waal, E., 2020. Wild primates copy higher-ranked individuals in a social transmission experiment. Nat Commun 11, 459. https://doi.org/10.1038/s41467-019-14209-8

Carmena, J.M., Lebedev, M.A., Crist, R.E., O’Doherty, J.E., Santucci, D.M., Dimitrov, D.F., Patil, P.G., Henriquez, C.S., Nicolelis, M.A.L., 2003. Learning to Control a Brain–Machine Interface for Reaching and Grasping by Primates. PLoS Biol 1, e42. https://doi.org/10.1371/journal.pbio.0000042

Cheung, S.W., Nagarajan, S.S., Schreiner, C.E., Bedenbaugh, P.H., Wong, A., 2005. Plasticity in primary auditory cortex of monkeys with altered vocal production. Journal of Neuroscience 25, 2490–2503.

Coleman, M.N., 2009. What Do Primates Hear? A Meta-analysis of All Known Nonhuman Primate Behavioral Audiograms. Int J Primatol 30, 55–91. https://doi.org/10.1007/s10764-008-9330-1

Cronin, K.A., Jacobson, S.L., Bonnie, K.E., Hopper, L.M., 2017. Studying primate cognition in a social setting to improve validity and welfare: a literature review highlighting successful approaches. PeerJ 5, e3649.

Cronin, K.A., Van Leeuwen, E.J., Vreeman, V., Haun, D.B., 2014. Population-level variability in the social climates of four chimpanzee societies. Evolution and Human Behavior 35, 389–396.

de Waal, F.B., Suchak, M., 2010. Prosocial primates: selfish and unselfish motivations. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 2711–2722.

Desai, N.P., Lehman, C., Munson, B., Wilson, M., 2018. Supervised and unsupervised machine learning approaches to classifying chimpanzee vocalizations. The Journal of the Acoustical Society of America 143, 1786–1786. https://doi.org/10.1121/1.5035847

Drea, C.M., Wallen, K., 1999. Low-status monkeys "play dumb" when learning in mixed social groups. Proc. Natl. Acad. Sci. U.S.A. 96, 12965–12969. https://doi.org/10.1073/pnas.96.22.12965

Egelkamp, C.L., Ross, S.R., 2019. A review of zoo-based cognitive research using touchscreen interfaces. Zoo Biology 38, 220–235. https://doi.org/10.1002/zoo.21458

Eliades, S.J., Wang, X., 2008a. Chronic multi-electrode neural recording in free-roaming monkeys. Journal of neuroscience methods 172, 201–214.

Eliades, S.J., Wang, X., 2008b. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453, 1102–1106.

Evans, T.A., Beran, M.J., Chan, B., Klein, E.D., Menzel, C.R., 2008. An efficient computerized testing method for the capuchin monkey (Cebus apella): Adaptation of the LRC-CTS to a socially housed nonhuman primate species. Behavior Research Methods 40, 590–596. https://doi.org/10.3758/BRM.40.2.590

Fagot, J., Gullstrand, J., Kemp, C., Defilles, C., Mekaouche, M., 2014. Effects of freely accessible computerized test systems on the spontaneous behaviors and stress level of Guinea baboons (*Papio papio*): Computerized Testing and Welfare in Baboons. Am. J. Primatol. 76, 56–64. https://doi.org/10.1002/ajp.22193

Fechner, G.T., 1860. Elemente der psychophysik. Breitkopf u. Härtel.

Fishman, Y.I., Kim, M., Steinschneider, M., 2017. A Crucial Test of the Population Separation Model of Auditory Stream Segregation in Macaque Primary Auditory Cortex. J. Neurosci. 37, 10645–10655. https://doi.org/10.1523/JNEUROSCI.0792-17.2017

Forss, S.I.F., Koski, S.E., van Schaik, C.P., 2017. Explaining the Paradox of Neophobic Explorers: The Social Information Hypothesis. Int J Primatol 38, 799–822. https://doi.org/10.1007/s10764-017-9984-7

Foster, J.D., Nuyujukian, P., Freifeld, O., Gao, H., Walker, R., Ryu, S.I., Meng, T.H., Murmann, B., Black, M.J., Shenoy, K.V., 2014. A freely-moving monkey treadmill model. Journal of neural engineering 11, 046020.

Freeman, H.D., Gosling, S.D., 2010. Personality in nonhuman primates: a review and evaluation of past research. Am. J. Primatol. 72, 653–671. https://doi.org/10.1002/ajp.20833

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007. Auditory attention—focusing the searchlight on sound. Current Opinion in Neurobiology 17, 437–455. https://doi.org/10.1016/j.conb.2007.07.011

Gazes, R.P., Diamond, R.F.L., Hope, J.M., Caillaud, D., Stoinski, T.S., Hampton, R.R., 2017. Spatial representation of magnitude in gorillas and orangutans. Cognition 168, 312–319. https://doi.org/10.1016/j.cognition.2017.07.010

Geissmann, T., 2002. Duet-splitting and the evolution of gibbon songs. Biological Reviews 77, 57–76.

Gold, J.I., Shadlen, M.N., 2001. Neural computations that underlie decisions about sensory stimuli. Trends in cognitive sciences 5, 10–16.

Gomez-Marin, A., Paton, J.J., Kampff, A.R., Costa, R.M., Mainen, Z.F., 2014. Big behavioral data: psychology, ethology and the foundations of neuroscience. Nat Neurosci 17, 1455–1462. https://doi.org/10.1038/nn.3812

Gouvêa, T.S., Monteiro, T., Soares, S., Atallah, B.V., Paton, J.J., 2014. Ongoing behavior predicts perceptual report of interval duration. Frontiers in neurorobotics 8, 10.

Graving, J.M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B.R., Couzin, I.D., 2019. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. eLife 8, e47994. https://doi.org/10.7554/eLife.47994

Greenwood, D.D., 1961. Auditory Masking and the Critical Band. The Journal of the Acoustical Society of America 33, 484–502. https://doi.org/10.1121/1.1908699

Hage, S.R., Gavrilov, N., Nieder, A., 2016. Developmental changes of cognitive vocal control in monkeys. Journal of Experimental Biology 219, 1744–1749. https://doi.org/10.1242/jeb.137653

Hage, S.R., Gavrilov, N., Nieder, A., 2013. Cognitive Control of Distinct Vocalizations in Rhesus Monkeys. Journal of Cognitive Neuroscience 25, 1692–1701. https://doi.org/10.1162/jocn_a_00428

Harris, J.C., 2021. Animal models of neurodevelopmental disorders with behavioral phenotypes. Current Opinion in Psychiatry 34, 87–93. https://doi.org/10.1097/YCO.0000000000000675

Harrison, R.R., Fotowat, H., Chan, R., Kier, R.J., Olberg, R., Leonardo, A., Gabbiani, F., 2011. Wireless Neural/EMG Telemetry Systems for Small Freely Moving Animals. IEEE Trans. Biomed. Circuits Syst. 5, 103–111. https://doi.org/10.1109/TBCAS.2011.2131140

Hirst, R.J., Cragg, L., Allen, H.A., 2018. Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. Neuroscience & Biobehavioral Reviews 94, 286–301. https://doi.org/10.1016/j.neubiorev.2018.07.012

Hopper, L.M., Spiteri, A., Lambeth, S.P., Schapiro, S.J., Horner, V., Whiten, A., 2007. Experimental studies of traditions and underlying transmission processes in chimpanzees. Animal Behaviour 73, 1021–1032.

Izumi, A., 1999. The Effect of Marker Frequency Disparity on the Discrimination of Gap Duration in Monkeys. Perception 28, 437–444. https://doi.org/10.1068/p2812

Jackson, A., Mavoori, J., Fetz, E.E., 2007. Correlations between the same motor cortex cells and arm muscles during a trained task, free behavior, and natural sleep in the macaque monkey. Journal of neurophysiology 97, 360–374.

Jacobson, S.L., Kwiatt, A.C., Ross, S.R., Cronin, K.A., 2019. The effects of cognitive testing on the welfare of zoo-housed Japanese macaques (Macaca fuscata). Applied Animal Behaviour Science 212, 90–97. https://doi.org/10.1016/j.applanim.2018.12.014

Jazayeri, M., Afraz, A., 2017. Navigating the neural space in search of the neural code. Neuron 93, 1003–1014.

Jeschke, M., Moser, T., 2015. Considering optogenetic stimulation for cochlear implants. Hearing Research 322, 224–234. https://doi.org/10.1016/j.heares.2015.01.005

Jin, D.Z., Kozhevnikov, A.A., 2011. A compact statistical model of the song syntax in Bengalese finch. PLoS computational biology 7, e1001108.

Johnson, J.A., Zatorre, R.J., 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. NeuroImage 31, 1673–1681. https://doi.org/10.1016/j.neuroimage.2006.02.026

Jordan, D., Kuehn, S., Katifori, E., Leibler, S., 2013. Behavioral diversity in microbes and low-dimensional phenotypic spaces. Proceedings of the National Academy of Sciences 110, 14018–14023.

Kane, S.A., Zamani, M., 2014. Falcons pursue prey using visual motion cues: new perspectives from animal-borne cameras. Journal of Experimental Biology 217, 225–234.

Kawagoe, R., Takikawa, Y., Hikosaka, O., 1998. Expectation of reward modulates cognitive signals in the basal ganglia. Nature neuroscience 1, 411–416.

Kayser, C., Petkov, C.I., Lippert, M., Logothetis, N.K., 2005. Mechanisms for Allocating Auditory Attention: An Auditory Saliency Map. Current Biology 15, 1943–1947. https://doi.org/10.1016/j.cub.2005.09.040

Kepecs, A., Uchida, N., Zariwala, H.A., Mainen, Z.F., 2008. Neural correlates, computation and behavioural impact of decision confidence. Nature 455, 227–231.

Knudsen, E.I., 2007. Fundamental Components of Attention. Annu. Rev. Neurosci. 30, 57–78. https://doi.org/10.1146/annurev.neuro.30.051606.094256

Knyazeva, S., Selezneva, E., Gorkin, A., Aggelopoulos, N.C., Brosch, M., 2018. Neuronal Correlates of Auditory Streaming in Monkey Auditory Cortex for Tone Sequences without Spectral Differences. Front. Integr. Neurosci. 12, 4. https://doi.org/10.3389/fnint.2018.00004

Konoike, N., Iwaoki, H., Miwa, M., Sakata, H., Itoh, K., Nakamura, K., 2022. Comparison of non-invasive, scalp-recorded auditory steady-state responses in humans, rhesus monkeys, and common marmosets. Sci Rep 12, 9210. https://doi.org/10.1038/s41598-022-13228-8

Koski, S.E., Buchanan-Smith, H.M., Ash, H., Burkart, J.M., Bugnyar, T., Weiss, A., 2017. Common marmoset (Callithrix jacchus) personality. Journal of Comparative Psychology 131, 326–336. https://doi.org/10.1037/com0000089

Koski, S.E., Burkart, J.M., 2015. Common marmosets show social plasticity and group-level similarity in personality. Sci Rep 5, 8878. https://doi.org/10.1038/srep08878

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., Poeppel, D., 2017. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 93, 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kuang, S., Zhang, T., 2014. Smelling directions: Olfaction modulates ambiguous visual motion perception. Sci Rep 4, 5796. https://doi.org/10.1038/srep05796

Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M.M., Di Santo, V., Soberanes, D., Feng, G., 2022. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nature Methods 19, 496–504.

Lee, S.B., Yin, M., Manns, J.R., Ghovanloo, M., 2013. A wideband dual-antenna receiver for wireless recording from animals behaving in large arenas. IEEE Transactions on Biomedical Engineering 60, 1993–2004.

Leinwand, J.G., Huskisson, S.M., Egelkamp, C.L., Hopper, L.M., 2020. Within-and between-species variation in the responses of three primate species to a touchscreen gambling task. Learning and Motivation 71, 101635.

Lewkowicz, D.J., Ghazanfar, A.A., 2009. The emergence of multisensory systems through perceptual narrowing. Trends in Cognitive Sciences 13, 470–478. https://doi.org/10.1016/j.tics.2009.08.004

Lewkowicz, D.J., Kraebel, K.S., 2004. The Value of Multisensory Redundancy in the Development of Intersensory Perception.

Luce, R.D., 1986. Response times: Their role in inferring elementary mental organization. Oxford University Press on Demand.

Lui, L.L., Mokri, Y., Reser, D.H., Rosa, M.G., Rajan, R., 2015. Responses of neurons in the marmoset primary auditory cortex to interaural level differences: comparison of pure tones and vocalizations. Frontiers in neuroscience 9, 132.

Macdonald, C., Whiten, A., 2011. The 'Living Links to Human Evolution' Research Centre in Edinburgh Zoo: a new endeavour in collaboration: EDINBURGH ZOO: 'LIVING LINKS TO HUMAN EVOLUTION' RESEARCH CENTRE. International Zoo Yearbook 45, 7–17. https://doi.org/10.1111/j.1748-1090.2010.00120.x

MacDonald, S.E., Ritvo, S., 2016. Comparative cognition outside the laboratory. Comparative Cognition & Behavior Reviews 11.

Mackey, C.A., Hauser, S., Schoenhaut, A.M., Temghare, N., Ramachandran, R., 2022. Psychometric and subcortical neurometric measures of temporal discrimination in rhesus macaques (preprint). Neuroscience. https://doi.org/10.1101/2022.08.05.502987

Mante, V., Sussillo, D., Shenoy, K.V., Newsome, W.T., 2013. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503, 78–84. https://doi.org/10.1038/nature12742

Marx, V., 2021. Neuroscientists go wireless. Nat Methods 18, 1150–1154. https://doi.org/10.1038/s41592-021-01281-6

Marx, V., 2013. The big challenges of big data. Nature 498, 255–260. https://doi.org/10.1038/498255a

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nature neuroscience 21, 1281–1289.

Mielke, A., Zuberbühler, K., 2013. A method for automated individual, species and call type recognition in free-ranging animals. Animal Behaviour 86, 475–482. https://doi.org/10.1016/j.anbehav.2013.04.017

Miller, C.T., Dibble, E., Hauser, M.D., 2001a. Amodal completion of acoustic signals by a nonhuman primate. Nature neuroscience 4, 783–784.

Miller, C.T., Dibble, E., Hauser, M.D., 2001b. Amodal completion of acoustic signals by a nonhuman primate. Nature neuroscience 4, 783–784.

Miller, C.T., Mandel, K., Wang, X., 2010. The communicative content of the common marmoset phee call during antiphonal calling. Am. J. Primatol. 72, 974–980. https://doi.org/10.1002/ajp.20854

Miller, C.T., Scarl, J., Hauser, M.D., 2004. Sensory biases underlie sex differences in tamarin long call structure. Animal Behaviour 68, 713–720.

Miller, C.T., Thomas, A.W., Nummela, S.U., de la Mothe, L.A., 2015. Responses of primate frontal cortex neurons during natural vocal communication. Journal of neurophysiology 114, 1158–1171.

Miller, C.T., Wang, X., 2006. Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. Journal of comparative physiology A 192, 27–38.

Miranda, H., Gilja, V., Chestek, C.A., Shenoy, K.V., Meng, T.H., 2010. HermesD: A high-rate long-range wireless transmission system for simultaneous multichannel neural recording applications. IEEE Transactions on Biomedical Circuits and Systems 4, 181–191.

Mohseni, P., Najafi, K., Eliades, S.J., Wang, X., 2005. Wireless multichannel biopotential recording using an integrated FM telemetry circuit. IEEE Transactions on Neural Systems and Rehabilitation Engineering 13, 263–271.

Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., Mathis, M.W., 2019. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat Protoc 14, 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Nathan, R., Monk, C.T., Arlinghaus, R., Adam, T., Alós, J., Assaf, M., Baktoft, H., Beardsworth, C.E., Bertram, M.G., Bijleveld, A.I., Brodin, T., Brooks, J.L., Campos-Candela, A., Cooke, S.J., Gjelland, K.Ø., Gupte, P.R., Harel, R., Hellström, G., Jeltsch, F., Killen, S.S., Klefoth, T., Langrock, R., Lennox, R.J., Lourie, E., Madden, J.R., Orchan, Y., Pauwels, I.S., Říha, M., Roeleke, M., Schlägel, U.E., Shohami, D., Signer, J., Toledo, S., Vilk, O., Westrelin, S., Whiteside, M.A., Jarić, I., 2022. Big-data approaches lead to an increased understanding of the ecology of animal movement. Science 375, eabg1780. https://doi.org/10.1126/science.abg1780

Okano, H., Kishi, N., 2018. Investigation of brain science and neurological/psychiatric disorders using genetically modified non-human primates. Current Opinion in Neurobiology 50, 1–6. https://doi.org/10.1016/j.conb.2017.10.016

Osmanski, M.S., Song, X., Guo, Y., Wang, X., 2016. Frequency discrimination in the common marmoset (Callithrix jacchus). Hearing research 341, 1–8.

Osmanski, M.S., Wang, X., 2011. Measurement of absolute auditory thresholds in the common marmoset (Callithrix jacchus). Hearing Research 277, 127–133. https://doi.org/10.1016/j.heares.2011.02.001

Perdue, B.M., Clay, A.W., Gaalema, D.E., Maple, T.L., Stoinski, T.S., 2012. Technology at the Zoo: The Influence of a Touchscreen Computer on Orangutans and Zoo Visitors: Technology at the Zoo. Zoo Biol. 31, 27–39. https://doi.org/10.1002/zoo.20378

Phaniraj, N., Wierucka, K., Zürcher, Y., Burkart, J.M., 2022. Optimising source identification from marmoset vocalisations with hierarchical machine learning classifiers (preprint). Animal Behavior and Cognition. https://doi.org/10.1101/2022.11.19.517179

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., Reby, D., 2016. Voice modulation: a window into the origins of human vocal control? Trends in cognitive sciences 20, 304–318.

Pomberger, T., Risueno-Segovia, C., Gultekin, Y.B., Dohmen, D., Hage, S.R., 2019. Cognitive control of complex motor behavior in marmoset monkeys. Nat Commun 10, 3796. https://doi.org/10.1038/s41467-019-11714-8

Portugues, R., Feierstein, C.E., Engert, F., Orger, M.B., 2014. Whole-brain activity maps reveal stereotyped, distributed networks for visuomotor behavior. Neuron 81, 1328–1343.

Prescott, M.J., Buchanan-Smith, H.M., 2003. Training Nonhuman Primates Using Positive Reinforcement Techniques. Journal of Applied Animal Welfare Science 6, 157–161. https://doi.org/10.1207/S15327604JAWS0603_01

Reader, S.M., Laland, K.N., 2001. Primate Innovation: Sex, Age ans SOcial Rank Differences. International Journal of Primatology 22, 787–805. https://doi.org/10.1023/A:1012069500899

Remington, E.D., Chen, C., Wang, X., 2022. Sound localization acuity of the common marmoset (*Callithrix jacchus*) (preprint). Neuroscience. https://doi.org/10.1101/2022.08.13.503849

Remington, E.D., Osmanski, M.S., Wang, X., 2012. An Operant Conditioning Method for Studying Auditory Behaviors in Marmoset Monkeys. PLoS ONE 7, e47895. https://doi.org/10.1371/journal.pone.0047895

Remington, E.D., Wang, X., 2019. Neural representations of the full spatial field in auditory cortex of awake marmoset (Callithrix jacchus). Cerebral Cortex 29, 1199–1216.

Richardson, W.K., Washburn, D.A., Hopkins, W.D., Savage-rumbaugh, E.S., Rumbaugh, D.M., 1990. The NASA/LRC computerized test system. Behavior Research Methods, Instruments, & Computers 22, 127–131. https://doi.org/10.3758/BF03203132

Rizk, M., Bossetti, C.A., Jochum, T.A., Callender, S.H., Nicolelis, M.A., Turner, D.A., Wolf, P.D., 2009. A fully implantable 96-channel neural data acquisition system. Journal of neural engineering 6, 026002.

Romero-Mujalli, D., Bergmann, T., Zimmermann, A., Scheumann, M., 2021. Utilizing DeepSqueak for automatic detection and classification of mammalian vocalizations: a case study on primate vocalizations. Sci Rep 11, 24463. https://doi.org/10.1038/s41598-021-03941-1

Ross, S.R., 2009. Sequential list-learning by an adolescent lowland gorilla (*Gorilla gorilla gorilla*) using an infrared touchframe apparatus. IS 10, 115–129. https://doi.org/10.1075/is.10.2.02ros

Roy, S., Wang, X., 2012. Wireless multi-channel single unit recording in freely moving and vocalizing primates. Journal of neuroscience methods 203, 28–40.

Roy, S., Zhao, L., Wang, X., 2016. Distinct neural activities in premotor cortex during natural vocal behaviors in a New World primate, the common marmoset (Callithrix jacchus). Journal of Neuroscience 36, 12168–12179.

Russell, W.M.S., Burch, R.L., 1992. The principles of humane experimental technique, Special ed., [Nachdr. der Ausg.] London 1959. ed. Universities Federation for Animal Welfare, Potters Bar, Herts.

Sadagopan, S., Wang, X., 2008. Level invariant representation of sounds by populations of neurons in primary auditory cortex. Journal of Neuroscience 28, 3415–3426.

Santos, L.R., Sulkowski, G.M., Spaepen, G.M., Hauser, M.D., 2002. Object individuation using property/kind information in rhesus macaques (Macaca mulatta). Cognition 83, 241–264.

Sasaki, E., Suemizu, H., Shimada, A., Hanazawa, K., Oiwa, R., Kamioka, M., Tomioka, I., Sotomaru, Y., Hirakawa, R., Eto, T., Shiozawa, S., Maeda, T., Ito, M., Ito, R., Kito, C., Yagihashi, C., Kawai, K., Miyoshi, H., Tanioka, Y., Tamaoki, N., Habu, S., Okano, H., Nomura, T., 2009. Generation of transgenic non-human primates with germline transmission. Nature 459, 523–527. https://doi.org/10.1038/nature08090

Sato, H., Berry, C.W., Peeri, Y., Baghoomian, E., Casey, B.E., Lavella, G., VandenBrooks, J.M., Harrison, J., Maharbiz, M.M., 2009. Remote radio control of insect flight. Frontiers in integrative neuroscience 24.

Schmitt, V., 2019. Implementing portable touchscreen-setups to enhance cognitive research and enrich zoo-housed animals. Journal of Zoo and Aquarium Research 7, 50–58. https://doi.org/10.19227/jzar.v7i2.314

Schmitt, V., 2016. Using touchscreen technology to facilitate comparative cognition and enhance animal welfare. https://doi.org/10.13140/RG.2.2.14150.47680

Schneider, D.M., Nelson, A., Mooney, R., 2014. A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature 513, 189–194. https://doi.org/10.1038/nature13724

Schwarz, D.A., Lebedev, M.A., Hanson, T.L., Dimitrov, D.F., Lehew, G., Meloy, J., Rajangam, S., Subramanian, V., Ifft, P.J., Li, Z., Ramakrishnan, A., Tate, A., Zhuang, K.Z., Nicolelis, M.A.L., 2014. Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. Nat Methods 11, 670–676. https://doi.org/10.1038/nmeth.2936

Sedaghat-Nejad, E., Herzfeld, D.J., Hage, P., Karbasi, K., Palin, T., Wang, X., Shadmehr, R., 2019. Behavioral training of marmosets and electrophysiological recording from the cerebellum. Journal of Neurophysiology 122, 1502–1517. https://doi.org/10.1152/jn.00389.2019

Selen, L.P., Shadlen, M.N., Wolpert, D.M., 2012. Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. Journal of Neuroscience 32, 2276–2286.

Shumaker, R.W., 2018. The history and status of cognitive research with great apes in the United States. Japanese Journal of Animal Psychology, The Japanese Journal of Animal Psychology 68, 105–119. https://doi.org/10.2502/janip.68.2.3

Šlipogor, V., Graf, C., Massen, J.J.M., Bugnyar, T., 2022. Personality and social environment predict cognitive performance in common marmosets (Callithrix jacchus). Sci Rep 12, 6702. https://doi.org/10.1038/s41598-022-10296-8

Song, X., Osmanski, M.S., Guo, Y., Wang, X., 2016. Complex pitch perception mechanisms are shared by humans and a New World monkey. Proc. Natl. Acad. Sci. U.S.A. 113, 781–786. https://doi.org/10.1073/pnas.1516120113

Stebbins, W.C., 1973. Hearing of old world monkeys (Cercopithecinae). Am. J. Phys. Anthropol. 38, 357–364. https://doi.org/10.1002/ajpa.1330380233

Steinmetz, N.A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., 2021. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. Science 372, eabf4588.

Steinmetz, N.A., Zatka-Haas, P., Carandini, M., Harris, K.D., 2019. Distributed coding of choice, action and engagement across the mouse brain. Nature 576, 266–273.

Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. science 304, 1782–1787.

Sutton, R.S., Barto, A.G., 2018. Reinforcement learning: An introduction. MIT press.

Szuts, T.A., Fadeyev, V., Kachiguine, S., Sher, A., Grivich, M.V., Agrochão, M., Hottowy, P., Dabrowski, W., Lubenov, E.V., Siapas, A.G., 2011. A wireless multi-channel neural amplifier for freely moving animals. Nature neuroscience 14, 263–269.

Takahashi, D.Y., Fenley, A.R., Teramoto, Y., Narayanan, D.Z., Borjon, J.I., Holmes, P., Ghazanfar, A.A., 2015. The developmental dynamics of marmoset monkey vocal production. Science 349, 734–738.

Takahashi, D.Y., Narayanan, D.Z., Ghazanfar, A.A., 2013. Coupled oscillator dynamics of vocal turn-taking in monkeys. Current Biology 23, 2162–2168.

Tanaka, M., 2016. Comparative cognition in zoo animals. Japanese Journal of Animal Psychology, The Japanese Journal of Animal Psychology 66, 53–57. https://doi.org/10.2502/janip.66.1.8

Tomioka, I., Nogami, N., Nakatani, T., Owari, K., Fujita, N., Motohashi, H., Takayama, O., Takae, K., Nagai, Y., Seki, K., 2017. Generation of transgenic marmosets using a tetracyclin-inducible transgene expression system as a neurodegenerative disease model†. Biology of Reproduction 97, 772–780. https://doi.org/10.1093/biolre/iox129

Turesson, H.K., Ribeiro, S., Pereira, D.R., Papa, J.P., de Albuquerque, V.H.C., 2016. Machine Learning Algorithms for Automatic Classification of Marmoset Vocalizations. PLoS ONE 11, e0163041. https://doi.org/10.1371/journal.pone.0163041

Van Essen, D.C., Donahue, C.J., Coalson, T.S., Kennedy, H., Hayashi, T., Glasser, M.F., 2019. Cerebral cortical folding, parcellation, and connectivity in humans, nonhuman primates, and mice. Proc. Natl. Acad. Sci. U.S.A. 116, 26173–26180. https://doi.org/10.1073/pnas.1902299116

Wagner, K.E., Hopper, L.M., Ross, S.R., 2016. Asymmetries in the production of self-directed behavior by chimpanzees and gorillas during a computerized cognitive test. Anim Cogn 19, 343–350. https://doi.org/10.1007/s10071-015-0937-2

Wakita, M., 2020. Common marmosets (Callithrix jacchus) cannot recognize global configurations of sound patterns but can recognize adjacent relations of sounds. Behavioural Processes 176, 104136. https://doi.org/10.1016/j.beproc.2020.104136

Wakita, M., 2019. Auditory sequence perception in common marmosets (Callithrix jacchus). Behavioural Processes 162, 55–63. https://doi.org/10.1016/j.beproc.2019.01.014

Walker, J.D., Pirschel, F., Sundiang, M., Niekrasz, M., MacLean, J.N., Hatsopoulos, N.G., 2021. Chronic wireless neural population recordings with common marmosets. Cell reports 36, 109379.

Washburn, D.A., Rumbaugh, D.M., 1992. Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. Behavior Research Methods, Instruments, & Computers 24, 157–164. https://doi.org/10.3758/BF03203490

Watkins, P.V., Barbour, D.L., 2011. Rate-level responses in awake marmoset auditory cortex. Hearing research 275, 30–42.

Whitehouse, J., Micheletta, J., Powell, L.E., Bordier, C., Waller, B.M., 2013. The Impact of Cognitive Testing on the Welfare of Group Housed Primates. PLoS ONE 8, e78308. https://doi.org/10.1371/journal.pone.0078308

Wienicke, A., Häusler, U., Jürgens, U., 2001. Auditory frequency discrimination in the squirrel monkey. Journal of Comparative Physiology A 187, 189–195.

Winter, Y., Schaefers, A.T., 2011. A sorting system with automated gates permits individual operant experiments with mice from a social home cage. Journal of neuroscience methods 196, 276–280.

Yin, M., Borton, D.A., Komar, J., Agha, N., Lu, Y., Li, H., Laurens, J., Lang, Y., Li, Q., Bull, C., Larson, L., Rosler, D., Bezard, E., Courtine, G., Nurmikko, A.V., 2014. Wireless Neurosensor for Full-Spectrum Electrophysiology

Recordings during Free Behavior. Neuron 84, 1170–1182. https://doi.org/10.1016/j.neuron.2014.11.010

Zhang, Y.-J., Huang, J.-F., Gong, N., Ling, Z.-H., Hu, Y., 2018. Automatic detection and classification of marmoset vocalizations using deep and recurrent neural networks. The Journal of the Acoustical Society of America 144, 478–487. https://doi.org/10.1121/1.5047743