

Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
Ivo Sonntag (M.Sc.)
born in Berlin, Germany
Oral-examination: 28th of February 2020

Fluidity of functional ensembles in the infralimbic cortex of rats during reward seeking

Referees: Prof. Dr. Thomas Kuner
Prof. Dr. Rainer Spanagel

Summary

The medial prefrontal cortex (mPFC), specifically the prelimbic (PL) and the infralimbic region (IL), plays a crucial role during reward seeking behaviour. The IL specifically is involved in the control of reward seeking and has been implicated in the representation of different rewards. However, the precise representation of reward seeking behaviour on a neuronal network level within the IL remains elusive.

To investigate neuronal ensembles during reward seeking in the IL of the mPFC in rats, an operant conditioning paradigm was combined with imaging of the intracellular calcium concentration ($[Ca^{2+}]_i$) as a proxy for neuronal activity. The latter is achieved using GRIN lenses and miniaturized head-mounted fluorescence microscopes. A frame trigger was used to synchronize the operant conditioning chambers with the $[Ca^{2+}]_i$ data and an analysis pipeline was developed using a combination of custom designed Matlab classes and available open source software. Neurons were identified for three saccharin self-administration (SA) and one reinstatement (RE) session and then matched across sessions. Periods during which rats interacted with the operant conditioning setup were identified (e.g. lever presses and head entries into the reward port) and the corresponding $[Ca^{2+}]_i$ transients were used to identify neurons coactive during distinct phases of reward seeking behaviour. Neurons were classified according to the time point of their activity relative to the sequence of actions consisting of the lever press and the time before, during, and after the head entry.

This analysis revealed that subsets of neurons are preferentially active during distinct events of the reward seeking. Also, cells tuned to time points during the reward seeking did not appear or show tuning in all of the sessions. If they did show tuning, however, the phase of the reward seeking to which they showed tuning generally remained the same. Hence, the specific ensemble which is active during the reward seeking in each session changes. Individual neurons that are recruited into these ensembles, however, keep their tuning. Also, the composition of tuned neurons active during a specific behavioural phase remains stable. In addition, neurons that are active and tuned in multiple sessions do not appear to be arranged in a topology that can be identified with the methods used.

In conclusion, the sequence of the reward seeking behaviour is encoded in neuronal ensembles of the IL cortex. These ensembles are formed from a larger pool of available neurons in each session. Neurons participating in these ensembles preferentially keep their tuning to a phase of the reward seeking, but may not be recruited to each of the ensembles. Thus, ensembles representing identical behavioural episodes in different sessions are not stable, but fluidly change their composition.

Zusammenfassung

Der mediale Präfrontale Kortex (mPFC), besonders die prä-, und infralimbische region, spielt eine entscheidende Rolle während des Belohnungsverhaltens. Für die IL Region wurde dabei gezeigt, dass sie eine wichtige Rolle in der Kontrolle von Belohnungsverhalten, sowie der Abbildung der Belohnung spielt. Präzise Abbildung des Belohnungsverhaltens auf Ebene neuronaler Netzwerke ist allerdings nicht.

Um neuronale Ensembles während der Belohnungssuche in der IL Region des mPFC bei Ratten zu untersuchen, wurde operante Konditionierung mit Aufnahmen der intrazellulären Calciumkonzentration ($[Ca^{2+}]_i$) als Korrelat für neuronale Aktivität kombiniert. Letzteres wird mit GRIN-Linsen und miniaturisierten, am Kopf montierten Fluoreszenzmikroskopen erreicht. Ein Bildsignal (frame trigger) wurde verwendet, um die Kammer zum operanten Konditionieren mit den $[Ca^{2+}]_i$ Daten zu synchronisieren. Ebenfalls wurde eine Analyse-Pipeline unter Verwendung einer Kombination aus selbst verfassten Matlab-Klassen und verfügbarer Open-Source-Software entwickelt. Die Neuronen wurden für drei Saccharin Selbstverabreichungs-, (SA) und eine Rezidivsituation (RE) identifiziert und dann über die Sitzungen hinweg verfolgt. Zeiträume in denen Ratten mit der Kammer zum operanten Konditionieren interagierten wurden identifiziert (z. B. drücken des Hebels und Einführen des Kopfes in den Belohnungsöffnung) und die entsprechenden $[Ca^{2+}]_i$ -Transienten wurden verwendet, um Neuronen zu identifizieren, welche während verschiedener Phasen des Belohnungsverhaltens koaktiv waren. Neuronen wurden nach dem Zeitpunkt ihrer Aktivität in Bezug auf die Abfolge der Aktionen, bestehend aus dem Drücken des Hebels und der Zeit vor, während und nach dem Kopfeintritt, klassifiziert.

Die Analyse ergab, dass Teilgruppen von Neuronen während bestimmter Ereignisse der Belohnungssuche bevorzugt aktiv sind. Außerdem haben Zellen, die während der Belohnungssuche eine Ereignispräferenz zeigten, diese nicht in allen Sitzungen. Ebenfalls sind Zellen nicht in allen Sitzungen anhand ihrer Aktivität identifizierbar. Wenn sie Ereignispräferenz auf eine Phase der Belohnungssuche zeigten, blieb diese in der Regel dieselbe. Das spezifische Ensemble, das während der Belohnungssuche in jeder Sitzung aktiv ist, ändert sich daher kontinuierlich. Einzelne Neuronen, die für diese Ensembles rekrutiert werden, behalten jedoch ihre Ereignispräferenz bei. Auch die Zusammensetzung der Neurone mit Präferenz, die während einer bestimmten Verhaltensphase aktiv sind, bleibt stabil. Darüber hinaus scheinen Neuronen, die in mehreren Sitzungen aktiv und abgestimmt sind, nicht in einer Topologie angeordnet zu sein, die mit den verwendeten Methoden identifiziert werden kann.

Declaration

I, Ivo Sonntag, hereby declare that the work in this thesis represents my original research results. The thesis has been written by myself using the references and resources indicated. Any work of others has been appropriately marked in the methods section (section 2). The work has been conducted under the supervision of Prof. Thomas Kuner at the Institute of Anatomy and Cell Biology, Department of Functional Neuroanatomy, Medical Faculty, Heidelberg University, Germany.

This thesis is being submitted for the degree of Doctor of Natural Sciences at Heidelberg University, Germany, and has not been presented to any other university as part of an examination degree.

Heidelberg, 22.11.2019

.....

Ivo Sonntag

List of abbreviations

[Ca²⁺]_i	Intracellular calcium concentration
AP	Anterior posterior (axis)
ACC	Anterior cingulate cortex
CaM	Calmodulin
CNMF-E	Constrained nonnegative matrix factorisation for endoscopy
CMOS	Complementary metal-oxide semiconductor
CS	Conditioned stimulus
DAQ	Data acquisition (board)
FI	Fixed interval
FOV	Field of view
FR	Forced ratio
GABA	Gamma-Aminobutyric acid
GECI	Genetically encoded calcium indicator
GFP	Green fluorescent protein
GRIN	Gradient index (lens)
HE	Head entry
HPC	High performance computing cluster
ICA	Independent component analysis
IEG	Immediate early gene
IL	Infralimbic (cortex)
LED	Light emitting diode
LP	Lever press
MAP	Microtubule associated protein
mPFC	Medial prefrontal cortex
Nac	Nucleus accumbens
oPFC	Orbito prefrontal cortex
PFA	paraformaldehyde
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCB	Printed circuit board
PL	Prelimbic (cortex)
PR	Progressive ratio
rAAV	Recombinant adeno associated virus
RE	reinstatement
SA	Self-administration
rLP	Rewarded lever press
SNR	Signal to noise ratio
VTA	Ventral tegmental area
VI	Variable interval
VP	Ventral pallidum

Table of content

Summary.....	i
Zusammenfassung.....	ii
Declaration.....	iii
List of abbreviations.....	iv
Table of content.....	v
1. Introduction.....	1
1.1. Operant conditioning and self-administration.....	1
1.2. The mesocorticolimbic system.....	3
1.2.1. Infralimbic cortex of the mPFC.....	4
1.3. Neuronal ensembles.....	6
1.4. Imaging of neuronal population activity.....	9
1.4.1. Types of cellular <i>in-vivo</i> Calcium imaging.....	11
1.4.2. Miniaturized epifluorescence microscopes.....	12
1.5. Analysis of <i>in-vivo</i> epifluorescence calcium imaging.....	13
1.5.1. Extraction of calcium transients.....	14
1.5.2. High level analysis.....	15
1.6. Objectives.....	15
2. Materials and Methods.....	17
2.1. Miniature fluorescence microscope: UCLA Miniscope.....	17
2.2. Animals.....	17
2.3. GRIN lens implantation.....	18
2.4. Modified Luigs & Neumann linear actuator.....	20
2.5. Fixation of baseplate for miniaturized fluorescence microscope.....	22
2.6. Behavioural procedures.....	24
2.6.1. Operant self-administration hardware.....	24
2.6.2. Operant self-administration.....	24
2.6.3. Extinction.....	25
2.6.4. Cue-induced reinstatement.....	25
2.7.1. Thin sectioning and mounting.....	25
2.7.2. Image acquisition of fixed tissue sections.....	26
2.8. Analysis.....	27
2.8.1. List of MATLAB classes used.....	27
2.8.2. List of external Matlab pipelines.....	29
2.8.3. $[Ca^{2+}]_i$ imaging data analysis.....	30
2.8.4. Analysis of the operant conditioning data.....	30
2.8.5. Merging of $[Ca^{2+}]_i$ imaging and behavioural data.....	32

2.8.6.	Matching of cells between sessions.....	33
2.8.7.	Bootstrapping of rLP $[Ca^{2+}]_i$ transients.....	33
2.8.8.	Assignment of behavioural phases.....	34
2.8.9.	High performance computing cluster	34
2.8.10.	Behavioural statistics.....	35
3.	Results	36
3.1.	GRIN lens implantation and micro endoscopic imaging.....	36
3.1.1.	Micro-endoscopic imaging.....	37
3.1.2.	Analysis pipeline	39
3.2.	Operant self-administration behavioural results.....	41
3.2.1.	Overall efficacy of GRIN implantations in rats.....	45
3.3.	Combination of operant self-administration data and calcium imaging.....	46
3.4.	Identification of neuronal tuning using shuffling	50
3.5.	Assignment of behavioural phases to tuned neurons and averaged self-administration behaviour	55
3.6.	Matching of cells across multiple sessions	58
3.7.	Infralimbic network dynamics during operant self-administration.....	64
3.8.	Recruitment and stability of neuronal tuning during the operant self-administration paradigm.....	69
4.	Discussion	74
4.1.	Summary of findings	74
4.1.1.	Neurons are tuned to specific time points of the self-administration sequence....	74
4.2.	Neuronal tuning and fluid network dynamics during reward seeking.....	76
4.3.	Neurons tuned in multiple sessions are likely to keep the initial tuning.....	78
4.4.	Lack of topological organisation of tuned neurons.....	79
4.5.	Technical limitations of microendoscopic recordings.....	79
4.5.1.	Issues during microendoscopic imaging in rats	80
4.6.	Caveats in the identification of tuned neurons through bootstrapping	81
4.6.1.	Alternative metrics for bootstrapping.....	82
4.6.2.	Concluding remarks on the bootstrapping	83
4.7.	Conclusions and outlook	83
4.7.1.	Future improvements potential extensions	84
	Bibliography.....	86
	Acknowledgements.....	97

1. Introduction

Reward seeking behaviour is an essential trait driving behavioural actions toward outcomes that are favourable to the organism (Spanagel, 2009). Seeking a reward is always driven through the motivation to obtain that particular reward, however, if this motivation is misbalanced the reward seeking becomes detrimental to the organisms wellbeing, a state which is referred to as addiction (G. F. Koob & Volkow, 2016; G. Koob & Simon, 2009; Volkow, Koob, & McLellan, 2016). While some components of the neural system that drive reward seeking have been identified, the precise mechanisms controlling reward seeking behaviour are still elusive. An area crucial for the formation of reward seeking is the prefrontal cortex (Otis et al., 2017), with the infralimbic (IL) region of the latter playing an essential role in its control (Pfarr et al., 2015) and to some extent the encoding of the reward between natural and drug rewards (Pfarr et al., 2018).

Neuronal ensembles in the IL are assumed to be involved in the control the reward seeking. The structure of such ensembles however, has not been studied over longer time periods and at a temporal precision high enough to allow correlation of the activity of individual neurons to behavioural events. Knowledge of the precise representation of the behaviour and its contextual information is an important step towards understanding the functional role of the IL in reward seeking. This work shows a combination of operant self-administration, combined with miniaturized epifluorescence microscopy to allow the observation of neuronal ensembles during reward seeking.

1.1. Operant conditioning and self-administration

Operant conditioning is a tool of central importance in the study of reward seeking and allows the investigation of associations an individual makes between a performed behaviour and its outcome (Skinner, 1938). As mentioned above, in order to study the activation of neuronal circuits in the IL during reward seeking behaviour, an operant self-administration paradigm was used. During operant conditioning the subject learns to strengthen a behaviour based on its punishment or reinforcement by the experimenter (Skinner, 1938). This voluntary active learning process has to be distinguished from 'classical' conditioning (Pavlov, 1927), which reinforces involuntary responses. Both the reinforcement as well as the punishment, are the main procedures which are used in operant conditioning paradigms. Reinforcements come in the form of positive and negative reinforcements, with the former referring to the rewarding of a behaviour and the latter to the alleviation of a negative stimulus. Hence both represent a form of reward to the subject, by resulting in a desirable outcome. Punishment also comes in the form of positive and negative punishment, with the former referring to the induction of a negative

stimulus following a certain behaviour and the latter referring to the removal of a positive stimulus. Hence, both are punishing the subject by resulting in an undesirable outcome to the subject. The last paradigm to be mentioned is the extinction. During extinction, the reinforcement of a behaviour is stopped, leading to a decrease in the likelihood of that behaviour being performed.

Self-administration refers to the training of animals to elicit an operant or instrumental response (lever press, nose poke etc.) which results in the acquisition of a natural or a drug reward and is a form of operant conditioning. Operant self-administration is commonly used as a model to study forms of addiction in different organisms (e.g. humans, non-human primates, rodents, birds). The paradigm generally requires the subject to perform an action (operant response), e.g. pressing a lever, which will result in the receipt of a natural or drug reward and thus increase the probability to repeat the action, a process termed positive reinforcement (Weeks, 1962). The first phase of an operant self-administration is the acquisition of the behaviour, during which the animal learns that a distinct behavioural action will result in the acquisition of a reward. This is often done using devices commonly referred to as 'Skinner Boxes'. Today these often consist of an enclosure with plain walls and two sides which both contain e.g. a lever with which the animal can interact. Dependent on the type of reward that is used e.g. natural or drug rewards, the latter of which can also be administered intravenously, each of the sides also contains a port in which the animal can collect the reward in case of oral administration. Additionally, means to present cues are added to the chamber e.g. cue-lights or speakers to present sounds. It is important to note that during a self-administration paradigm, the subjects will likely follow a dose-response curve (Pomrenze, Baratta, Cadle, & Cooper, 2012), which will stabilize the blood levels of the reward. Also depending on the time period for which self-administration is allowed, an escalation of the behaviour can be exhibited for longer durations, believed to resemble dependence states (G. F. Koob et al., 2004).

After the acquisition of the behaviour, experimental variables can be changed. This can be separated into various reinforcement schedules like increasing the number of e.g. lever presses which are required to obtain a reward. A few variations are commonly used, with the term forced ratio (FR) referring to a fixed number of operant responses, while a progressive ratio (PR) refers to an increasing number of operant responses which will be necessary to obtain a reward. During the latter, the subject will eventually cease the operant responses, referred to as the break point. Another distinction can be made between fixed interval (FI) and variable interval (VI) schedules, referring to either fixed or variable time periods between reward acquisitions.

Another possibility of altering experimental parameters is the discontinuation of the positive reinforcement following the operant response. This is referred to as extinction and will lead to

the gradual cessation of the behaviour. This is often followed by a phase of restoration of the operant response, referred to as reinstatement and can be subdivided in two categories. The first is the cue-induced reinstatement, which requires the previous presentation of an environmental cue (e.g. light, sound or odour). Upon presentation of this cue to the operant response, the reward seeking is briefly increased, even after a successful extinction. The second type is a drug-induced reinstatement, upon which the operant response is reinitiated after exposure to the initial positive reinforcement.



Figure 1: Rat in an operant behaviour box. Drawing by Laura Schaaf.

In case of this study, a positive reinforcement in form of the natural reward saccharin was used with a fixed ratio (every lever press resulted in a reward). After a set of self-administration sessions, first extinction and then a cue-induced reinstatement was done, with the flashing of a light used as cue (see 2.6 for more details).

1.2. The mesocorticolimbic system

The regions of the brain mentioned above in relation to addiction and reward seeking behaviour are as a whole often referred to as the mesocorticolimbic system, which can be separated into a cortical and limbic component.

The limbic components of the basal ganglia are involved in motivational, emotional, contextual and affective information on behaviour. The limbic nuclei, as well as the amygdala, hippocampus

and the medial prefrontal cortex send glutamatergic connections to the nucleus accumbens (NA). The NA can be subdivided in a core and a shell region (Heimer et al., 1997). It has two GABAergic projections, one to the ventral tegmental area (VTA), and one to the ventral pallidum, both of which send GABAergic projections to the medial dorsal thalamus. From there glutamatergic connections are formed to the mPFC (Alexander & Crutcher, 1990; Heimer et al., 1997; Kalivas & Nakamura, 1999).

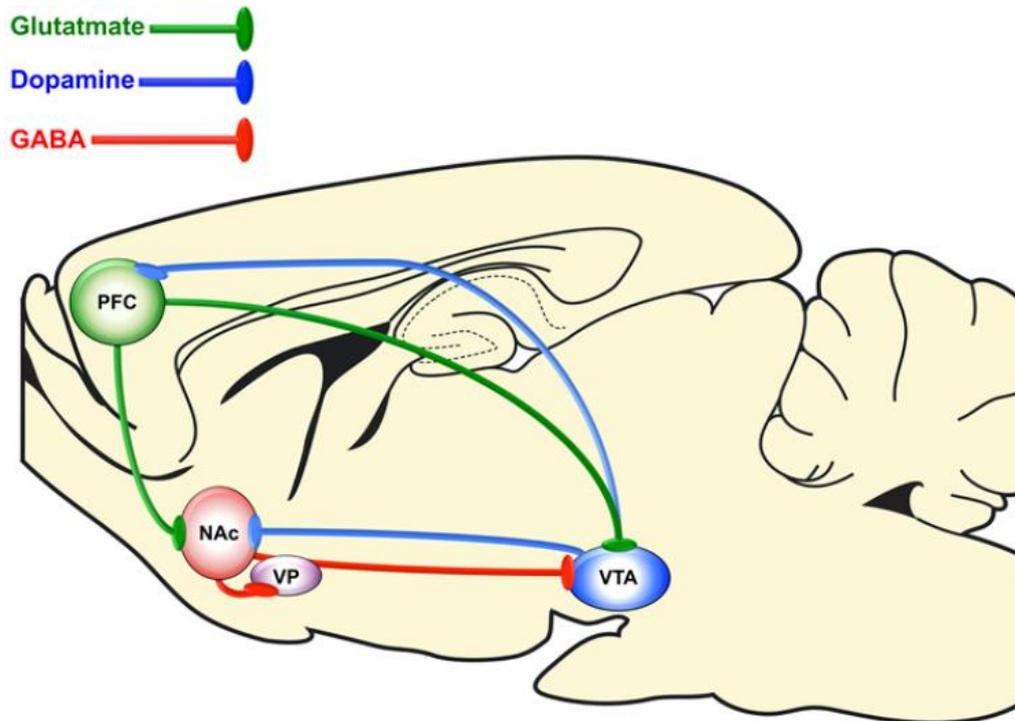


Figure 2: Limbic circuitry. Green arrows indicate glutamatergic pathways; red arrows indicate GABAergic pathways; blue arrows indicate dopaminergic pathways. Figure from Scofield & Kalivas, 2014.

The VTA is a source of many dopaminergic projections to the NA, amygdala, hippocampus, mPFC and the ventral pallidum and plays a key role in the modulation of the limbic circuit (Carr, O'Donnell, Card, & Sesack, 1999; Kalivas & Nakamura, 1999; Maslowski-Cobuzzi & Napier, 1994; Susan R. Sesack, Carr, Omelchenko, & Pinto, 2003; Wise, 2002).

1.2.1. Infralimbic cortex of the mPFC

The mPFC is part of the mesocorticolimbic system that drives both reward as well as fear related conditioning (Corbit & Balleine, 2003; Gourley, Lee, Howell, Pittenger, & Taylor, 2010). It receives

dopaminergic projections from the VTA (A10 neurons, Swanson, 1982), amygdala, hippocampus and others (Hoover & Vertes, 2007) and is thought to integrate those inputs to evaluate them for their salience and motivational significance regarding the associated context of the stimulus (Kalivas, 2009; Lasseter, Xie, Ramirez, & Fuchs, 2010). Executive control over the initiation of a response (both reward seeking as well as fear) is then driven over its projections to the nucleus accumbens (NAc), which in turn projects to areas like the dorsal striatum and the ventral pallidum (Kalivas, 2009; Lasseter et al., 2010; Vertes, 2004).

It is separated into multiple structures, the anterior cingulate cortex, prelimbic cortex (PL), infralimbic cortex (IL) and orbitofrontal cortex (oPFC) (Ongur & Price, 2000).

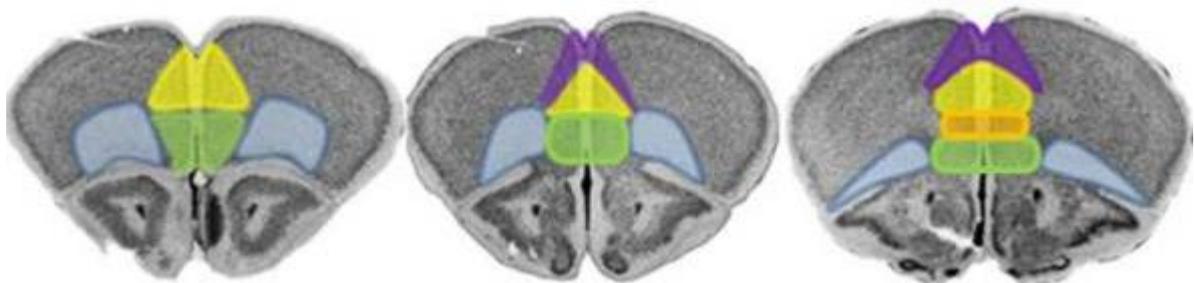


Figure 3: Anterior cingulate cortex (ACC) is shown in purple. Infralimbic cortex (IL) is shown in orange. Prelimbic cortex (PL) is shown in yellow. Medial orbitofrontal cortex (oPFC) is shown in green (medial) and blue (lateral). From (Gourley & Taylor, 2016).

Both the prelimbic (PL) as well as the infralimbic (IL) cortices are crucial in the formation of reward seeking behaviours. While the view of the PL as a 'GO' structure and the IL as a 'NO-go' structure (Heidbreder & Groenewegen, 2003) is heavily debated (Gourley & Taylor, 2016), there are distinct functional and anatomical differences between the areas which will be addressed in the following paragraphs.

The IL/PL boundary is defined through its anatomical connectivity, with their projections targets being largely different (Hoover & Vertes, 2007; Vertes, 2004). The PL innervates the NAc and the basolateral and lateral nuclei of the amygdala, while the IL targets the NAc shell and the basal, medial and central compartments of the amygdala (McDonald, Mascagni, & Guo, 1996; S. R. Sesack & Bunney, 1989).

Regarding functional differences, under controlled conditions, the ACC as well as the dorsal PL are thought to guide the continuation of behaviour (GO), while the ventral PL, IL and oPFC are thought to suppress behaviour (no-GO) (Heidbreder & Groenewegen, 2003).

Studies in fear conditioning showed, that the PL region is needed for the expression of conditioned fear (Do-Monte, Manzano-Nieves, Quiñones-Laracuente, Ramos-Medina, & Quirk,

2015) and disinhibition of the PL increases conditioned fear expression (Courtin et al., 2014). This is not the case for the IL, however it is thought to be involved in the extinction process. This interpretation can be misleading, as it is thought to be a novel process of learning, rather than the erasure of a memory (Todd, Vurbic, & Bouton, 2014). During extinction, the presentation of the conditioned stimulus, without the aversive or rewarded stimulus, is supposed to reduce the fear response or reward seeking behaviour. Lesion studies of the IL have shown, that levels of fear response will increase despite extinction training, while the initial acquisition of the behaviour will not be changed (Morgan, Romanski, & LeDoux, 1993; Quirk, Russo, Barron, & Lebron, 2000). Another study investigated the conditioned self-administration of ethanol and ablated neurons in the IL which were active during the conditioning paradigm, leading to an increase in reward seeking during the reinstatement (Pfarr et al., 2015). It was also shown, that increases in the activation of the IL (e.g. through optogenetic manipulation) will decrease fear responses during the presentation of the CS after extinction (Do-Monte et al., 2015).

A remaining open question is the structure of the neuronal ensemble activity during reward seeking and how stable the activation of these ensembles is during various stages of the operant self-administration. Especially in light of theories regarding the IL as a 'no-GO' structure, at least in the context of operant conditioning paradigms, elucidating the presence or absence of differences in neuronal responses to the reward seeking between phases of self-administration and reinstatement can help to understand the role of the IL during reward seeking. Based on the current knowledge of neuronal dynamics in the mPFC during reward seeking, it can be hypothesized for example to find neurons which are specifically active before or after the operant response, indicating either the induction of a behaviour, or retrospectively assigning a valence to it. An example for similar responses has already been demonstrated by Peters, O'Donnell, & Carelli, 2005, who showed increased activation of neurons in the mPFC (specific region unspecified) before an operant response.

1.3. Neuronal ensembles

The simultaneous investigation of large numbers of neurons has only become possible in recent decades. However, the functional implications of information processing through large neuronal networks have been postulated more than a century ago by Charles Sherrington in his book 'the integrative Action of the Nervous system' (Sherrington, 1906). There he suggested that the nervous system acts as a large integrative pathway, converging onto α -motoneurons, resulting in a meaningful output.

A more refined theory was later postulated by Donald O. Hebb, and was based on his theory called Hebb's Law, which stated:

"When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." (Hebb, 1949).

This is commonly paraphrased as 'Neurons that wire together fire together' and based on this, Hebb came to the conclusion that groups of neurons can form many different variations of connections which can act as processing units and referred to those units as 'assemblies'.

The goal of both was to understand the neural coding, as in, the relationship between a stimulus and the resulting neuronal response. Currently, there is no consensus about the precise mechanism of neural coding and whether it relies on information encoded by individual neurons or through assemblies of neurons. There is a large collection of theories on neural coding, and the concept of neuronal assemblies, also referred to as ensembles, as a group of neurons which form a functional unit, is a recurrent idea. For example, it is debated whether information is conveyed through the rate of action potential firing (rate code), or through precise spike timing of groups of neurons (temporal code) (Richmond, 2009), or both. Further examples are how information can be encoded through population activity or sparse coding. With the former, individual neurons would follow a noisy distribution of activity around a particular stimulus and hence recording from multiple neurons would increase the amount of information gathered about the stimulus or even allow to decipher it in the first place. An example for this is a population vector (Scott, 2000). Sparse coding on the other hand refers to the strong activation of only very few neurons that encode a particular piece of information (Beyeler, Rounds, Carlson, Dutt, & Krichmar, 2019; Richmond, 2009).

One of the reasons why the study of neuronal ensembles is a relatively recent endeavour may have been technical limitations. Another explanation may be, that the study of individual neurons resulted in the discovery of countless examples of neurons which appear to encode a particular stimulus or context through the firing of action potentials. Hubel and Wiesel were the first to describe neurons in the primary visual cortex, which respond to oriented slits of light (Hubel & Wiesel, 1959). Later also cells responding to orientations, sizes, positions and forms were discovered. Another example are place cells in the hippocampus (O'Keefe & Dostrovsky, 1971). However, something that all of these stimulus-tuned neurons have in common, is that that the response of individual neurons is always noisy (Wallis, 2018).

So far, there is no precise definition of a neuronal ensemble. The main identifying feature is likely, that at least two neurons have to show changes in activity correlated to each other and ideally

also to some external stimulus that allows the inference of their functional significance. Also given the multitude of technologies that are used to investigate neuronal ensembles, the amount of information from different dimensions often differs between studies. For example, ocular dominance columns could be defined as neuronal ensemble either based on their unique connectivity, or through their tuning to a particular receptive field (Horton & Adams, 2005).

A common tool to identify neuronal ensembles is the use of immediate early genes (IEG, George & Hope, 2017). This method utilizes a cellular mechanism during which strong persistent activation leads to increased intracellular calcium levels which activate the MAP kinase pathway and then lead to the activation of transcription factors for e.g. the C-fos or Arc promoter. The use of IEGs allows the detection of cells which were active during a particular time period and also allow the further characterisation of these neurons through immunohistochemistry. A caveat however is that the time period is not very precise (minutes up to an hour) and the readout is currently limited to two consecutive time points (Pfarr et al., 2018).

Alternatively, electrode arrays can be used (Malagon-Vina, Ciochi, Passecker, Dorffner, & Klausberger, 2018), which allow a very high temporal resolution (milliseconds), but at the expense that spatial information is lost and that particular cell types can only be identified through optogenetic tagging (Lima, Hromádka, Znamenskiy, & Zador, 2009). Also it remains challenging to track the same neurons over multiple days or even weeks.

Lastly, calcium imaging, which is discussed in more detail in chapter 1.4, allows the investigation of neurons over longer time periods, however at the expense of temporal resolution because the $[Ca^{2+}]_i$ concentration is used as an indirect measure of the neurons firing activity.

Regarding the question of whether individual neurons or neuronal ensembles encode information, one possible answer is both. While individual neurons can show tuning to a particular stimulus, in the case of neurons which are upstream of direct sensory inputs, this tuning has to be derived from the combination of inputs the neuron receives and integrates, creating the tuning. Hence, the sum of the input neurons can be regarded as a neuronal ensemble. Also, the underlying neuronal ensemble also had to carry the same information which the integrating neuron then relays in form of action potentials.

A good example for this is the processing of information in the visual system, which represents a multistage hierarchical network (Roelfsema & de Lange, 2016; Seabrook, Burbridge, Crair, & Huberman, 2017). For neuronal tuning of higher order representations, e.g. a neuron in the murine higher order visual cortex that represents a complex shape (Tschechne & Neumann, 2014), an individual neuron may be active only in response to a particular shape and could therefore encode that shape. However, the information encoded in the activation of this neuron

is just the convergence of the activity of its upstream network and contains its integrated information, making the neurons which are positioned upstream a neuronal ensemble.

Neuronal ensembles, defined as a group of neurons showing correlated activity to a behaviour or stimulus, can be stable as well as transient and even both. The latter was described by Liberti et al., 2016 in the HVC, a pre-motor nucleus of songbirds. They show that the amplitude as well as firing probability of principal neurons were drifting, while inhibitory neurons remained stable. Another example for a stable neuronal ensemble are striatal D1R and D2R neurons (Sheng, Lu, Shen, & Poo, 2019). During a cued lever pushing task, neuronal ensembles would emerge, showing a stable sequence of firing that only changed during other cued motor tasks. An example for instability of neuronal ensembles was shown through firing states of neuronal populations in the PL and ACC in a strategy switching task (Malagon-Vina 2018). If the animals are required to apply multiple strategies to obtain a reward, the firing state will change along with the strategy. If a rule that was presented before, is presented a second time, the firing state will not revert back to its state during the first period of that rule, but form a new state.

The concept of neuronal ensembles is not well defined and care has to be taken ensure that different interpretations of the term do not stand in the way of communicating results. As for this work, the term neuronal ensemble is defined as a group of neurons which are repeatedly active during the same stimulus or behaviour. Regarding the question on the existence of neuronal ensembles in the mPFC during reward seeking, Peters et al., 2005 have already shown the existence of groups of neurons which will repeatedly activate in a reproducible pattern during operant self-administration. Given what is known about the role of the IL during reward seeking (1.2.1), and the existence of neuronal tuning towards reward seeking in the mPFC in general, observing many neurons in the IL over longer time periods to characterize the stability of potential ensembles is crucial to better understand the role of IL during reward seeking.

1.4. Imaging of neuronal population activity

In order to be able to record the same neurons over long time periods, rather than using electrophysiology, the intracellular concentration of calcium ions can be used ($[Ca^{2+}]_i$) to infer activity levels. Monitoring the $[Ca^{2+}]_i$ can be achieved with two different classes of $[Ca^{2+}]_i$ indicators. First, chemical indicators like fura-2, indo-1 or fluo-3 and second genetically encoded indicators (GECIs) like the GCaMP's. While chemical indicators were first available to monitor neuronal activity (Koester & Sakmann, 2000, Helmchen & Sakmann, 1996) GECIs largely replaced chemical indicators, mostly due to limitations in the control of their cellular localisation and also the targeted cell. Chemical indicators also only allow $[Ca^{2+}]_i$ imaging for a limited time window

before they are removed from the cells again. One of their benefits however is the broad range of Ca^{2+} affinities for which different indicators are available (Paredes, Etzler, Watts, & Lechleiter, 2009).

The development of GECIs brought many advantages over the formerly used chemical indicators. They can be delivered specifically to various cell types with temporal control and show, dependent on the type of promoter and virus serotype used, constant expression over the remaining live span of the organism (Hammond, Leek, Richman, & Tjalkens, 2017; Walter, You, Hagstrom, Sands, & High, 1996). Thus they can be used to monitor the same cells over weeks at a time.

While various single fluorophore indicators have been developed, e.g. Camgaroo (Baird, Zacharias, & Tsien, 1999), pericam (Nagai, Sawano, Eun Sun Park, & Miyawaki, 2001) or jRCaMP/jRGECO (Dana et al., 2016), by far the most prominently used one today is the GCaMP family (Nakai, Ohkura, & Imoto, 2001), with the latest generation of jGCaMP7 (Dana et al., 2019) and its predecessor GCaMP6 which was used in this study (Chen et al., 2013), both of which have been optimized through mutagenesis screenings.

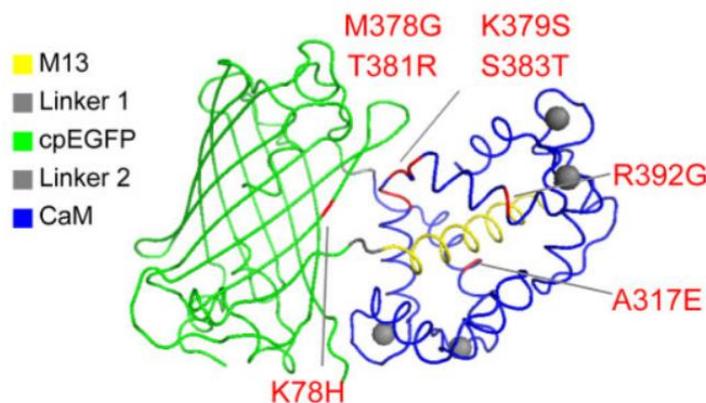


Figure 4: GCaMP structure and its mutation sites. The sites shown in red highlight mutation sites between GCaMP6 and GCaMP 5. Image from Chen et al., 2013.

The GCaMP variants are all based on a circularly permuted green fluorescent protein (GFP), a calcium binding protein called calmodulin (CaM) and a CaM interacting M13 peptide from the myosin light chain kinase (Chen et al., 2013). Changes in the calcium concentration will lead to conformational changes of the CaM/M13 complex and will cause increases (if $[\text{Ca}^{2+}]_i$ is increased) or decreases (if $[\text{Ca}^{2+}]_i$ is decreased) in its brightness.

There are a number of notable variants in both generations of GCaMP6 and GCaMP7. GCaMP6 introduced three variations called s, m and f, with the s variant showing the largest fluorescence transients, but longer decay times and the m, and f variants showing lower spike detection rates,

however with faster kinetics (Chen et al., 2013). The novel jRCaMP7 variants expanded on this concept and aimed to provide calcium sensors specifically targeted for various purposes. Again, there are a s and an f version which are either providing a high signal to noise ratio (SNR), but slower kinetics (s version), or the other way around in case of the f version (Dana et al., 2019). In addition, the c and the b variant were introduced, with the specific purpose to provide either a high (b) or a very low (c) background fluorescence. The former was designed with the imaging of very small structures, like dendrites, in mind, while the latter is supposed to decrease background fluorescence for wide field applications, like epifluorescence microscopy, which can suffer from out of focus fluorescence (Aharoni & Hoogland, 2019; Dana et al., 2019).

1.4.1. Types of cellular *in-vivo* Calcium imaging

The described $[Ca^{2+}]_i$ indicators have been used excessively in monitoring neurons *in vitro* and *in vivo* (Cai et al., 2016; Lütcke et al., 2010). Monitoring $[Ca^{2+}]_i$ activity of neurons *in vivo* is a demanding task relying on modern imaging techniques and different approaches have been developed. The most important aspects in achieving *in vivo* calcium imaging is the type of microscope used as this allows and limits application options.

Because of limitations intrinsic to confocal scanning and wide field epifluorescence microscopy, like light scattering and reduced imaging depth, 2-photon microscopy was the first technique that allowed *in vivo* $[Ca^{2+}]_i$ imaging of neural tissue (W. Denk et al., 1994; Winfried Denk, Strickler, & Webb, 1990; Helmchen & Denk, 2005). 2-photon-excitation was first used because it has multiple benefits over conventional confocal scanning microscopy and wide field fluorescence microscopy. The high spectrum of the excitation wavelength limits phototoxicity, allows comparably deeper access into the tissue and drastically decreases out of focus fluorescence (Carrillo-Reid, Yang, Kang Miller, Peterka, & Yuste, 2017). Its caveats however are, the complex technology and corresponding high cost that is involved, its ultimately limited depth within the tissue and, for now, the need to perform most behavioural imaging tasks in awake animals using head-fixed configurations. In contrast epifluorescence wide field microscopy allows utilization of head-mounted miniature that allow *in vivo* imaging in freely moving animals. In the following paragraphs, miniaturized epifluorescence microscopy is discussed, which was increasingly used in the last years to perform $[Ca^{2+}]_i$ imaging in areas of the brain that are out of reach for 2-photon imaging.

1.4.2. Miniaturized epifluorescence microscopes

Advances in complementary metal-oxide semiconductor technology (CMOS) and the improvement of GRIN lens manufacturing allowed the development of head mountable epifluorescence microscopes (Ghosh et al., 2011). Today there is a multitude of different head mountable microscope systems available, both commercial as well as open-source (self-assembly). These microscope systems are capable of dual colour imaging (Jacob et al., 2018), simultaneous manipulation of neuronal populations (Doric system, nVoke), imaging of volumes at high frame rates (Skocek et al., 2018) and more (Aharoni & Hoogland, 2019). Currently two commercial systems are available (Doric, Inscopix) and a number of open source versions with the first that became available being the UCLA miniscope (Cai et al., 2016) and others like the CHEndoscope (Jacob et al., 2018), FinchScope (Liberti, Perkins, Leman, & Gardner, 2017) and miniScope (Zhang et al., 2019) which followed.

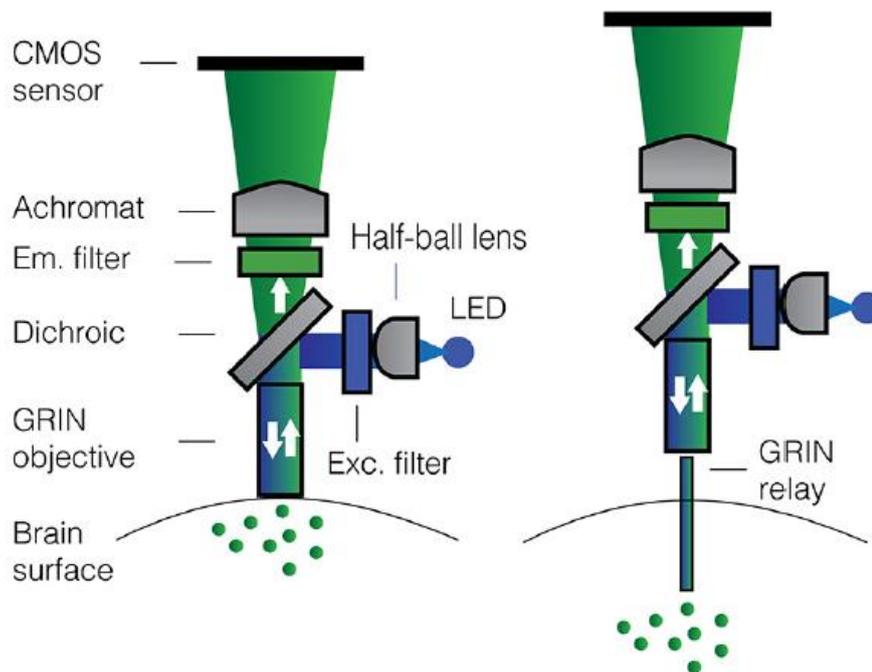


Figure 5: Typical design and application of miniaturized epifluorescence microscopes. Image from (Aharoni & Hoogland, 2019).

All miniaturized epifluorescence microscopes mostly follow the same design principles. The light of an LED (usually $\sim 470\text{nm}$) is collimated through a half-ball lens and then passes an excitation band pass filter (e.g. $470\pm 20\text{ nm}$ for UCLA Miniscopes). A dichroic mirror (e.g. 495 nm longpass for UCLA Miniscopes) then directs the light through a GRIN lens which serves as an objective and focusses the light into the specimen. Emission light is directed through the GRIN objective and the dichroic mirror and then passes an emission filter to block photons outside of the emission

spectrum (e.g. 525±25 nm for UCLA Miniscopes). An achromatic lens then focusses the collimated beam from the GRIN objective onto the CMOS sensor (Figure 5)

Two different types of imaging modalities can be distinguished. The first uses the GRIN objective lens for superficial imaging of either the cortex, or of the hippocampus (Cai et al., 2016), in preparations where the GRIN lens was implanted after removal of the cortex. Because of its design principles, there are boundaries to how small the objective GRIN lens can be manufactured, limiting its application for implantations if one wants to avoid removing too much neural tissue. To overcome this limitation, couplings with smaller GRIN lenses have been used to allow imaging of deeper structures. The smaller GRIN lenses have working distances which are close to the surface of the lenses, allowing the imaging in deep structures with epifluorescence illumination.

While the use of epifluorescence illumination has issues regarding photo bleaching, potential photo toxicity if the excitation power is too high (Carrillo-Reid et al., 2017) and only provides limited spatial resolution, it provides many benefits. Among those currently limited to epifluorescence applications are the possibility to broadly use behavioural paradigms which are not restricted to head fixed applications and most importantly its cost effectiveness (Aharoni & Hoogland, 2019). Given the simplicity of the technology and reliance on off-the-shelf components, most laboratories are able to afford using the technique. Hence for applications that do not require high resolution imaging of structures below the size of somata, but do require not head-fixed behavioural experiments, miniaturized epifluorescence microscopy provides an ideal solution.

1.5. Analysis of *in-vivo* epifluorescence calcium imaging

The analysis of $[Ca^{2+}]_i$ imaging data has always been a bottleneck, and usually, as soon as major challenges have been solved, new technological breakthroughs of the imaging result in new challenges for the analysis. E.g. after many problems like image stabilisation and automated identification of neurons had been more or less solved for 2-photon imaging data, the emergence of miniaturized epifluorescence microscopy required many adaptations of the existing workflows.

The analysis of *in-vivo* epifluorescence calcium imaging datasets can be conceptually split into multiple steps, each either correcting or further condensing the original imaging data. The first step following image acquisition and the storage in an appropriate format (e.g memory maps) is motion correction to account for movements of the animal during the acquisition. This is then followed by a source extraction to identify the spatial footprint and temporal calcium dynamic of each neuron. This can be further processed into estimations of spike trains. Lastly, the neuronal

signals are either used by themselves to infer e.g. ensemble activity or in-, decreases in synchronicity, or the signal are related to behavioural information. Because many available software packages focus on motion correction and source extraction, the following paragraphs are split between all steps required for the extraction of $[Ca^{2+}]$ transients and then the approaches used for high level analysis.

1.5.1. Extraction of calcium transients

In order to extract $[Ca^{2+}]$ transients, the same area of the tissue should be captured by the same pixel of the final image throughout the whole recording. Motion correction algorithms correct such movements and can be separated into two categories. The first, rigid motion correction, only corrects any linear displacements of the two axes of the image. Hence, in cases where tissue deformations due to excessive movement of the animal led to morphing of the tissue, a rigid approach may not be able to fully remove all artefacts (Pnevmatikakis, 2019). A rigid correction however, is less computationally expensive and approaches utilizing a dual approach have been implemented (Lu et al., 2018). The second approach is non-rigid motion correction, which aims to correct motion artefacts in separate areas of the image independently (Greenberg & Kerr, 2009; Pnevmatikakis & Giovannucci, 2017).

The next step is the source extraction of calcium traces. The result is generally a set of spatial footprints (2D), and $[Ca^{2+}]$ transients (1D) for each neuron. A particular problem of epifluorescence imaging is the relatively poor spatial resolution and the contamination with out of focus light caused by the large light excitation volume. This leads to mixing of signals from multiple light sources (e.g. overlapping neurons, neuropil) in each individual pixel. While methods exist which employ 'classical' segmentation approaches of the neuronal footprints and then extract the $[Ca^{2+}]$ transient through a spatial weighted average of the individual pixels (Chen et al., 2013), matrix-factorisation approaches have been most commonly used recently, because of their better performance.

Matrix-factorisation approaches assume, that the fluorescence of each neuron is represented as an active signal, and that each neuronal signal can be represented by a vector in space (its spatial footprint) and a vector in time (its $[Ca^{2+}]$ transient), with the original set of images representing the product of those two vectors. This concept was first implemented through a combination approach of principal and independent component analysis (PCA/ICA, Mukamel, Nimmerjahn, & Schnitzer, 2009) and then extended through non-negative matrix factorization approaches (Diego & Hamprecht, 2014; Maruyama et al., 2014), for which a number of open source pipelines exist today. Some of the available pipelines have to be combined because they either only correct

motion artefacts (NoRMCorre, Pnevmatikakis & Giovannucci, 2017) or extract $[Ca^{2+}]_i$ traces (CNMF-E, Neufeld et al., 2018), while others contain modules for both motion correction, as well as signal extraction (Caiman and Min1pipe, Giovannucci et al., 2019; Lu et al., 2018).

1.5.2. High level analysis

After the extraction of calcium traces, the number of possibilities for further analysis are endless and should be carefully adapted to the hypothesis and the amount of meta information which is available besides the calcium transients. These could include behavioural information (potentially multidimensional), information about the neuronal cell type, including potential projections to other areas or even information about perturbations during the recordings.

An important aspect when imaging neurons over the course of weeks and therefore likely multiple sessions, is to be able to identify the same neurons in each of the sessions. Among recent publications (Cai et al., 2016; Corder et al., 2019) which solved this issue independently, Sheintuch et al., 2017 released an analysis pipeline to precisely match overlapping neurons between sessions with high accuracy (2.8.6).

Analysis approaches also exist to analyse calcium transients without external information. One such approach aims to quantify the level of co-activation between neurons to compare healthy and disease states (Hamm, Peterka, Gogos, & Yuste, 2017). Another approach, limiting the initial requirement for external information is a pipeline called LeMoNaDe by Kirschbaum et al., 2018 which aims to identify neuronal ensembles based on the raw imaging data (without calcium trace extraction) through the use of variational autoencoders.

As described, there are a number of options available to perform and analyse *in vivo* calcium imaging in freely moving animals. These options have to be carefully chosen in accordance to the experimental questions. In this study, the focus was on the neural representation of reward seeking behaviour during different stages of the operant self-administration. Hence the collected information included signals from the lever and the head entry port with which the animal could interact. Also it was crucial to image the same region during all of the stages of the behavioural paradigm to be able to track the activity of the same neurons across the sessions.

1.6. Objectives

In order to understand how neural circuits control reward seeking, it is essential to monitor cellular activity during reward seeking behaviour. Novel methods allow the monitoring of large

numbers of neurons using genetically encoded calcium imaging over the time span of multiple weeks in freely moving animals (Ghosh et al., 2011, Cai et al., 2016). Based on previous studies it is known that subsets of neurons in the IL of the mPFC are active during operant self-administration (Pfarr et al., 2015), and that neurons in the mPFC in general are active at specific events during reward seeking behavior (Peters et al., 2005). Since the precise dynamics and stability of the neuronal ensemble that is active during reward seeking behavior remains unknown, the aim of this study was to elucidate ensemble dynamics during multiple operant self-administration sessions. The goal is to identify neurons that are active during reward seeking, determine if the activity is time locked to behavioral events and track the activity of these neuronal ensembles through multiple SA sessions. This work shows the first exemplary approach towards the dissection of neural networks in the IL of the medial prefrontal cortex (mPFC), essential in the formation of reward seeking behaviour by using an operant self-administration paradigm in rats combined with miniaturized epifluorescence microscopy.

2. Materials and Methods

2.1. Miniature fluorescence microscope: UCLA Miniscope

UCLA Miniscopes were manufactured based on the descriptions of the miniscope.org webpage. The CMOS imaging printed circuit boards (PCB) and the data acquisition (DAQ) PCBs were obtained from SierraCircuits based on the designs the developers released on the GitHub repository (PCB and DAQ version 3.2 (Aharoni, 2016d, 2016c)). Later additional CMOS and DAQ PCBs were bought from labmaker (Berlin, Germany). The housing (focus slider, main body and filter cover, all black Delrin) as well as the baseplates (aluminium) and the baseplate-covers (black Delrin) were manufactured by the local machine-workshop and were also based on the UCLA Miniscope design (Aharoni, 2016g). Achromatic and half-ball lenses were bought from Edmund optics (5x15 mm achromatic lens, ID:45207; 5x12.5 mm achromatic lens, ID:49923; 5x7.5 mm achromatic lens, ID:45407; half ball lens 3 mm, 47269). GRIN lenses were bought from Inscopix. The circuit boards for connecting the emission diode and the coaxial cable connector were obtained from SilverCircuits (Houston, USA) and based on the designs the developers released on the GitHub repository (Aharoni, 2016g). The emission LEDs and miscellaneous electronics and items were bought from Digikey (Thief River Falls, USA; LED Luxeon Rebel Blue SMD, ID: 1416 1028 1 ND; SMA Connector 50 Ohm, ID: consma013.062; Light Pipe clear 3mm, ID: VLP 550 F; IC Eeprom 1Mbit 400Khz, ID: 24AA1025 I/P; shunt jumper .1", ID: 969102 0000 DA; IC DIP socket 8pos, ID: DILB8P 223TLF; SMA jack R/A 50 Ohm, ID: consma002). Optic filters were bought from Chroma (Bellows Falls, USA; excitation filter, ET470/40x 3.5x4x1mm, ID: IN054535; emission filter, ET525/50x 4x4x1mm, ID: IN054538; dichroic mirror, T495lpxr 4x6x1mm, ID: IN054536). Magnets for the baseplates and main bodies were purchased from KJ Magnetics (Pipersville, USA; 1/16" dia. x 1/32" thick, ID: D101-N52). The housing for the DAQ boards was 3D printed on an Ultimaker 2 using 3 mm thick, white PLA at 100µm z-resolution (Aharoni, 2016e). Two different laptops (Dell Inspiron 15, MacBook Pro 15, 2017) were used for the recordings. Both had windows 7 installed as the operating system and the recording software (Aharoni, 2016f) supplied by the UCLA Miniscope team was used for calcium imaging.

2.2. Animals

Wistar rats were ordered from Charles River and kept at an inverted 12-hour dark/light cycle with access to food and water ad libitum. At the start of the experiments animals were 8 weeks old. All procedures were approved by the local animal welfare body (Regierungspräsidium Karlsruhe, G-221/15, In-vivo Mikroendoskopie im Gehirn von Ratten).

2.3. GRIN lens implantation

Animals were anesthetized with Isoflurane (1.5 - 3%, Univentor 1200 - anaesthesia Unit, Univentor, Zejtun, Malta) and head-fixed in a Kopf Stereotax frame (Model 900, David Kopf Instruments, Tujunga, USA). To promote survival of the animals during the surgical procedure, a constant supply of oxygen was supplied to the inhalation gas. Occasional casualties were restricted to the time during or after the insertion of the GRIN lens. This could be solved by the injection of a small dose of dexamethasone to reduce the intracranial pressure (Savard, Lema, Hélie, & Vachon, 2009) caused by the inhalation gas (Hockel, Trabold, Schöller, Török, & Plesnila, 2012) as well as the insertion of the GRIN lens (1 mm x 9 mm, Inscopix, Palo Alto). The head was shaved to expose the skin and then disinfected with ethanol swabs. All surgery equipment was sterilized using a bead sterilizer during the surgery (Steri 250, Simon Keller AG). The skin was cut along the anterior-posterior axis and the connective tissue above the cranium was gently removed with sterile cotton swabs to expose the bone. Care was taken to not damage the periosteum or the bone to avoid degeneration of bone-tissue and subsequent loss of implants. Three small craniectomies were drilled using a dental drill and Fine Science Tools (FST, Foster City, USA) bone screws were fixed to the skull. Then the skull was adjusted along the anterior-posterior axis (AP-axis) to be parallel to the AP-axis of the stereotaxic frame. A craniectomy (~2mm Ø) was made 3 mm anterior and 0.5 mm medial to bregma. A pulled glass injection pipette (Puller: Model P-97, Sutter Instrument Co., Novato, USA; Pipette: Blaubrand intraMark micropipettes, product ID 708707) was then loaded with a 1:20 dilution (PBS) of AAV1.Syn.GCaMP6f.WPRE.SV40 (UPenn Vector Core, Philadelphia, USA, Lot#: CS1107, titer: 2.13×10^{13} GC/ml) using a syringe and the pipette was lowered to 5 mm ventral to bregma in two positions around the center of the craniectomy (+200µm AP and ML). To lower the pipette into the brain, a modified Luigs&Neumann linear actuator was used (2.4). 500 nl of AVV was then slowly injected at a speed of 100 nl/min. After the injection was finished, the pipette was left in place for at least 3 minutes to allow the pressure to equilibrate. Then the pipette was slowly removed. After the AAV injection, the skull was cleaned with saline solution (NaCl 0.9%, B. Braun). To open up the tissue for the insertion of the GRIN lens, a glass fiber (Ø 127µm) that was threaded into a hollow pipette tip and fixed with super glue (super glue: Pattex, Düsseldorf, Germany, flüssig; pipette: Kwik-Fil, borosilicate glass capillaries, 1B150F-4) was attached to the stereotaxic frame and slowly lowered into the brain using the modified Luigs&Neumann linear actuator (2.4) to a depth of 4.7 mm from the surface of the skull. The glass fiber was inserted into the tissue over a span of approximately 30 minutes by repeatedly lowering the fiber 300 µm into the tissue and then raising it again 200 µm to avoid building up too much pressure. Once the fiber was within 1

mm of the target depth, the alternation rate was changed to lowering of 30 μm followed by raises of 20 μm . After the fiber reached the target depth of 300 μm above the injection site (4.7 mm) it was left in place for 2-3 minutes and then slowly (30 $\mu\text{m/s}$) removed from the brain using the linear actuator. After removal of the fiber, the skull was again cleaned with saline. The GRIN lenses (1 mm x 9 mm, Inscopix, Palo Alto) were attached to a custom built holder (Cambridge Neurotech, Cambridge, UK) that was attached to the linear actuator. Vacuum grease was attached to the top of the lens to make sure it could not drop out of the holder. The bottom of the lens was carefully inspected for cracks and imperfections and then carefully cleaned with lens paper and ethanol (Lens cleaning tissue, GE Healthcare Life Sciences). The lens was then lowered above the craniectomy and the protocol for lowering the linear actuator was repeated until the lens was lowered to approximately 300 μm above the injection site. After the lens reached its position, the area around the craniotomy was carefully cleaned from any blood and the skull was carefully dried with cotton tissue swabs. Great care was taken to not touch the lens, its holder or move the cranium in any way. To attach the lens to the cranium, the tip of a 30G needle was crushed with forceps until it resembled a small spoon. Skin glue (N-Butyl cyanoacrylate, Surgibond, Cisek, Poland) was then applied to the tip until a small drop formed. The surface of this drop was then carefully lowered towards the lens until it touched the lens. Care was taken to neither touch the lens holder, nor the skull first. The former would lead to the attachment of the lens to the holder, while the latter could result in the formation of a thin film of polymerized glue on its surface, due to liquid coming out of the craniectomy, which results in an incomplete attachment of the lens to the skull. After the skin glue visibly polymerized the lens holder would be slowly raised. The z-axis of the stereotaxic frame was zeroed for this to ensure the lens could be lowered back to the original position in case of an incomplete attachment to the skull. After the successful removal of the lens holder a layer of OptiBond™ FL Kit (Kerr, Bioggio, Switzerland) was applied on the skull according to the manufacturer's instructions. This was done to improve adhesion of the implant to the skull. Next, multiple layers of dental cement were applied around the GRIN lens to protect it and to build the foundation for the attachment of the baseplate. For this, dental cement liquid (Paladur, Henry Schein, Gallin, Germany) and black dental powder (Contemporary Ortho-Jet™, black, Lang Dental, Wheeling USA) were mixed and then slowly applied around the lens. Care was taken to not attach any cement to the skin and the surface area of the cement base was kept at minimum at about the size of the baseplates. The cement was applied until about 1 mm of the GRIN lens was left sticking out of the surface and the area around the lens was gently pressed down to create a flat surface. In order to keep the animals from breaking the GRIN lenses, a circle of cement was raised around the lens at a minimum distance of at least 1 mm. After the circle was raised above the lens, it was covered in more vacuum grease, followed by a few layers of lens paper to fill the 'cavity' surrounding the lens. The top was then covered in dental cement. Care

was taken to ensure that no cement touches the lens, or gets too close to it, to avoid its destruction during the process of recovering it for the baseplate surgery. The skin was then disinfected with iodine (7.5%, Baunol, B. Braun, Melsungen, Germany) solution and the anterior and the posterior areas of the cut were sutured to close the wound and avoid exposure of the cranium.

GRIN lens implantations were performed with assistance of Simone Pfarr and Janet Barroso-Flores.

2.4. Modified Luigs & Neumann linear actuator

To enable the slow and reproducible insertion of the GRIN lenses, a Luigs&Neumann linear actuator (mini25, Luigs&Neumann, Ratingen, Germany) was coupled to a commonly available controller unit for a 3D printer (RAMPS 1.4, RepRap, reppap.org). The linear actuator contains a P430 stepper motor and has a SUB-D 9 plug attached to it which can be used to access the 4 phases of the stepper motor and both of the end-stops and their common ground. The spindle gradient is 0.5 mm per turn and the motor has 100 full-steps per turn, yielding a 5 μm resolution for full-steps. To control the linear actuator, the RAMPS-1.4 shield for the Arduino-Mega was used. The Marlin firmware ("Marlin Firmware," 2015, version 1.1) was used and changed to account for the spindle gradient and number of steps the motor provides. To control the linear actuator through the RAMPS-1.4 board the software Pronterface (Printrun, 2015) was used. The stepper motor was connected to the y-axis controller on the RAMPS-1.4 board. Pronterface allows the direct manipulation of the linear actuator through g-code commands or by executing a g-code script. In order to automatically insert the lenses, a custom Matlab script was written, which creates a g-code file that contains a list of all g-code commands. This file can then be run by Pronterface to execute the commands.

Command	Function
G1 <i>xnnn ynnn znnn</i>	This command will move the specified axis (x/y/z in any combination) to its absolute position <i>nnn</i> in mm.
G92 <i>xnnn ynnn znnn</i>	This command will set the position of the specified axis (x/y/z in any combination) to the specified position <i>nnn</i> in mm.
M114	This command displays the current absolute position of the axis.

Table 1: Relevant g-code commands.

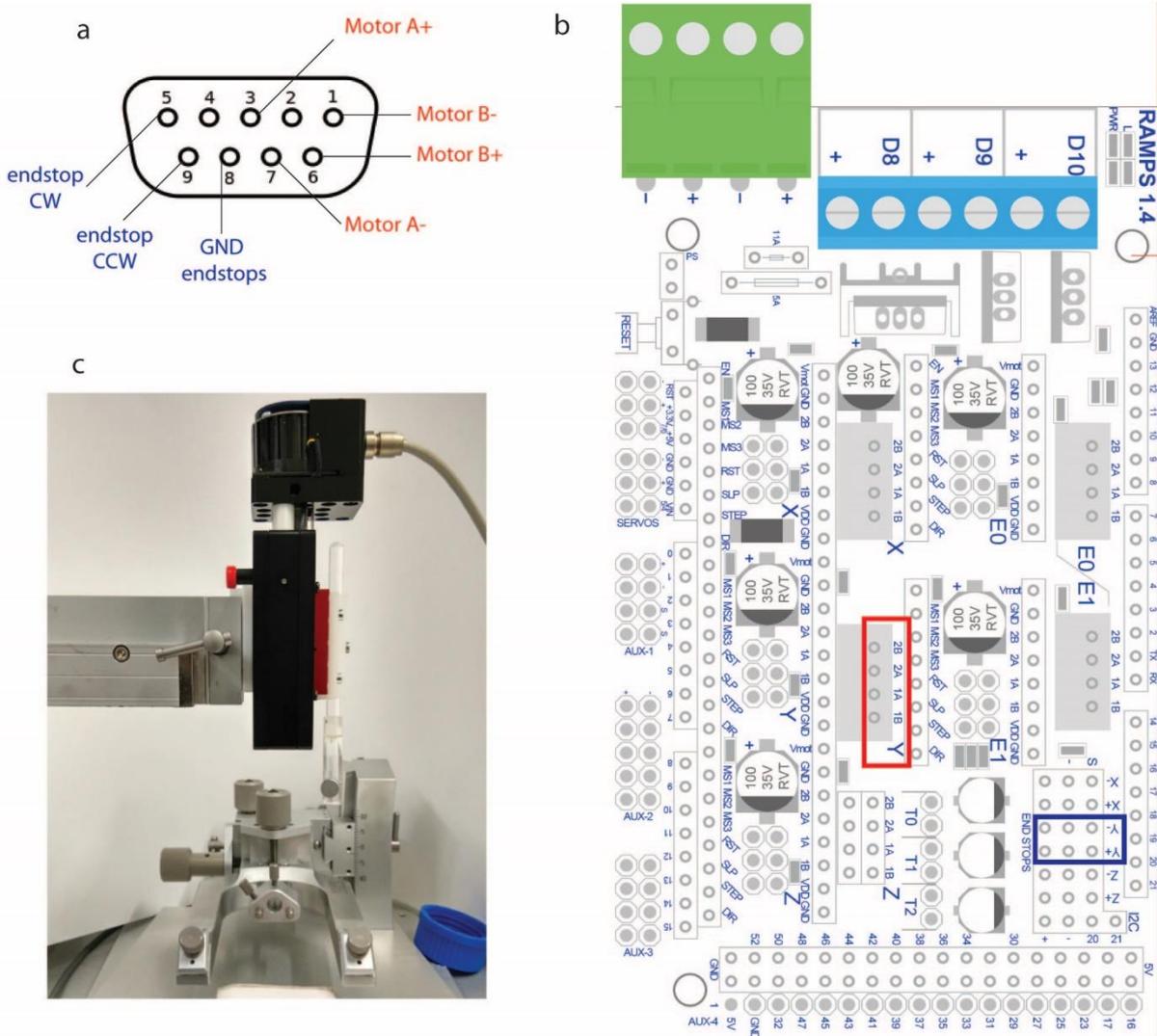


Figure 6: Luigs&Neumann linear actuator and RAMPS 1.4 board.

2.5. Fixation of baseplate for miniaturized fluorescence microscope

Baseplates for the UCLA miniscope system were equipped with magnets and a set screw to allow a tight connection to the miniaturized epifluorescence microscopes. Before the microscope was attached to the stereotax, the focus was centered to its mid-range to allow adjustments in both directions after the fixation of the baseplate. Animals were anesthetized with isoflurane (1.5 – 3%, Univentor 1200, anaesthesia Unit, Univentor, Zejtun, Malta) and head-fixed in a stereotaxic alignment system (Kopf Stereotax, model 900, Tujunga, USA). The protective cement cover of the implanted GRIN lens was carefully removed using a dental drill and the lens paper in the cement cavity was carefully removed with forceps. Care has to be taken to avoid scratching the surface of the GRIN lens. After the surrounding cement was drilled down to the level of the lens and all the lens paper was removed, the lens was cleaned with lens paper and ethanol. Care was taken to

remove all remaining vacuum grease. In case of contaminations of the lens with dental cement, acetone was used to dissolve the cement and clean the surface of the lens. Care has to be taken to not put acetone in contact with the wound or the skin of the animal. Once the lens was clean and could be reached with the objective GRIN lens, the miniaturized microscope, attached to the stereotaxic frame by a custom built holder, with the attached baseplate was lowered down to the implanted GRIN lens. The miniaturized microscope was attached to a PC running Windows 7 (Dell, Inspiron 15) through the DAQ board and the software supplied by the developers (Aharoni, 2016f) was running to provide visual feedback on the field of view (FOV) of the microscope. Once the GRIN lens was within the focal distance of the microscope the imaging parameters were initially adjusted to optimize the range of the histogram and facilitate the process of focusing. The image repetition rate was generally set to 20 Hz (corresponding to exposure times of approximately 50 ms per frame), except for cases of weak calcium signals. The gain was always set to the maximum value and the LED power was adjusted until the background was visible, but no over-saturation occurred. Because of the anaesthesia, it was rarely possible to see calcium transients of individual cells. To find an appropriate field of view for the recordings, the bottom of the GRIN lens was brought into focus (close to the top of the GRIN lens). From there the microscope was raised until no more scarring on the bottom of the lens was visible, but there were still blood vessels in focus. If a satisfying focal plane was identified, dental cement liquid (Paladur, Kulzer Mitsui Chemicals Group, Hanau, Germany) and black dental powder (Contemporary Ortho-Jet™, black, Lang Dental, Wheeling USA) were mixed and the four corners of the baseplate were attached to the implant with the fresh cement. Care was taken to avoid using cement that was too liquid to avoid the cement 'running' below the baseplate and covering the lenses. After the four corners of the baseplate were attached to the implant, the remaining sides were covered in cement. To avoid detachment of the aluminium baseplates from the cement, super glue (Pattex, flüssig, Düsseldorf, Germany) was carefully applied to the edges between the aluminium baseplates and the dental cement. Care was taken to avoid getting super glue between the baseplate and the microscope to avoid permanent attachment of the two. After the baseplate was cemented in place, the microscope set screw was carefully opened and the microscope could be carefully removed. The baseplate was then covered with a baseplate cover (Aharoni, 2016g) (UCLA, Zentralbereich Neuenheimer Feld, machine shop) to protect the lens from dirt.

2.6. Behavioural procedures

2.6.1. Operant self-administration hardware

All experiments were performed in operant chambers (MED Associates chamber, ID: MED-008-CT-B4) enclosed in ventilated sound-attenuating cubicles. The chambers were equipped with a response lever on each side of the chamber. Responses at the appropriate lever activated a dipper cup in which the reward was presented inside of a head entry port. A light stimulus was placed above the right response lever of the self-administration chamber. An IBM-compatible computer controlled the delivery of fluids, presentation of stimuli, and data recording. TTL pulses from the operant chamber for the levers and head entry ports on both sides, as well as the cue light for the rewarded side were acquired/digitized using an Arduino Mega running the Firmata (Firmata, 2015) firmware, along with the frame trigger of the DAQ board of the miniaturized microscope. The data that was received by the Arduino was then saved using Bonsai (Lopes et al., 2015) as .csv file.

2.6.2. Operant self-administration

Saccharin self-administration training and testing sessions were performed 3 h after beginning of the dark phase, 5–6 days per week. Animals were trained to self-administer 0.2% (w/v) saccharin in daily 30 min sessions on a fixed-ratio 1 schedule using a saccharin fading procedure modified from Samson, Pfeffer, & Tolliver, 1988. During the first 3 days of training, the animals were kept water deprived for 20 hours per day. Responses at the right lever were reinforced by the delivery of 0.2% (w/v) saccharin solution. For the remaining experiments the animals underwent the same procedure without water deprivation. Responses at the left lever were recorded but had no consequences. A (visual) cue was used for conditioning. The discrete visual stimulus was presented after correct responses resulting in saccharin delivery (right lever). As visual stimulus, a 5 s blinking light was used, which was activated after a response at the active lever and was therefore directly connected to saccharin availability. The 5 s period served as a “time out”, during which responses were recorded, but did not lead to an additional reward delivery (conditioned stimulus, 1st 1.5s, 2nd 1s, 3rd 1s; 1s off in between pulses). For the stimulus-conditioning training, the animals had to complete 10 sessions of 30 minutes, with the conditioned stimulus (CS), without simultaneous calcium imaging, and after baseplate fixation enough sessions to reach previous lever response levels with simultaneous calcium imaging. A Med Associates Super Port Input module (ID: DIG-712) was used to relay TTL pulses indicating behaviourally relevant events in the chamber (lever press left/right, head entry left/right, cue

light left/right). Upon pressing the rewarded lever (always the right side), the dipper cup was placed in the liquid and the reward was presented in the head entry port.

The pre-training was done by Simone Pfarr, Janet Barroso-Flores, Laura Schaaf and Rebecca Hoffmann.

2.6.3. Extinction

After successful completion of the stimulus conditioning phase all animals underwent 4-5 daily, 30 min extinction sessions, which was sufficient to reach an extinction criterion of 25% of the baseline activity at the active lever per session. During extinction sessions, both levers were extended. Responses at the previously active lever activated the dipper cup, which did not result in reward delivery or presentation of the discrete CS (blinking light). The TTL-pulse was still transmitted to aid the post-hoc analysis.

2.6.4. Cue-induced reinstatement

For the cue-induced reinstatement sessions, both levers were extended during the sessions and the animals were presented with the same conditioned stimuli (CS) as during the conditioning phase. After pressing the rewarded lever, the dipper cup would move; however, no reward was presented in the head-entry port.

2.7. Histology

The animals were deeply anaesthetized with isoflurane and transcardially perfused with 100 ml 1xPBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 1.46 mM KH₂PO₄, pH 7.4), containing 10000 units of Heparin sodium/l. Next, the animals were perfused with 50 ml of 4% paraformaldehyde (PFA) in 1xPBS solution (pH 7.4). The complete heads of the animals were post fixed in fixative solution for at least 7 days at 4 °C. After removal from the skull, brains were stored in 4 °C PBS.

Perfusions were performed by Simone Pfarr and Janet Barroso-Flores.

2.7.1. Thin sectioning and mounting

After at least one week of incubation in fixation solution at 4 °C the brains were dissected from the skulls. The GRIN lenses were removed along with the top of the skull before dissection. Care was taken to avoid damaging the tissue with the implanted GRIN lenses. After the dissection the

brains were rinsed 2x in fresh PBS. Then the brains were cut in half along the anterior-posterior axis to separate the hemispheres. The side which contained the GRIN lens was then cut a second time approximately 3 mm posterior to the entry site of the GRIN lens and the anterior half was mounted in a vibratome Slicer (Leica VT1000S) and cut into 100 μm thick sections. The sections were collected in 24 well plates containing PBS with DAPI (#D9542-10MG, final concentration 0.05-0.1 mg/ml, Sigma-Aldrich, St. Louis, USA). Those slices close to the area which contained the GRIN lens were mounted on 24 mm x 70 mm glass specimen slides (Marienfeld, Lauda-Königshofen, Germany), covered with SlowFade (#S36936, Thermo Fisher Scientific, Waltham, Massachusetts) and a coverslip (24 mm x 60 mm, Carl Roth, Karlsruhe, Germany) and then sealed with nail polish.

2.7.2. Image acquisition of fixed tissue sections

Overview images were acquired using a Leica DM6000 upright, fluorescence microscope, equipped with a 1.25X/0.04 NA (Leica Microsystems, #11506215) and a 10X/0.4 NA (Leica Microsystems, #11506284) objective, and using the LasX acquisition software. The emission light-path contained an additional magnification step of 1.2x to account for the size of the camera. High-resolution close-up images were acquired using a Leica SP8 confocal microscope, equipped with a 10X/0.4NA (Leica Microsystems, #11506293) objective, using the same software. The image resolution was kept at 1024x1024 pixels, while the scan speed was set to 200 Hz with solely unidirectional scanning. All fluorophores were imaged sequentially. Image processing was performed using Fiji (Rueden et al., 2017).

2.8. Analysis

The analysis pipeline was designed by Ivo Sonntag and written by Ivo Sonntag and Ariel Iporre-Rivas, with help of Janet Barroso-Flores. The code is available at request from Ivo Sonntag.

2.8.1. List of MATLAB classes used

<i>Class name</i>	<i>Class description</i>
Miniscope	The Miniscope class provides basic access to the metadata that is saved during each recording session. This includes the imaging parameters (LED-power, gain) as well as the time stamps for all frames that were saved on the hard-drive (essential for the sequence adjustment).
Bonsai	The Bonsai class parses the .csv files saved by the Bonsai software. It contains information about each of the data channels (frame trigger, right and left head entry, left and right lever, cue light) as well as a timestamp (μs). The time stamps in combination with the frame trigger information are used by the SequenceAdjustment class to synchronize the behaviour and the calcium imaging datasets.

Table 2: Basic classes that can be used independently of the rest of the pipeline to get fast and easy access to the information about the calcium imaging or the behaviour sessions.

<i>Class name</i>	<i>Class description</i>
SequenceAdjustment (SeqAd)	The SequenceAdjustment class synchronizes the behaviour and calcium imaging dataset using the time stamps for each frame and behaviour data point that was recorded and the frame trigger that is linked to the behaviour information through Bonsai. For each frame of the calcium imaging it provides the corresponding Bonsai data points and vice versa. Its inputs are the Bonsai and the Miniscope object for the respective recording session, as well as the calcium traces.
PatternDetector (PD)	The PatternDetector class can be used to categorize the behaviour based on the properties that are assigned to each lever press and head entry through the Tag classes. Its input is a SequenceAdjustment object

	of a recording session. It contains a list of LeverTag and HeadEntryTag objects for every LP and HE of its recording session.
SignalTag	The SignalTag class is used to bind the LeverTag and HeadEntryTag class to the PatternDetector class.
LeverTag	The LeverTag class is used to describe the properties of an individual LP. These include whether the LP was rewarded (boolean), whether the reward was collected (boolean), how much time the animal took to enter the HE port (double), how much time passed to the previous and next rLP (double) and the index within the behaviour datastream (double), which can be used to link the time point of the LP to a frame of the calcium recording.
HeadEntryTag	The HeadEntryTag class is used to describe the properties of an individual HE. These include whether the HE took place during a cue light sequence (boolean), whether there should have been a reward in the HE port (boolean, 2.8.4), the duration of the HE (double), how much time passed to the previous and next rLP (double) and the index within the behaviour datastream (double), which can be used to link the time point of the HE to a frame of the calcium recording.
SessionBootstrap	The SessionBootstrap class is used to calculate into which percentile the variance of an averaged set of calcium transients, aligned to a rLP, falls within a distribution of randomly picked time points. It requires an Animal object as input.
Animal	The Animal class was used to combine all essential steps of the analysis for further processing. It provides methods to extract calcium traces which are aligned to distinct events in the behavioural recordings (PatternDetector, SequenceAdjustment) and matches the cells across sessions (CellRegWrap). It requires a list of PatternDetector objects as inputs.
CellRegWrap	The CellRegWrap is a wrapper class to provide fast access to the CellReg (Sheintuch, 2017; Sheintuch et al., 2017) pipeline (2.8.6)
SessionStitcher	The SessionStitcher class was used to correct recording sessions during which the calcium imaging was split into multiple recordings. It requires a Bonsai and a Miniscope object as input and a user input to define the time points at which the stitching should take place.

Table 3: Classes used to merge the behaviour and the calcium imaging datasets (SequenceAdjustment), classify the behaviour data into relevant time periods (PatternDetector),

match the cells between different sessions (CellRegWrap), organize the data to combine the processed data for individual animals (Animal) and run the bootstrapping (SessionBootstrap).

<i>Class name</i>	<i>Class description</i>
MemoryMap	The MemoryMap generates the memory maps and can be used to access them through the various processing stages. It requires the directory of the recording files as well as the dimensions of the recording as input.
MotionCorrector	The MotionCorrector applies the motion correction parameter set in a MotionCorrection object to a MemoryMap object.
MotionCorrection	The MotionCorrection object serves as wrapping class to streamline the process of first preparing the data and then applying a motion correction to it.

Table 4: The following classes were used to convert the .avi files of the initial UCLA Miniscope recordings into a motion corrected memory map for further processing by the CNMF-E pipeline.

2.8.2. List of external Matlab pipelines

Name	Description
CNMF-E (Neufeld et al., 2018; Zhou, 2016)	The CNMF-E pipeline extends the CNMF pipeline by adding support for single photon datasets. It provides the separation of individual signal sources in calcium imaging datasets into their spatial and temporal components (e.g. ‘footprints’ of neurons and their calcium traces).
Minipipe (Lu, 2018; Lu et al., 2018)	The Minipipe pipeline provides an all-in one solution for the motion correction and subsequent extraction of calcium traces for single photon datasets. It provides non-rigid motion correction and uses a CMF approach after a ‘neural enhancing’ step to remove background signal.
CellReg (Sheintuch, 2017; Sheintuch et al., 2017)	The CellReg pipeline uses a series of models to estimate the most likely combination of matching cell pairs in consecutive imaging sessions.

Table 5: List of external open-source analysis pipelines that were used.

2.8.3. $[Ca^{2+}]_i$ imaging data analysis

To extract calcium traces of individual cells from the *in vivo* imaging data, two different pipelines were used. The first is a custom built pipeline that combined multiple available open source repositories (Aharoni, 2016e; Neufeld et al., 2018; Zhou, 2016; Aharoni, 2016a) (as indicated) in combination with custom written software in Matlab. It is modularized in different Matlab classes and runs in batches on large recordings to reduce the amount of memory required. First .avi files of the miniscope software were memory mapped. The memory maps contain the raw as well as a filtered version of the data to enhance the background for motion correction. A rigid motion correction algorithm was then used to calculate the shifts in each image pair (Aharoni, 2016b). After the shifts are calculated, they can be applied to the raw data, which can be saved with both spatial and temporal down sampling to reduce memory demands during the calcium trace extraction. To extract calcium traces, a constrained non-negative matrix approach, optimized for microendoscopic imaging data, was used (Neufeld et al., 2018; Zhou, 2016). The constrained-non-negative matrix-factorisation (CNMF) approach aims to extract individual signal sources from the data by separating them into their spatial and temporal components. The spatial component is the visible 'footprint' the cell occupies in the FOV and the temporal component is the $[Ca^{2+}]_i$ transient of the cell/neuron. If cells overlap, so would the $[Ca^{2+}]_i$ transients if they were extracted simply based on the pixels a cell/neuron occupies in the recorded imaging data. The CNMF-E algorithm separates the two signals, even if they overlap on an individual pixel level.

The second approach utilized an open-source MATLAB pipeline called 'Min1pipe' (Lu, 2018)(Lu et al., 2018), which is an all-in-one solution for the analysis of microendoscopy data. It was used for some of the datasets that showed extensive motion artefacts and therefore required a non-rigid motion correction. Because the results of both pipelines for the same dataset are very similar, results of both pipelines were used for the following analysis. Animal 1 and 2 were processed with the customized pipeline while animal 3 was processed with 'Min1pipe'.

After the extraction of the $[Ca^{2+}]_i$ transients, the z-score ($Z = \frac{x-\mu}{\sigma}$; z is the z-score, x the data point, μ the mean of the population and σ the standard deviation of the population) of each transient was calculated and then used for further analysis.

2.8.4. Analysis of the operant conditioning data

As described above (2.6.1), the operant conditioning setup provides 5 data streams (cue-light, left and right lever press, left and right head entry, also see Figure 1 and Figure 10). The three of them which are on the rewarded side of the setup (always right) were used for further analysis. The

cue-light was used as an indicator on whether the setup registered a rewarded lever press (rLP). A MATLAB class for both the lever presses (LP) as well as the head entries (HE) was written. Within their respective classes, each LP or HE was treated as a tag with a time point and contained various properties that add information about the behaviour of the animal during the time before and after the tag. These properties include discrete (true, false), as well as continuous values (e.g. time in ms).

For the lever presses the discrete properties are whether the lever press was rewarded (rLP) and whether the reward was collected. An rLP was defined as a lever press that triggered the cue-light and therefore the presentation of a reward in the head entry port of the same side. If a head entry was detected during the cue-light sequence, the lever press was defined as one where the reward was collected. Because the setup did not contain a lick sensor for the dipper cup in which the reward is stored, it is not certain that the reward was actually collected. The continuous properties are the 'initial index', 'end index' (both in frame number), the time the animal took to collect the reward (ms, only for rLPs), the time until the next rLP (ms) and the time since the last rLP (ms).

Property name	Data type	Description
rewarded	boolean	Set to 'true' if the LP triggered the cue-light and the dipper cup was moved. Otherwise it is set to false.
reward collected	boolean	Set to 'true' if a HE was detected during the cue-light sequence of the rLP. Set to 'false' if no HE was detected.
initial index	integer (32bit)	The first time point of the LP signal.
end index	integer (32bit)	The last time point of the LP signal.
time until collection	integer (32bit)	The time that passed between the 'end index' of the rLP and the beginning of the HE signal. Only if the HE signal started within the cue-light sequence.
time until next rLP	integer (32bit)	The time that passed until the next rLP.
time since previous rLP	integer (32bit)	The time that passed since the last rLP.

Table 6: LP tag properties

For the head entries the discrete properties are whether the head entry happened during a cue-light and whether there should have been a reward in the dipper cup. The latter was determined by tracking the number of times the dipper cup was filled with a reward and the consecutive head entries. If a reward was triggered, the first consecutive head entry was marked as one that contained a reward in the port. All head entries afterwards, until the next rLP, were then marked as ones that did not contain a reward in the head entry port.

Property name	Data type	Description
during cue-light	boolean	Set to 'true' if the HE took place during a cue-light sequence.
reward in HE	boolean	Set to 'true' if the dipper cup should have contained a reward (see text above for detailed description).
duration	Integer (32 bit)	The duration of the HE signal.

Table 7: HE tag properties.

Using the respective classes for both 'lever tags' as well as 'head entry tags', distinct events can be defined and the time points of all their occurrences can be extracted from the data. For example, all lever presses that were rewarded and for which a head entry took place during the display of the cue light.

2.8.5. Merging of $[Ca^{2+}]_i$ imaging and behavioural data

Because the behaviour data was not collected in a continuous way, but only when the operant conditioning setup triggered a signal, the data does not have a linear ascending series of timestamps and is also asynchronous to the calcium imaging data. In order to be able to link the two datasets, the frame trigger of the UCLA Miniscope system, as well as a timestamp was collected along with the data of the operant conditioning setup (2.6.1). Hence, every single frame of the calcium recording can be linked with its timestamp to every event that was recorded from the operant conditioning setup. To achieve this, a separate MATLAB class (Sequence Adjustment, 2.8.2) was written which extracts and synchronizes the datasets and in case of the behaviour data also interpolates the data points to create data that linearly progress in time.

After the synchronization time points can be chosen based on information from the lever press or head entry tags (2.8.4). The MATLAB class then uses a 1D-interpolation based on the timestamps to resample the behaviour data into 1 ms bins. The resampled data traces are then extracted within a range of time that can be freely chosen. For the further analysis, 1 s before until 6 s after the rLPs, 2 s before until 5 s after the beginning of the head entry (with reward) and 5 s before until 2 s after the end of the head entry (with reward) were chosen. The different time windows were chosen to account for the sequential nature of the behavioural task and they ensure that an approximately similar time window around the cue-light sequence is used.

2.8.6. Matching of cells between sessions

To match the cells between sessions an open-source pipeline called 'CellReg' (Sheintuch, 2017; Sheintuch et al., 2017) was wrapped in a class that could be inserted into the analysis pipeline. It uses a probabilistic model of the spatial footprint similarities of neighbouring cell-pairs from different sessions after initial alignment to a reference coordinate system using rigid-body transformation. The initial cell registration uses an optimized registration threshold, which is refined using a clustering algorithm. As input to the pipeline, spatial components (2.8.3) of all neurons were used for each session.

2.8.7. Bootstrapping of rLP $[Ca^{2+}]_i$ transients

Bootstrapping was used to assess the tuning of a cell towards the rLP sequences during operant conditioning. For this, the previously described (2.8.5 and 2.8.1) MATLAB class was used to extract all $[Ca^{2+}]_i$ transients that were synchronized to the rLPs. Another MATLAB class was then created to handle the necessary bootstrapping calculations.

First, the average $[Ca^{2+}]_i$ transient for each neuron was calculated (Figure 12). The variance of this average was calculated to estimate whether the $[Ca^{2+}]_i$ is elevated during distinct time points of the chosen sequence. The approach was only applied to the rLP sequences because the variance is not dependent on a single large event, but will also be elevated if the transient is elongated in time and has a larger amplitude. Therefore, also neurons that are active during interactions with the head entry port will be detected, even though the averaged transients were aligned on the lever press as a fixed point, which is independent of the time point of the head entry and exit and can lead to a 'smearing' effect along temporal axis.

The variance was then used as a neuron's reference value for the chosen behavioural sequence (sequences around the rewarded lever presses which triggered a cue-light were evaluated). To

evaluate whether a certain variance value is meaningful or not, the variance was calculated again for a sequence of randomly chosen time points instead of those of the actual rLPs. This was repeated 50000 times, resulting in a distribution of possible variance values. The variance value for the 'true' rLPs can then be assigned a percentile, based on its position within the distribution of values for randomly chosen time points.

Because every behaviour session yields different numbers of rLPs, the influence of the number of available rLPs was calculated by repeating the calculation of the percentile for lower numbers of rLPs. For this, different numbers of rLPs were used to calculate the percentile. The lowest number was 4 and then the numbers were increased in intervals of 2 until the maximum number of available rLPs was reached. Every iteration was repeated 10 times. All calculations were done on the local HPC (2.8.9)

2.8.8. Assignment of behavioural phases

Because the animals could freely choose when to press the lever and when to collect the reward, these events in the behavioural sequence are independent. To better estimate the tuning of a particular neuron to one of these events, the alignment of the averaged $[Ca^{2+}]_i$ transient of all neurons that reached at least the 100th percentile in the bootstrapping approach (2.8.7) was calculated for the beginning of the head entry (-2 s until +5 s, respective to the HE beginning) and the end of the head entry (-5 s until +2 s, respective to the HE end). For this first analysis, all alignments were calculated for the $[Ca^{2+}]_i$ transients of all sessions combined. The neurons were then each classified into one of three phases based on the two additional alignments. First, all alignments to the end of the head entry were used to find all neurons that had the maximum averaged $[Ca^{2+}]_i$ transient after the end of the head entry. These neurons were then generalized to be 'post head entry' neurons and removed from the list of neurons that had not been assigned a phase yet. In the next step, those neurons that were not classified as post-HE neurons and had their maximum averaged $[Ca^{2+}]_i$ transient after the beginning of the head entry were assigned as 'head entry' neurons. The remaining neurons were then assigned as 'pre head entry' neurons.

To assess whether neurons 'keep' their assigned phases, the same calculation was repeated for each of the self-administration sessions and the reinstatement session.

2.8.9. High performance computing cluster

Because of the high computational demand of some of the used analysis methods the local high performance computing cluster of the university was used (bwUniCluster). Moba X-term was

used to access the cluster. All data was used on a remote drive with direct access to the HPC. To run calculations with the Min1pipe Matlab package the 'fat'-nodes (40 cores, 1536 GB memory) were used. For all other applications the 'standard'-nodes (16 cores, 128 GB memory) were used.

2.8.10. Behavioural statistics

All behavioural statistics between multiple self-administration, extinction and reinstatement sessions were tested with a two-way ANOVA in Matlab (2018b), using a multiple comparison post-hoc test to identify significant differences (criterion: $p < 0.05$, tukey-kramer).

3. Results

The results are structured in three parts. First, the overall outcome of establishing the surgeries as well as the analysis are presented (3). The second part details the results of the analysis of the individual datasets. This includes the behavioural data, the $[Ca^{2+}]_i$ sequences and the position of the neurons in their respective FOVs throughout consecutive sessions (3.2 - 3.6). Lastly the results of the combined analysis, using all previous results, are presented (3.7 - 3.8).

3.1. GRIN lens implantation and micro endoscopic imaging

All experiments followed the same sequence of procedures (Figure 7). First, rAAVs for the expression of GcaMP6f under the human-synapsin promotor were injected and then GRIN-lenses were implanted within the same surgery. This was followed by 1-2 weeks of recovery, after which the initial training for the self-administration took place (2.6). Then the baseplates for the fixation of the miniaturized epi-fluorescence microscopes were attached in a second procedure. If, in the field of view, strong signals of clearly identifiable cells were visible, the animals were habituated to the attachment of the microscope in the following days. After these 'preparatory' steps, the animals performed three additional self-administration sessions, followed by extinction and a reinstatement session (2.6). During these sessions the behavioural data was collected simultaneously with the video recording of the fluorescence microscope. The animals generally reached the extinction criterion within 4-5 sessions. The extinction criterion was, that the number of lever presses does not exceed 25% of the lever presses of the previous SA session with the most lever presses. Subsequently, 1-2 reinstatement sessions were recorded and the session with the higher number of lever presses was chosen for the analysis.

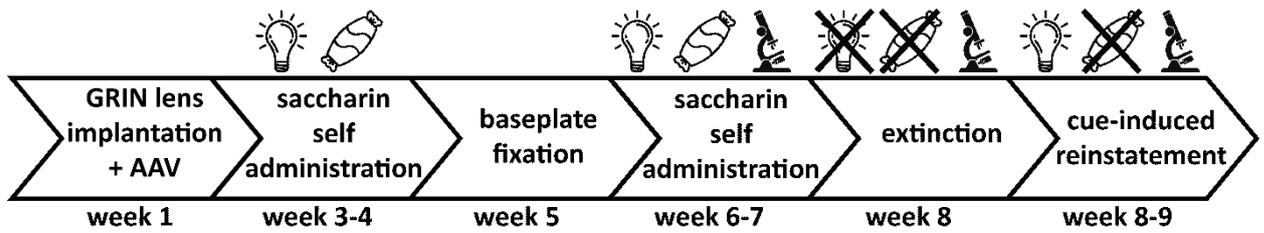


Figure 7: Schematic representation of the experimental time course. Time course from left to right. Pictograms indicate the cue-light (light-bulb), the saccharin reward (candy) and the miniaturized epifluorescence microscope (microscope). After a recovery phase from the surgery animals undergo a training phase of saccharin self-administration without attachment of the miniscope (2.6.2). After at least four weeks of GCaMP6f expression, baseplates for the miniaturized epifluorescence microscopes were mounted (2.5). After a brief habituation to the microscopes (not shown), self-administration (SA) sessions are combined with the calcium imaging (2.6.1). During the extinction the cue-light as well as the reward are missing (2.6.3). For the reinstatement (RE) only the saccharin reward is missing.

3.1.1. Micro-endoscopic imaging

The microendoscopic imaging data is saved in 8bit grayscale format and because of the optical limitations of GRIN lenses generally allows little inference about histological features, besides occasional blood vessels and neurons which are active (Figure 8a). To quickly assess the quality of a recording, processing of the image stacks to show the change in pixel intensity over time, also called dF/F , can be used (Figure 8b, single image of dF/F stack). Alternatively calculations of the standard deviation of each pixel over time condenses the stack to a single image and allows the quick evaluation of the presence of neurons that can be found within the field of view (Figure 8c). After motion correction of the image stacks the extraction of neuronal $[Ca^{2+}]_i$ traces can be started using demixing techniques (2.8.3). These generally start with a correlation image (Figure 8), to identify seed pixels for the nonnegative matrix factorisation. This procedure identifies individual sources of signal in the image stacks, even if the signal is overlapping on individual pixels. The signal is then represented in form of two components, the first showing the spatial extent within the field of view (Figure 8e) and the second its temporal progression (Figure 8f), showing intracellular $[Ca^{2+}]_i$ fluctuations over time.

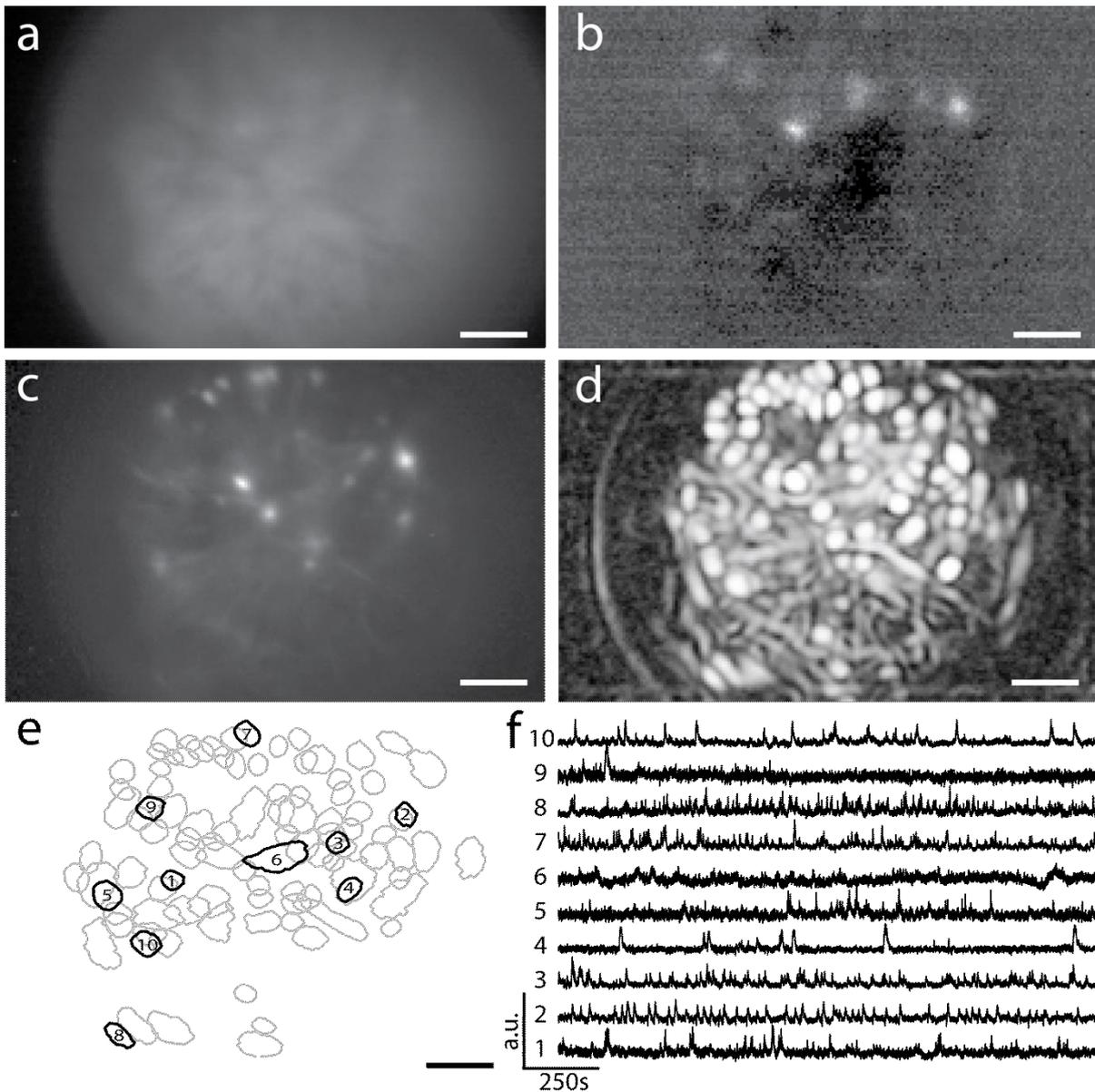


Figure 8: data processing steps from raw videos to $[Ca^{2+}]_i$ traces for individual neurons. **a** example of a single frame as raw data **b** same frame depicted as dF/F image (dF from 1000 frame average) **c** standard deviation projection of 1000 frames **d** correlation image used to pick seed pixels for the non-negative matrix factorisation **e** examples of spatial components extracted using non-negative matrix factorisation, 10 randomly chosen neurons are outlined in black and the temporal components are shown in **f**. All data is derived from Animal 1 SA session 1. Scale bars show 100 μm .

Generally, a single $[Ca^{2+}]_i$ recording was synchronized with the respective behavioural data of the same session. Due to issues with the hard-, and software, multiple separate $[Ca^{2+}]_i$ recordings were acquired for some sessions. These separated calcium recordings were re-connected using the SessionStitcher class (2.8.1). In addition, the calcium imaging recordings also occasionally showed dropped frames. These did not exceed more than a 20-30 seconds per session and individual phases of frame dropping remained sufficiently short (<1 second) to not alter the data.

To compensate the dropped frames a 1d-interpolation of the calcium transients with their actual time stamps as well as the aspired frame rates was computed prior to the synchronisation with the behavioural data (2.8.5).

To assess placement of the GRIN lenses and the expression of GCaMP6f, thin sections were cut after intracardial perfusion and imaged using an epifluorescence microscope. Figure 9 shows the path of the GRIN implant after its removal. To better preserve the outline of the GRIN lens, the tissue was left in the fixation solution for at least 7 days. Figure 9a shows a representative overview of the region in which the lens was implanted. The lens is clearly positioned in the IL, with the anterior olfactory nucleus visible in the bottom and parts of motor cortex to the left. Figure 9b-d shows close ups of the bottom of the GRIN implant.

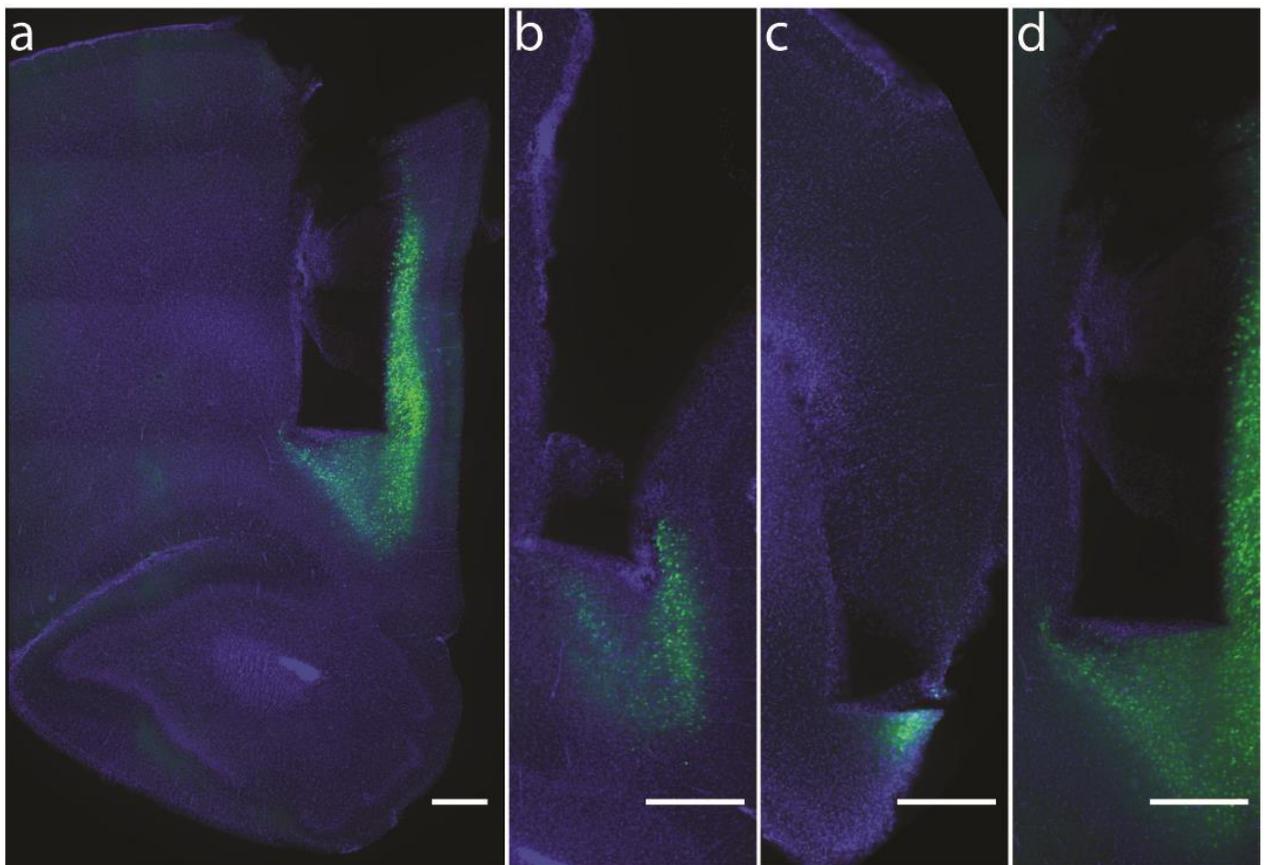


Figure 9: GCaMP6f expression and imprint of the GRIN lens. a to d, blue channel is DAPI, green is GCaMP6f autofluorescence. a shows an example of the GRIN lens placement. b-d show close ups of the recording site. Scale bar represents 500 μm .

3.1.2. Analysis pipeline

To cope with the complexity of the datasets, MATLAB classes were written to separate individual parts of the analysis and then combined in a class combining all experimental data for an

individual animal (2.8). Because the data streams ($[Ca^{2+}]_i$ traces and behavioural data) are asynchronous, we used a 1-D interpolation of the time points with a trigger signal that was present in both sets of data. After synchronisation of the data streams, specific events in the behaviour data can be identified and the corresponding calcium transients can be identified. An example of the synchronized data streams is shown in Figure 10.

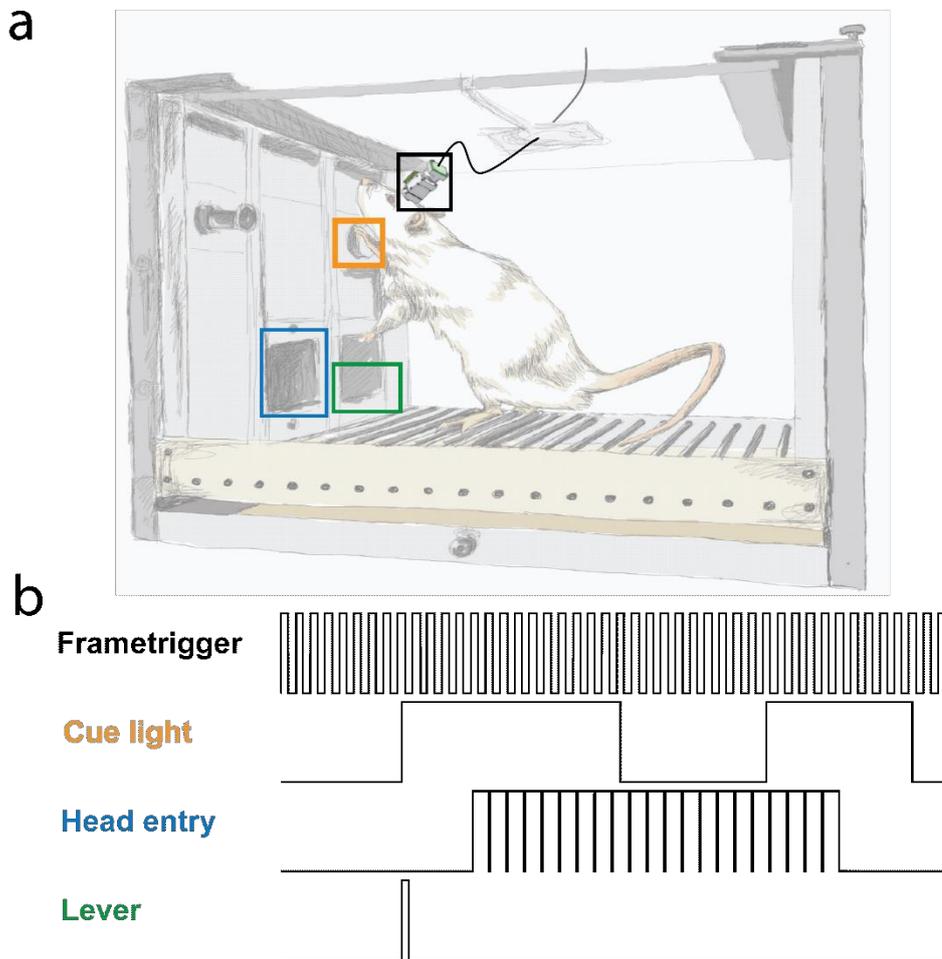


Figure 10: Schematic of the behavioural setup and the collected data traces. A rat interacting with one of the sides of the operant conditioning setup. The cue-light, lever and head-entry port are highlighted with coloured boxes, example data traces are shown in b.

Two distinct types of events were defined that could be used for this extraction. The first being the instances when the animals pressed a lever (lever presses; LPs) and the second being the instances in which the animal inserted its head into the head entry port (head entries; HEs) to pick up the reward. It is important to remember that only the lever on one of the sides is rewarded, and that the side does not change during the complete duration of the recordings. The reward is presented in the HE port on the same side of the rewarded lever. Each occurrence (LP and HE) was defined as a 'tag' and carried a range of attributes that describe its properties. For

lever presses these include, among others, the starting and ending time point, whether the lever press resulted in the presentation of a reward and duration until the next head entry (2.8.4). For head entries, the properties include e.g. the duration or whether there should have been a reward present in the head entry port (see methods for detailed description). Using these 'tags', we separated the lever presses into those that were rewarded and those that were not rewarded and subsequently only focused on lever presses that resulted in the placement of a reward in the head entry port and consequently the presentation of the cue light. We then used the time points associated with each 'tag' to extract the calcium transients for the time period of 1 second before the occurrence of the lever press until 6 seconds after the lever press, or in cases of alignment to the beginning or the end of the HE, 2 seconds before until 5 seconds after (beginning HE) and 5 seconds before until 2 seconds after (end HE). These precisely defined time windows allowed the registration of neuronal ensemble activity with distinct episodes of the reward behaviour.

3.2. Operant self-administration behavioural results

To ensure that the animals properly learned how to interact with the operant conditioning chamber, they were first trained on an operant self-administration paradigm without the presence of the miniature microscopes (pre-training phase, 2.6.2). After proficiency of the paradigm was established, the baseplates were attached to the implanted GRIN lenses (2.5). The pre-trained animals generally needed a few (1-10, data not shown) self-administration sessions to start pressing the lever again after the baseplate was fixed. This habituation to the presence of the miniaturized microscope was needed because the animals were often focussed on exploring the microscope cable (data not shown).

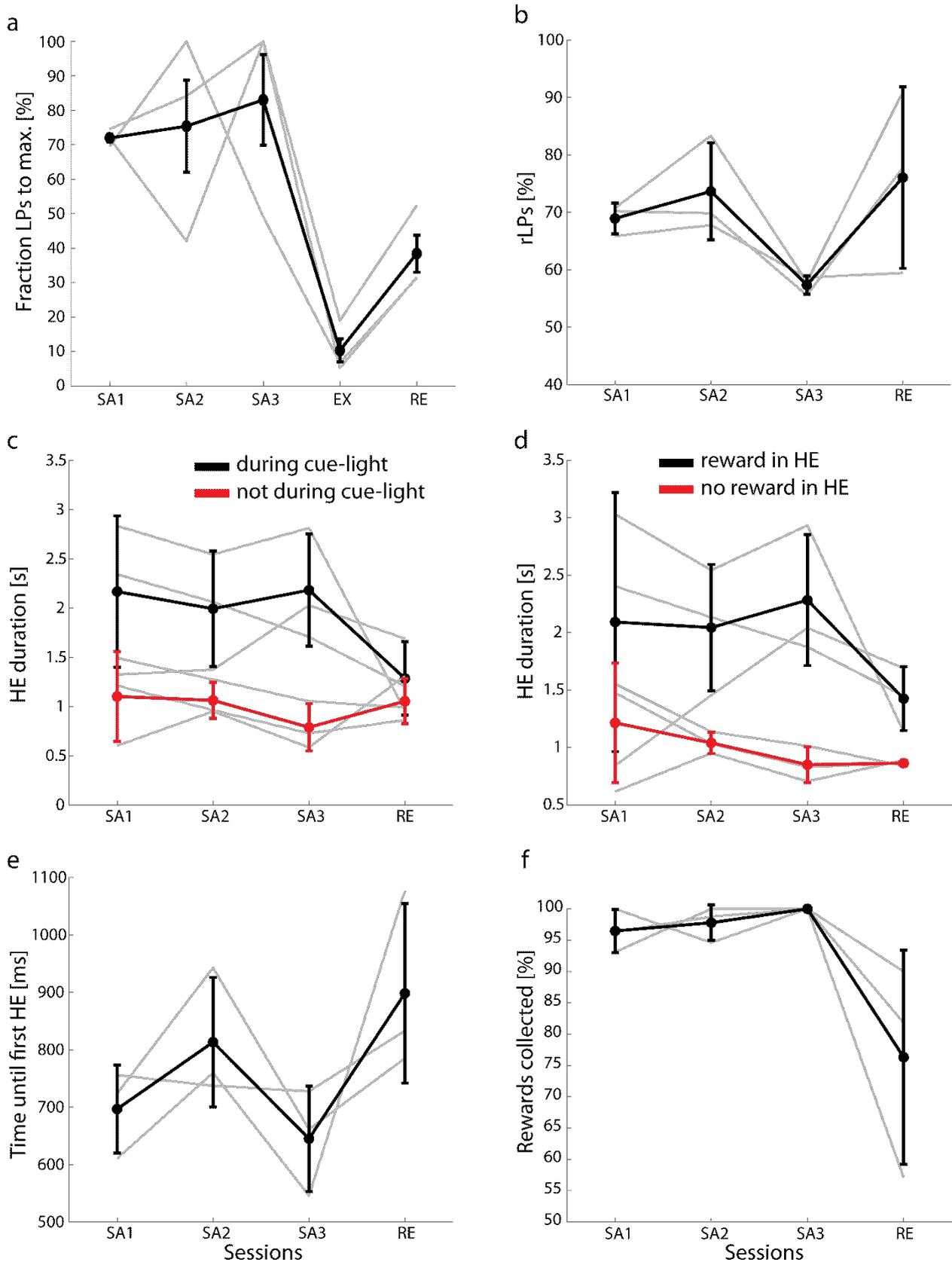


Figure 11: Behaviour Data. **a** Fraction of LPs in %, of the highest number of LPs of the three SA sessions for each animal. **b** Percentage of rewarded LPs from total right LPs per session. **c** The duration of the HE during the cue-light sequence (black) and outside of it (red). Individual animals

are shown in grey. **d** The duration of the HE for HEs which directly did (black) or did not (red) follow a rLP, hence for which the HE-port contained a reward or would contain a reward under SA session conditions (e.g. during RE sessions). Individual animals are shown in grey. **e** Time until the first HE after a rLP. **f** Percentage of the rewards collected after the rLPs. Individual animals are shown in grey, the mean value is shown in black (**a-f**) and red (**c-d**).

Figure 11a shows the mean number of rewarded lever presses (rLPs) normalized to the highest number of rLPs of all self-administration sessions for each animal were $72.9 \pm 25.6\%$, $71.3 \pm 28.9\%$ and $86.2 \pm 12.8\%$ (mean \pm std). The extinction only shows the percentage for the last extinction sessions prior to the reinstatement, because this is the session in which the criterion (<25% of maximum LPs) was met. During this last session the rLPs were thus below the criterion for all animals (6.8%, 5.3% & 19.0%). During the reinstatement the rLP rates reliably increased again to values above the extinction criterion (31.3%, 31.5% and 52.4%). The difference between the extinction session and the RE session was not significant (two-way ANOVA, criterion: $p < 0.05$, post-hoc test: tukey-kramer). The only significant differences could be found between the SA sessions and the extinction session (two-way ANOVA, criterion: $p < 0.05$). Even though the differences between the extinction and the reinstatement are not significant, the trend towards an increase in LPs during the reinstatement session, in which the cue-light is presented after the LP, is visible.

The animals often pressed the lever during the cue-light period, which would not result in a reward. This is shown in Figure 11 b as the percentage of all lever presses that resulted in the presentation of the cue-light and a rewarded lever press (rLP). Between the individual animals the values ranged from 55.5% to 83.3%, ($68.9\% \pm 2.7\%$, $73.6\% \pm 8.4\%$, $57.3\% \pm 1.6\%$, [mean \pm std]). The reinstatement shows the largest differences between the animals with value of $76.1\% \pm 15.8\%$ (mean \pm std). The extinction is not shown because of the absence of cue-light presentation.

The animals did not enter the HE port during all presentations of the cue-light and hence did not always collect the reward. In these cases the reward remained in the HE port and could be collected the next time, the animal entered the HE port, regardless of whether the rLP was triggered before. Also, the animals would sometimes enter the HE port multiple times during a cue-light presentation. Hence to see if the animals were aware of the experimental contingencies and acted by a reward-driven motivation and not through a habitual automation, both the dependency of the HE duration by presence of a cue-light as well as the presence of a reward is detailed below.

The individual times spent by the animals in the HE port varied strongly between animals, as is shown in Figure 11 c. HEs that occurred during the cue-light sequences range from 1.33 to 2.83 s, however, the average HE durations between the self-administration sessions are fairly stable (SA 1: 2.16 ± 0.77 s, SA 2: 1.99 ± 0.59 s, SA 3: 2.18 ± 0.57 s, [mean \pm std]). The duration decreased in the RE sessions (1.29 ± 0.37 s, [mean \pm std]). Interestingly, exploratory HE made in the absence of reward predicting signals, i.e. outside of the LP/cue-light sequences, were consistently shorter compared to those during the cue-light sequences (SA 1: 1.10 ± 0.46 s, SA 2: 1.06 ± 0.18 s, SA 3: 0.79 ± 0.24 s, RE: 1.05 ± 0.23 s, [mean \pm std]) (see Figure 12 for a detailed description of the sequence). The only significant difference was found for the SA 3 session (two-way ANOVA, criterion: $p < 0.05$, post-hoc test: tukey-kramer). In the RE session the animals spent about the same time in the HE port during the cue-light sequences, as they did when there was no cue-light present.

A HE with a reward was defined as one that followed an rLP. It was then assumed that if a reward is present it was consumed by the animal and all following HEs are regarded as ones with no reward present, until the next rLP. Because the cue-light does not reliably indicate whether a HE contained a reward on entry, the dependence of the HE duration to the presence of a reward was investigated next. As expected, the duration for HEs most likely containing a reward differed from non-reward containing HEs (Figure 11 d). In all SA sessions rewarded HEs range between 0.84 seconds to 3.03 seconds (SA 1: 2.09 ± 1.12 s, SA 2: 2.04 ± 0.55 s, SA 3: 2.28 ± 0.57 s, [mean \pm std]). The durations of HEs for which no reward was present range between 0.61 s to 1.55 s during the SA sessions (SA 1: 1.21 ± 0.52 s, SA 2: 1.04 ± 0.09 s, SA 3: 0.85 ± 0.16 s, [mean \pm std]). During the reinstatement session, there was, of course, no reward present. However, if a LP resulted in triggering the cue-light sequence, the next HE was counted as one that would have been expected to contain a reward. The HE durations for those HEs that carried the expectation of a reward ranged between 1.14 s to 1.55 s (1.43 ± 0.28 s, [mean \pm std]). Those that did not contain a reward were again lower and ranged between 0.84 s to 0.89 s (0.86 ± 0.02 s, [mean \pm std]). The animals spent less time in the HE port if it would have contained a reward under SA session conditions. Interestingly there was still a difference between HE durations for those HEs that would have contained a reward in SA sessions (reinstatement HEs after LPs which triggered a cue-light). Therefore, the presence of the reward and the time it takes to consume it does not solely account for the time differences of the HE durations between SA and RE sessions. This difference however was not significant (two-way ANOVA, criterion: $p < 0.05$, post-hoc test: tukey-kramer).

The differences in HE durations during cue-light presentation, as well as times when there was a reward present in the HE port (both conditions often coincide, but do not have to), show that the animals are aware of the correlation of the cue-light and the presence of a reward, and only take

a short amount of time to evaluate whether a reward is present or not, as can be seen in the decrease of the HE duration in the RE sessions.

Next, the time until the first HE after an rLP was investigated, to check how quickly the animals collected the reward and to check for inter animal differences (Figure 11 e). For the SA sessions these times range between 546 ms to 942 ms and show high fluctuations between sessions (SA 1: 696 ± 76 ms, SA 2: 813 ± 112 ms, SA 3: 645 ± 91 ms, [mean \pm std]). In comparison to the times during the SA sessions, the animals required slightly more time to reach the HE port during the RE sessions (898 ± 156 ms, [mean \pm std]), probably because of remaining uncertainty about the presence of a reward in the HE port. This difference was not significant (two-way ANOVA, criterion: $p < 0.05$, post-hoc test: tukey-kramer).

The percentage of rewards that were collected within the cue-light sequence was always above 90% for all SA sessions and animals, ranging between 93.1% to 100% (SA 1: 96.5 ± 3.5 %, SA 2: 97.8 ± 2.8 %, SA 3: 100 ± 0 %, [mean \pm std]; Figure 11 f). It is important to note, that there is no certainty on whether the reward was actually collected because of the lack of a lick sensor. The criterion to define a collected reward, was the registration of a HE within the cue-light sequence following a rLP. Interestingly, during the RE sessions, this drops to a mean value of $76.3\% \pm 17.1$ (mean \pm std) with values for individual animals ranging from 57.1% to 90%. Hence, within a single session, the animals realized that pressing the lever and the presence of the cue-light was not indicative of a reward in the HE port. There was no significant difference between any of the SA and the RE session however (two-way ANOVA, criterion: $p < 0.05$, post-hoc test: tukey-kramer).

Because of the limited number of inactive lever presses, these were not included in the analysis.

3.2.1. Overall efficacy of GRIN implantations in rats

Due to issues with the surgical procedure only a small number of animals were left to perform $[Ca^{2+}]_i$ imaging in combination with operant self-administration (see 2.3 for details of the procedure, roughly 60% of the animals remained). During the attachment of the baseplates some animals had to be discarded because the implants appeared to be loose or they did not show sufficient expression for imaging (roughly 20% of the animals remained). The issue of implant detachment continued throughout the following operant conditioning sessions, resulting in the loss of many animals due to the self-inflicted removal of the implant during the recordings (roughly 10% - 15% of the animals remained). The loss of implants was presumably due to a mixture of both infections as well as the excessive damaging of the periosteum that results in the degeneration of the cranium and consequently in an instability of the implants (Gardiner & Toth,

1999). This was eventually solved by taking special care to reduce the damage to the periosteum of the cranium as much as possible and by performing surgeries in strictly sterile environments. An additional problem was the incorrect placement of the AAV injection in relation to the GRIN implant, leading to virus expression outside of the FOV range of the lens. For the remaining animals, recordings could be started, however only 3 animals yielded a sufficient number of recording sessions to complete the analysis of 3 SA and 1 RE session. This was caused by either loss of the implants during recordings, frequent destruction of the miniscopes, ending the sessions prematurely, insufficient signal to reliably extract $[Ca^{2+}]_i$ signals and in case of a single animal, absence of neural tuning to the reward seeking, caused by mistargeting the lens.

3.3. Combination of operant self-administration data and calcium imaging

In order to better understand the relation between the self-administration behaviour and the $[Ca^{2+}]_i$ transients and to extract meaningful information from it, both datasets were synchronized. This was done using the time-stamped calcium imaging frames and all time indices of the behavioural data (2.8.5). Figure 12a shows an example of the distribution of multiple rLPs during a 30 minute SA session and Figure 12b illustrates the typical sequence of behavioural signals that are collected following an rLP.

Pressing of the lever on the rewarded side by the animal initiates the presentation of a reward inside of the respective (same side) head entry port, as well as a sequence of the cue light flashing three times (Figure 12b). During this time the animal can press the lever again, however no additional reward will be presented in the head entry port and no extra sequence of the cue light will be triggered. The behavioural signals monitored, include the lever presses for both sides of the operant conditioning box, the triggered cue light as well as the head entry port IR sensor for both sides. Essentially two independent signals coming from the behaviour of the animal could be determined. Namely the lever press and the head entry, from which then the time between the lever press and the start of the head entry could be collected, as well as the duration of the head entry.

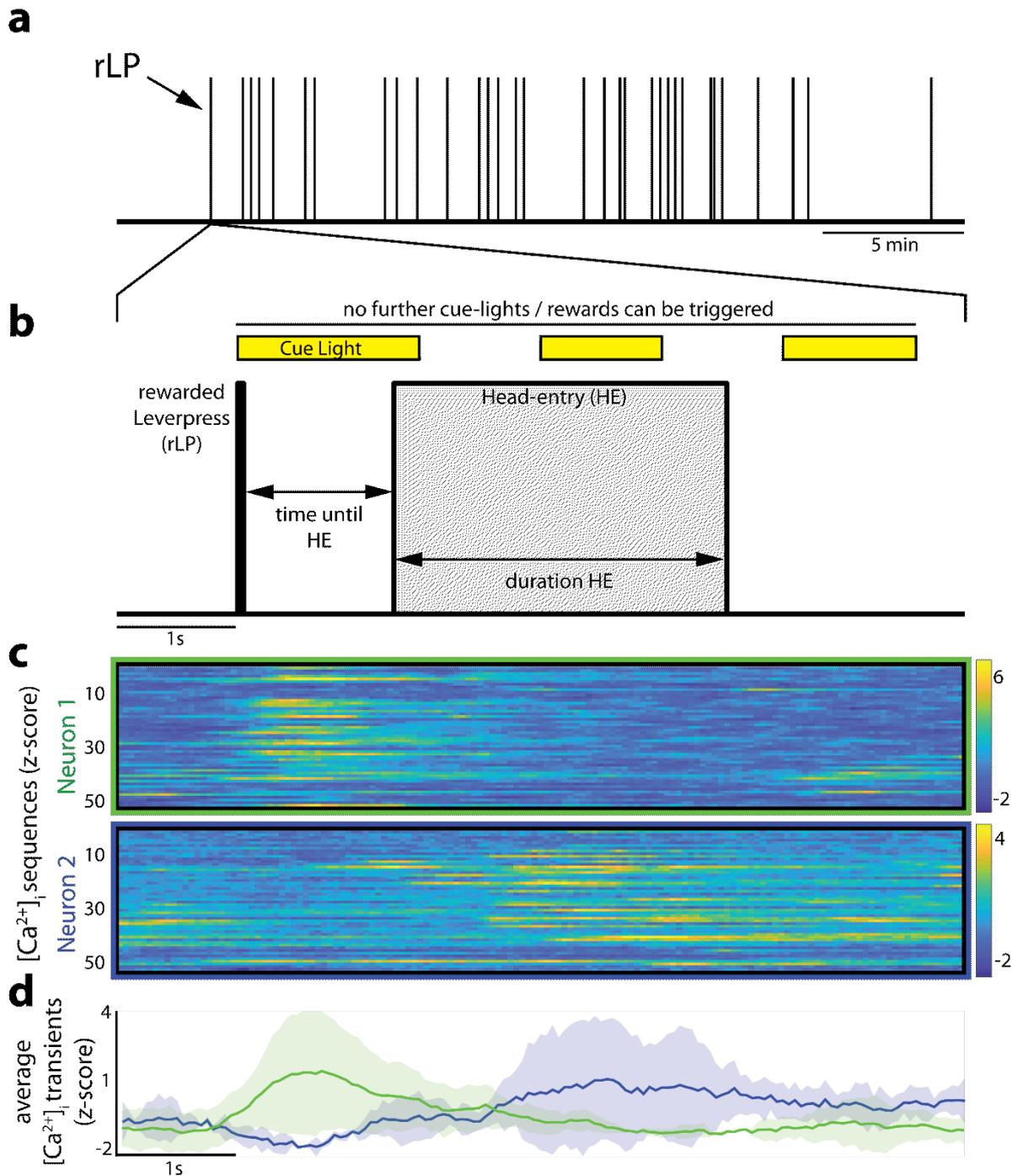


Figure 12: Self-administration sequences throughout the length of the session with simultaneously recorded $[Ca^{2+}]_i$ transients of individual neurons. a rLPs during typical 30 min recording session (SA and RE). **b** The temporal sequence (duration: 7 s) of behavioural events of the rLP highlighted in **a**. The rLP (black horizontal bar) which triggers the cue-light (yellow) and the presentation of a reward in the head entry port. During the cue-light sequence no additional LPs will not be rewarded. Arrows indicate the time between the rLP and the HE and the total duration of the HE (gray rectangle). Both durations vary between trials. **c** $[Ca^{2+}]_i$ transients recorded during the behavioural sequence shown in **b**. Each box (outlined in green and

blue) depicts the $[Ca^{2+}]_i$ sequences for two individual neurons. Each row shows a single color-coded $[Ca^{2+}]_i$ sequence (n=54), where yellow indicates higher $[Ca^{2+}]_i$ levels and blue areas lower ones. **d** The average $[Ca^{2+}]_i$ transients of the two neurons shown in **c**. The colours correspond to the colour coding used in **c**. **b-d** are synchronized.

As indicated in Figure 12b and quantified in Figure 11, the time until the animals put their heads into the head entry port for the first time, as well as the duration of time which is spent there depends on the animals individual behaviour. This variability of the behavioural signals becomes most apparent in the further processing of the data when analysing the synchronized $[Ca^{2+}]_i$ signal for the self-administration sequences relative to the rLP or the beginning or the end of HE. In case of Figure 12 the calcium signals relative to the time point of the rLP are shown. It is clearly visible that both neurons tend to be active during similar time points after the rLP with neuron 1, showing most activity (high $[Ca^{2+}]_i$ compared to the rest of the sequence) just after the rLP and with a more consistent offset in the activation relative to the LP, compared to neuron 2. Nonetheless, not all rLP sequences of the two individual neurons shown are accompanied by an increase in $[Ca^{2+}]_i$. This means that the neurons were not reliably active in every single reward seeking sequence, however if they were, the time point of their activation was somewhat correlated to one of the behavioural signals that was collected.

The averaged transients for both neurons are shown in Figure 12d, with colour coding corresponding to the outlines of the two boxes in Figure 12c. The dark line shows the average value, with the standard deviation shown as the transparent outline. Subsequently, the average activity of neuron 1 is higher in close temporal proximity to the rLP and is more temporally defined compared to neuron 2 (Figure 5d). This difference in the 'tuning' is likely a consequence of the behavioural cue (LP or beginning/end HE) which is used to align the $[Ca^{2+}]_i$ transients and does not indicate that neuron 1 is more relevant to the behaviour than neuron 2.

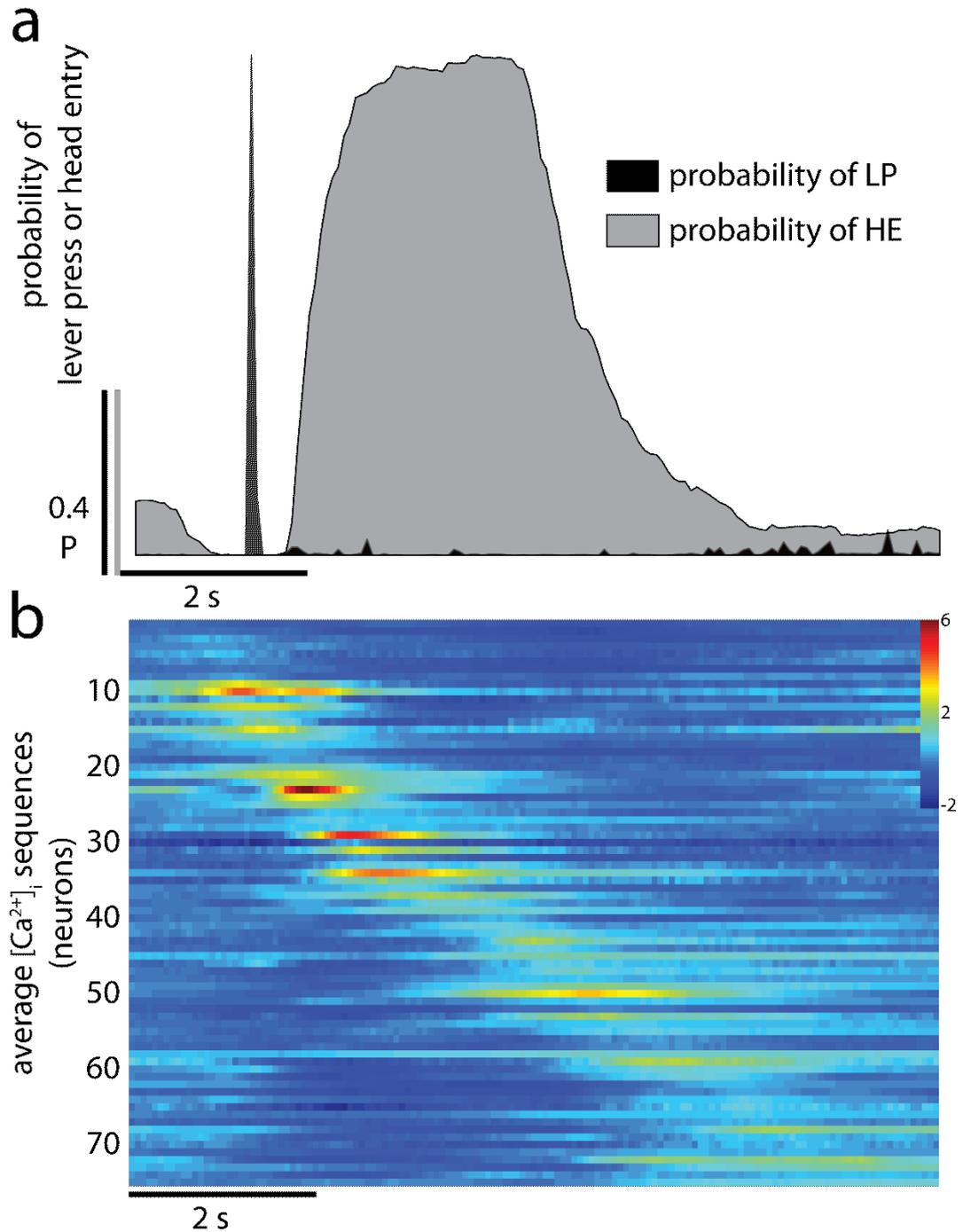


Figure 13: Average $[Ca^{2+}]_i$ sequences of individual neurons, synchronized to the probability of a lever press or head entry. **a** Probability of the occurrence of a LP (black) and a HE (gray) from the operant conditioning setup. The black area shows the LP probability and the gray area the HE probability. **b** Each row corresponds to the average $[Ca^{2+}]_i$ sequence (as shown for two examples in Figure 12d) of a single neuron, relative to respective rLPs of all three SA sessions. The sequence displays the time window of 1 s before the rLP until 6 s after the rLP. The cells are sorted by the time of the peak activity. Color coding (z-scores) indicates higher $[Ca^{2+}]_i$ levels.

The alignment of the $[Ca^{2+}]_i$ transients to the time points of the rLPs reveals two distinct features (Figure 13).

First, as the sorting of the neurons by the time of their peak averaged $[Ca^{2+}]_i$ transient shows (Figure 13), is the temporal progression of the behaviour almost fully covered by neurons that show preferences for distinct time points of the sequence. This means that most of the time, any one of the neurons will be active, and also there does not appear to be much clustering around a specific time point. It has to be noted, though, that this does not mean that all neurons are tuned to the rLP. Second, the peaks of the averaged $[Ca^{2+}]_i$ transients appear to fade with increasing temporal distance to the behavioural signal (the rLP). This could mean that the behavioural signal (in this case the rLP) which is used to average the $[Ca^{2+}]_i$ signal, may not correlate well to the activity of the neuron (see Figure 16 for alignments to different signals).

Figure 13a shows the probability of a lever press, colour coded in black, and the probability of the head entry being occupied in grey. The data was pooled from all rewarded lever presses of all animals. Because the data is selected relative to the rewarded lever press, the probability of the lever being pressed at 1 s is 100%. There was also a low probability for the lever to be pressed within a second after the initial lever press (double presses) and towards the end of the cue light sequence. The probability of the head entry port being occupied shows a sharp increase within 0.5 seconds to 1.5 seconds after the rLP, lasting for about 2 seconds before declining again and then gradually decreasing. This tail of the head entry occupation coincides with the 'false' lever presses, likely because the animal returns to the head entry port for a reward after pressing the lever again. If this happens within the time period during which the cue light is still active, the presentation of another reward will be blocked.

3.4. Identification of neuronal tuning using shuffling

As shown previously (Figure 12 and Figure 13) individual neurons show a preferred time point during the behavioural sequence at which they are more likely to be active. In order to quantify, how likely it is that a particular neuron is preferably active at a certain time point during the self-administration sequence, shuffling of the data was used to determine the likelihood of a coincidental increase in $[Ca^{2+}]_i$ in relation to the behavioural signals. This was done by randomly picking time points instead of using the actual time points of the rLPs. Using the distribution of the randomly picked time points it can be estimated whether the averaged $[Ca^{2+}]_i$ sequences during the rLP are statistically different from randomly picked time points.

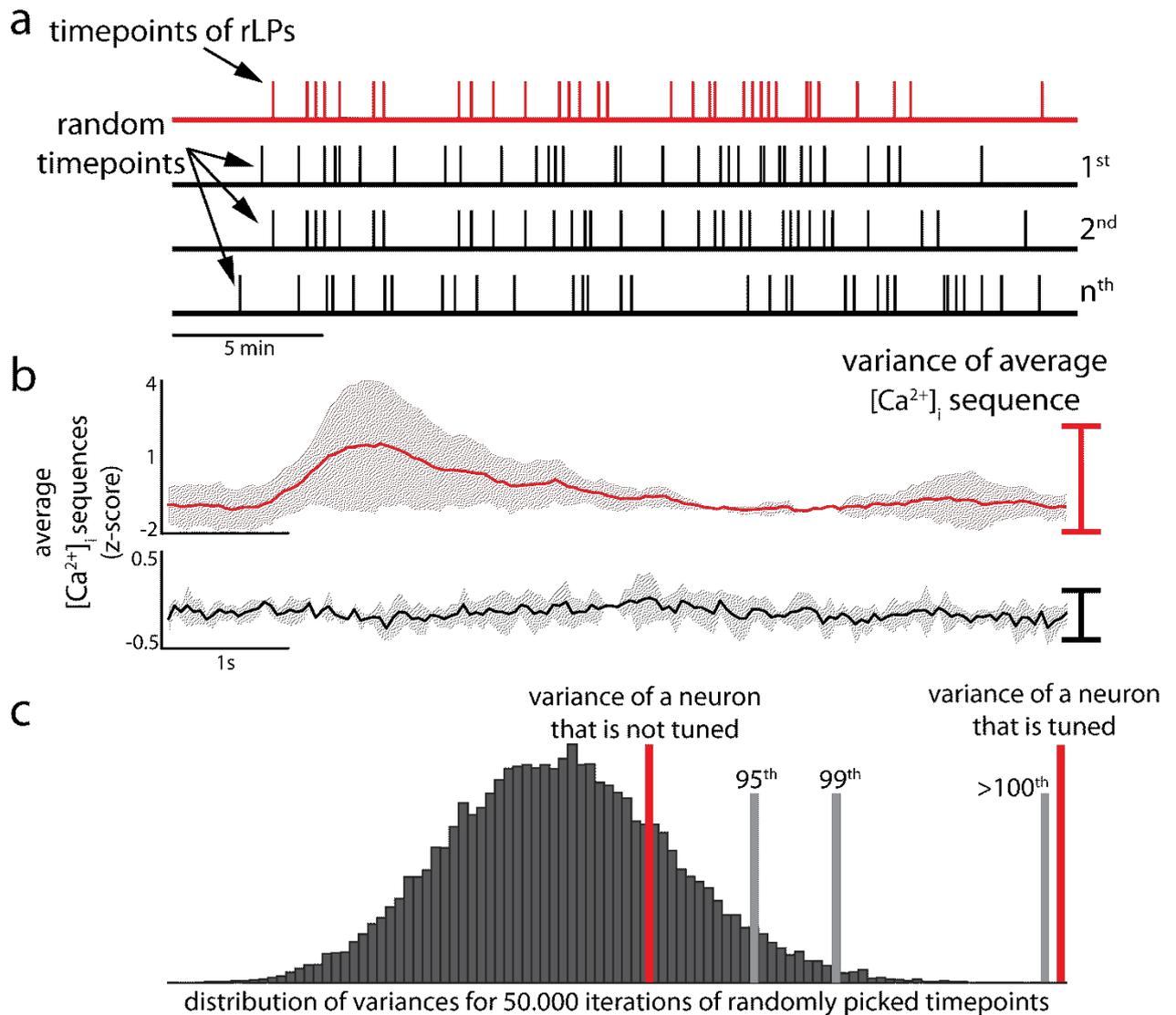


Figure 14: Bootstrapping the variance of averaged $[Ca^{2+}]_i$ sequences to identify tuned neurons. a four examples of LP occurrences during a 30 min session. Red shows an example of actual LPs, while random LP time points are black. b average $[Ca^{2+}]_i$ sequences of LP time points for two neurons, red showing an example of high and black of low variance. The average of the red is derived from LP time points, while the black average used randomly assigned time points. c histogram of 50000 variance scores of an individual neuron after randomized sampling. Gray bars indicate upper confidence intervals (95th, 99th and 100th).

As a metric to describe the likelihood of a repeated calcium response, the variance of the averaged traces was used as reference (see 2.8.7 and discussion for details). Then the variance of the averaged calcium trace for all rLP sequences was calculated and compared to the distribution of variances one would get from randomly chosen time points (50k iterations). The significance of the variance value for the ‘true’ rLP can then be determined by checking the distribution of the variances for randomly selected time points. The confidence interval within which the ‘true’ (rLP) variance lies can then be used to separate statistically more ‘tuned’ neurons. If the variance of a neuron falls into a percentile which lies above a defined threshold, it can be counted as specifically

responsive during the behavioural sequence (see 2.8.7 and Figure 14).

It is important to note, that this analysis and the resulting identification of 'true responder neurons' (neurons with consistently higher than expected variance) is strongly biased by two parameters which influence the number of neurons that will be identified. The first is the number of lever presses that are available to the bootstrapping (Figure 15a,c,e) and the second is the percentile (or confidence interval) that is used as the threshold for identification of responder cells. In order to determine these parameters, the bootstrap was performed for a range of different values for the number of rLP and the percentile. To estimate the influence of the number of rLPs available to the bootstrap, different (increasing) fractions of the available rLPs were picked at random and the analysis was run using only these smaller subsets (2.8.7). As a cut off to define neurons which are tuned to the self-administration sequence, the >100th percentile was used. This means that none of the 50k randomly chosen time points exceeded the variance value that was calculated for the actual time points of the rLPs (Figure 14).

It is apparent that none of the bootstrapped datasets for the SA sessions show a plateau towards the end of the behavioural sequence. If a plateau was formed, this would mean that the number of responding neurons does not continue to increase as the number LPs increases. The lack of the plateau formation means that the overall number of rLPs available in the individual sessions is likely too low to detect all potentially tuned cells. In theoretical sessions with infinite rLPs, the percentage of neurons that are above a certain percentile will eventually be independent of the number rLPs which were used for the analysis. Therefore, it is likely that the percentage of responders is underestimated as the real number of rLPs is too low. Also, while some of the data in Figure 15 a,c,e show similar curves for some of the SA sessions of the individual animals, the number of rLP available for analysis of the sessions that directly preceded the reinstatement (shown in black with dark gray outline) is higher than the RE session. Also, two of the animals had SA sessions for which the number of responder neurons in dependence of the rLPs for the RE lies above or on the same level. This means, that the number of identified responder neurons defined by the bootstrap, does not differ between SA and RE sessions. It is important to look at the dependence of the number of tuned neurons to the number of available lever presses instead of simply calculating the number of tuned neurons for the maximum number of LPs. Because of the two factors (#LPs and percentile threshold) mentioned above, looking at the relation of the dependence curves is more informative. If the curve for either the SA or the RE sessions would show a different progression, e.g. through its slope, this would be a more credible indication about differences in the overall tuning of the network towards the behaviour.

Overall neurons in the RE sessions tend to be less tuned to the reward seeking behaviour compared to the SA sequences, however, the effect was not pronounced in all sessions and it

cannot be excluded that this is observed due to the overall variance of the data. In addition, the curves of the number of tuned neurons over the available number of lever presses shows visible variance, despite randomly selecting the time points 50.000 times and then repeating the procedure 10 times for each set of randomly picked rLPs (2.8.7). In case of animal 2, this results in a noticeable decrease in the number of identified responder cells of the last SA session prior to the RE session, resulting in overlapping curves panel d (showing the dependence to the percentile using the maximum number of available lever presses). This can be misleading because the curves for rLP number dependence shows a clear difference in the slope as the numbers of rLPs increase, with a higher number of tuned neurons during the SA sessions.

After determining the influence of the number of lever presses, we analysed how the chosen threshold will influence the results (Figure 15). The number of responder neurons identified is dependent on the chosen percentile (Figure 15b,d,f). The red graph shows the reinstatement and the black graph the last self-administration session prior to the reinstatement. All animals show a decline in responder neurons as the threshold increases, which is expected because a higher threshold is more restrictive. The number of responder neurons identified during the reinstatement were lower, compared to the last self-administration session, for animal 1 and 2. However, for animal 3 they were nearly identical. This might be an artefact of the dip in the rLP curve shown in panel c (Figure 15) for the SA session shown in black

Strikingly, the dependence of identified responder cells on the number of rLPs as well as the threshold, is different between the sessions. Animal 1, which showed consistent numbers of responder neurons, for both the rLP dependence as well as the different percentile cut-offs, exhibits a lower number of responder neurons for all of the tested percentiles. All of the sessions (SA and RE) show a sharper decline between the 100th and >100th percentile. This sharper decline for the higher percentiles is less pronounced for the other animals.

This shows that the choice in percentile expectedly influences the resulting number of neurons that will be identified as responders, as does the number of rLPs available. However, it did not alter the trend that RE sessions show fewer responder neurons compared to the SA sessions, independent of the choices of the two parameters.

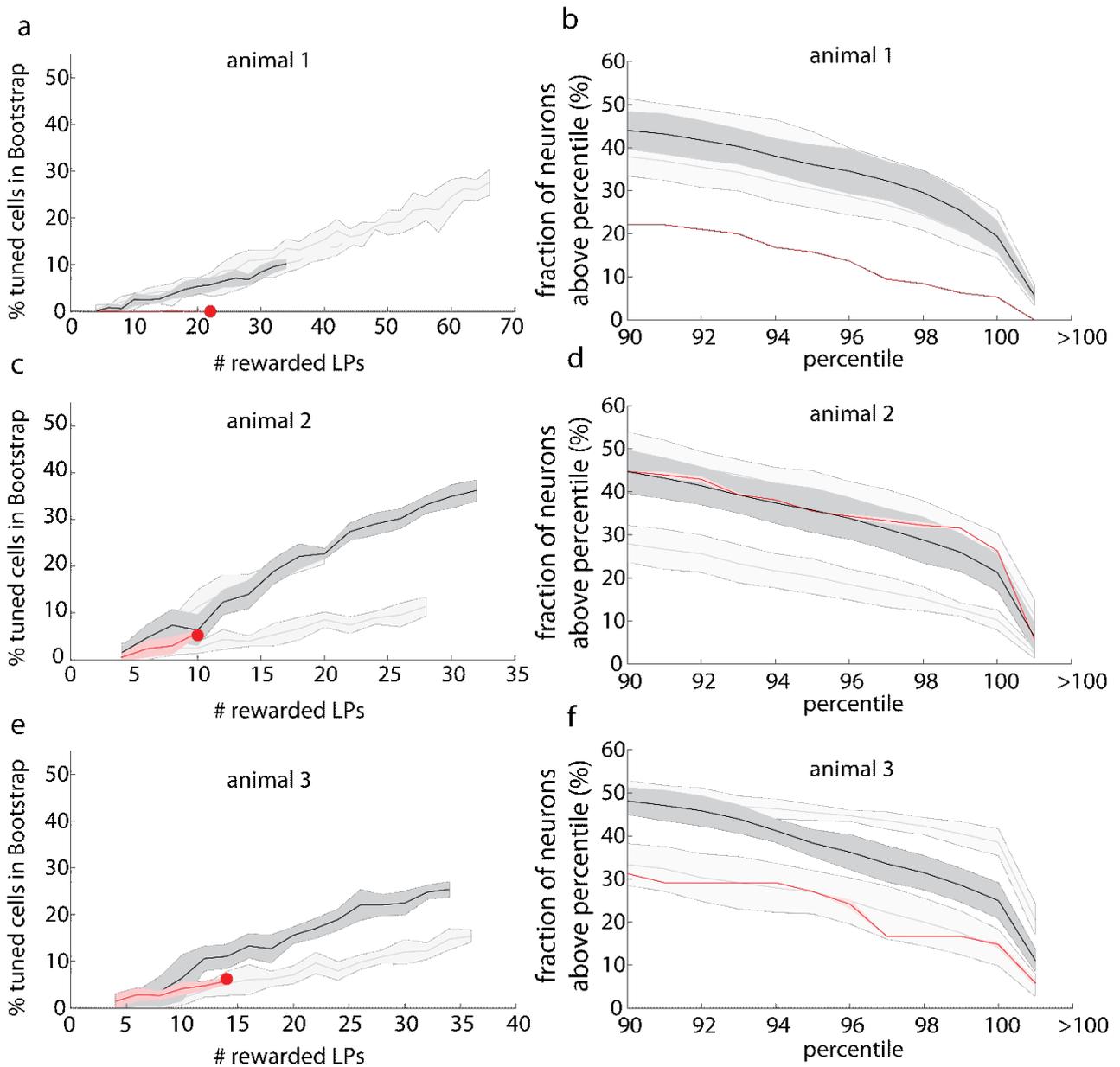


Figure 15: Bootstrapping results. Each row shows the data for an individual animal (**a-b**; **c-d**; **e-f**). The left column (**a,c,e**) shows the percentage of cells $>100^{\text{th}}$ percentile in dependence of the number of rLPs that were used for the calculation of the bootstrapped distribution (50.000 iterations). The calculation for each number of rLPs was repeated 10 times (2.8.7). The standard deviation is shown as lighter colored boundary to the darker lines. Black line show data for the last SA sessions prior to the RE session. Red lines show data for the RE session. The red dot indicates the maximum number of rLPs available for the RE session. The right column (**b,d,f**) shows the percentage of cells above the percentile on the x-axis. The maximum is the $>100^{\text{th}}$ percentile, meaning none of the 50 000 randomly chosen time points exceeded the variance value that was calculated for the actual time points of the rLPs. The plots are based on the same data as those in **a,c**, and **e**.

3.5. Assignment of behavioural phases to tuned neurons and averaged self-administration behaviour

As briefly mentioned before (3.3), the selection of the time windows of the $[Ca^{2+}]_i$ sequences relative to a behavioural events results in different ‘sharpness’ of the resulting averaged $[Ca^{2+}]_i$ sequence. To determine whether or not the apparent broadening of averaged $[Ca^{2+}]_i$ sequences with increasing temporal distance to the rLP timepoint used for the analysis was merely an artefact of the analysis, the same analysis of the calcium sequences was performed with the time window selected relative to the beginning and end of each HE. Using a different behavioural event indeed results in either a more defined $[Ca^{2+}]_i$ peak signal, or the opposite, indicating that the neurons are active in relation to different events that happen during the trial (Figure 16). This was used to define phases of the reward seeking behaviour, based on the behavioural signal that were available (in this case the HE). Due to the lack of other behavioural markers, no additional alignments could be included, hence there are potentially more distinct behavioural correlates to which the neurons may have reacted.

The phases are the pre-HE, the HE and the post-HE phase and are assigned by mutual exclusion from the end of the analysed sequence to its beginning. First calcium signals from all neurons that showed variances $>100^{\text{th}}$ percentile in the bootstrap were extracted from the end of each HE to 5s after the end of the HE. $[Ca^{2+}]_i$ transients of all three available SA sessions were used (Figure 16) and classified according to their maximum value of the averaged $[Ca^{2+}]_i$ peak. If the maximum value was detected after the exit of the HE, the neuron was defined as a post-HE phase neuron. These neurons were subsequently removed from the pool of neurons that need a phase assignment and the procedure was repeated. Similarly, a neuron with a maximum in the averaged $[Ca^{2+}]_i$ transients during the HE was defined as HE-phase neuron. The remaining neurons are assigned as pre-HE neurons (2.8.8).

With the neurons classified in the named phases, the time windows for the analysis of their calcium traces was selected as follows: For pre-HE neurons the analysis window was relative to the rLP, and the HE and post-HE neuron time windows were selected relative to the beginning or end of the HE, respectively. Notably, all neurons show sharp peaks around the time point of their behavioural alignment (Figure 16 panel b). For the pre-HE cells, none of the neurons show a strong averaged $[Ca^{2+}]_i$ transient immediately before the rLP. However given the temporal limitations of $[Ca^{2+}]_i$ imaging, it is hard to determine with which delay, or whether at all the neuron’s $[Ca^{2+}]_i$ transients increased before, during or after the rLP with high temporal precision.

The HE-phase cells have the strongest $[Ca^{2+}]_i$ peaks shortly (within 1-2 seconds) after the beginning of the HE (Figure 16 c). Interestingly, the peaks that are temporarily further away from the start of the HE, appear less sharp than those close to the HE. Also, only very few cells show higher $[Ca^{2+}]_i$ levels outside of the range of the HE.

The post-HE cells show strong averaged $[Ca^{2+}]_i$ peaks shortly after the end of the HE (Figure 16 d). The duration to define the time window before and after the alignment points (rLP, start of HE, end of HE) were chosen to ensure that similar time periods are used for all extracted $[Ca^{2+}]_i$ sequences. This means that each of the time windows that were picked to define the three phases always contain the time-point of the rLP as well as the complete duration of the HE, including a brief period of time after the HE. It also prevents overlap with the $[Ca^{2+}]_i$ sequences of following rLPs, but prevents the possibility to make a definitive statement about the potential loss in tuning of the averaged $[Ca^{2+}]_i$ peaks with greater distance to their alignment. Interestingly however, the averaged $[Ca^{2+}]_i$ transients of post-HE neurons appear to show a decrease in $[Ca^{2+}]_i$ during the time of the highest probability for HE occupation (Figure 16 panel g). This is not the case for the time period before the bulk of the HEs, which indicates a potential inhibition of the cells for times of high probability of a HE occupation.

Overall, this analysis highlights that the time window picked for the analysis is crucial in order to identify whether different cells show a particular tuning to a behavioural event.

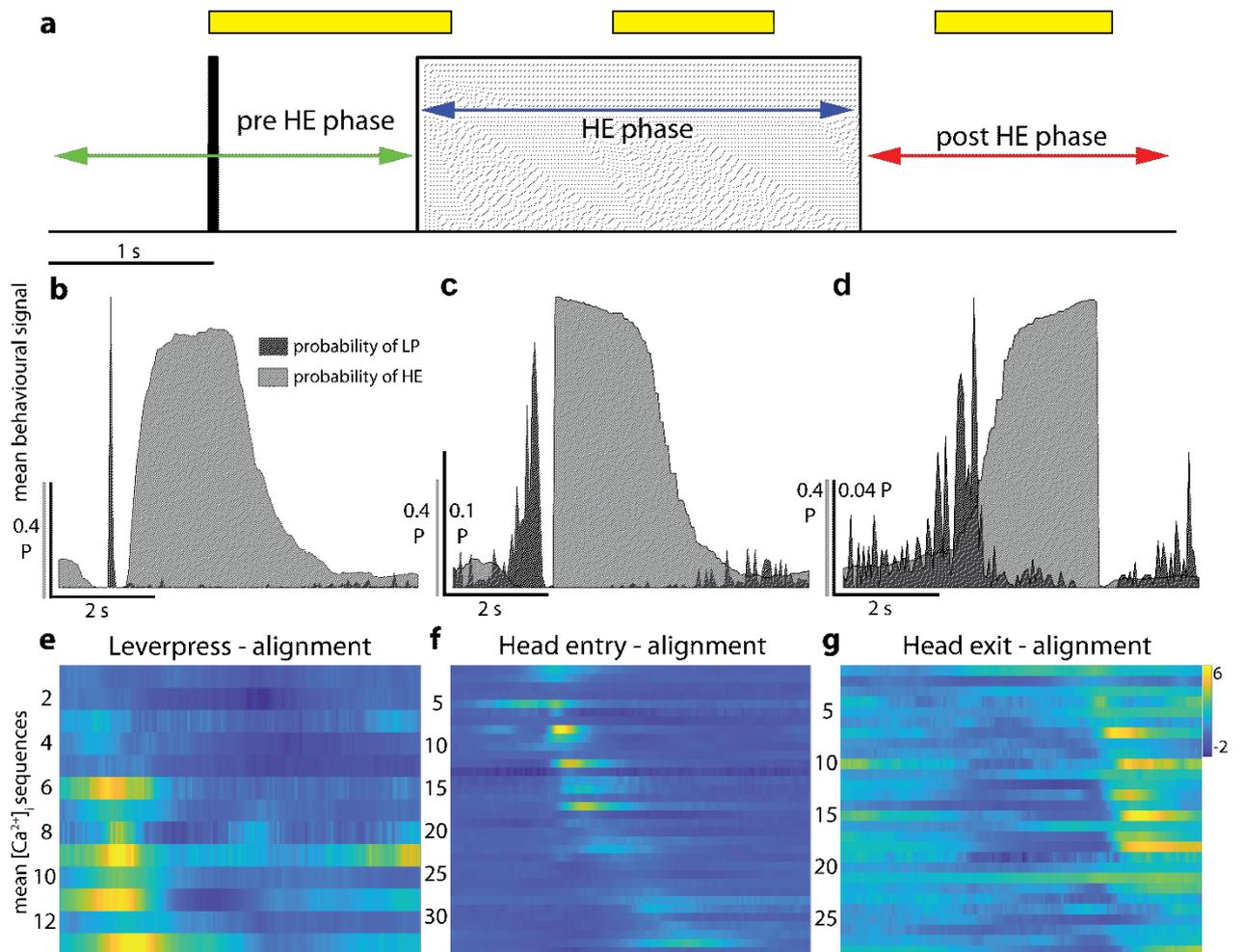


Figure 16: Behavioural phases and the corresponding averaged $[Ca^{2+}]_i$ transients. **a** Schematic of the typical rLP behavioural sequence (7 s) showing the the rLP (black line), the cue-light sequence (yellow rectangles) and the duration of the head entry (grey box). **b-d** show the aligned behaviour data. The probability of an LP is shown in black. The probability of the HE port being occupied is shown in grey. **b** shows the pre-HE phase based on the rLP. **c** shows the HE phase based on the beginning of the HE. **f** shows the post-HE phase based on the end of the HE. **e-g** Averaged $[Ca^{2+}]_i$ transients (z-scored) from all SA sessions. Each row corresponds to a responder cell as identified by the bootstrap analysis (criterion $>100^{\text{th}}$ percentile). $[Ca^{2+}]_i$ transients of all three SA sessions were pooled. **e** shows pre-HE cells with the $[Ca^{2+}]_i$ relative to the rLP. **f** shows HE cells based on the beginning of the HE. **g** shows the post-HE cells based on the end of the HE.

3.6. Matching of cells across multiple sessions

In order to determine, whether the same neurons are active in all or only certain sessions of the behaviour paradigm, the individual spatial footprints of each neuron from each session had to be matched. This is crucial to investigate how, or if at all, the neuronal ensemble which is activated during a particular session changes its composition.

To achieve the tracking of single neurons across multiple sessions, the modified CellReg (Sheintuch, 2017; Sheintuch et al., 2017) pipeline was used (2.8.6). The output is a list of global (for all sessions) cell IDs linking them to each individual session. This information also includes their 'global' centroids, e.g the centroid which accounts for the most likely matching of pairs of spatial footprints between sessions.

The number of neurons that showed detectable $[Ca^{2+}]_i$ transients in each session is shown in Figure 17. Between all sessions the number of detected neurons ranged from 48 to 121 neurons per animal. Overall I observed little difference in the number of detected neurons between sessions within individual animals (SA1: 79 ± 18 neurons; SA2: 88 ± 31 neurons; SA3: 76 ± 26 neurons; RE: 73 ± 24 neurons; overall: 79 ± 22 neurons, [mean \pm std]). However, I found a great variation between individual animals with the average number of neurons detected between sessions being 102 ± 12 neurons for animal 1, 80 ± 4 neurons for animal 2 and 54 ± 7 neurons for animal 3 (mean \pm std). However, these differences were not significant (two-way ANOVA, criterion: $p < 0.05$).

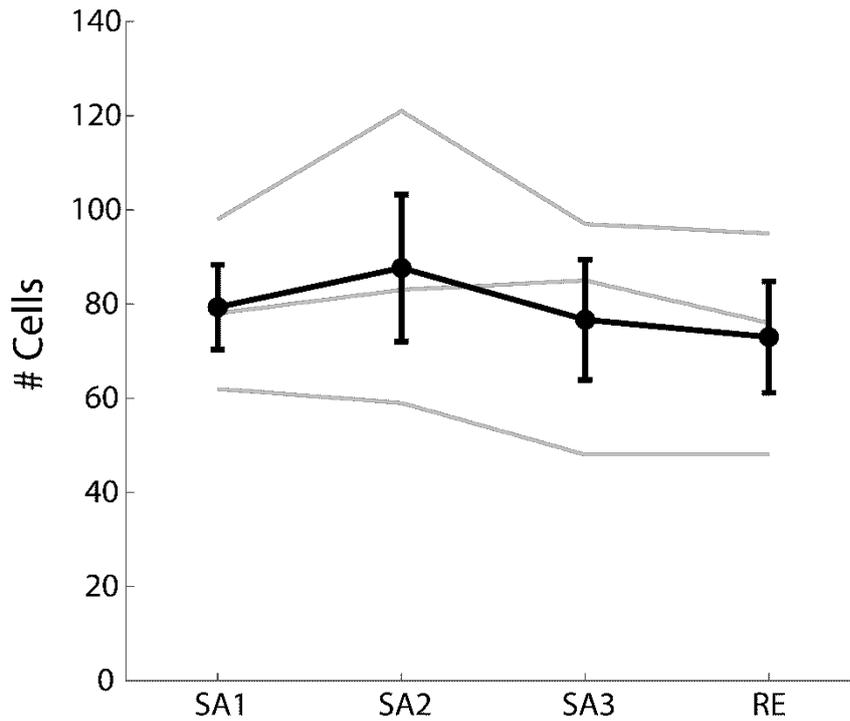


Figure 17: Number of cells detected per session. Grey lines show individual animals; black line shows the average between the animals. N = 3.

Figure 18 shows both the numbers of cells that could be matched across sessions, as well as the accuracy. It is important to note that only cells that showed $[Ca^{2+}]_i$ transients during a session could be detected. Hence, if a neuron was active in session 1 and 3, but not in session 2, the matching percentage between session 1 and 3 will be 100%, and for both compared to session 2 it will be 0% (in case of a single neuron being matched). It is evident that the matching decreases as the number of sessions and thus their difference in time increases (Figure 18a-d). Hence the lowest percentage of overlap can be found between the first SA the RE sessions. Panel d, shows a different representation of the same data, by showing plots of the different possible pairings of sessions.

The differences in the matching of neuronal footprints are high, with values between 32.5% to 68.5% overlap between the SA sessions and as low as 10.3% for some SA sessions compare to the RE session. The averages of the overlap of SA 1 to the other sessions are $56.9\% \pm 12.2$ STD for SA 2, $47.3\% \pm 16.5$ STD for SA 3 and $33.9\% \pm 14.4$ STD for the RE session. For SA 2 the remaining averages are $43.8\% \pm 17.7$ STD for SA 3 and $32.9\% \pm 19.1$ STD for the RE session. For SA 3 the average overlap to the RE session is $38.1\% \pm 24.2$ STD.

This shows that the number of neurons that could be matched constantly declined as the number of recorded sessions increased. This is the case, even though no recognisable change in the field of view (FOV) could be observed between the recorded sessions (data not shown).

Because of the limited number of recorded sessions, it is not immediately accessible whether the relatively low overlap between cells throughout the sessions is a property of the local network, or rather an artefact of the analysis. To investigate the accuracy of the cell matching, the displacement plot of all centroids to their matched pairs was determined (Figure 18 e-g). This revealed that the number of cell pairs within a specific x and y displacement were highly concentrated within a radial distance of 3 μm to the centre of the centroids used. This minute difference indicates, that the spatial arrangement of the neuronal footprints used for the analysis is highly similar, again, suggesting different activity patterns of the network within each session, rather than an error in the measurement.

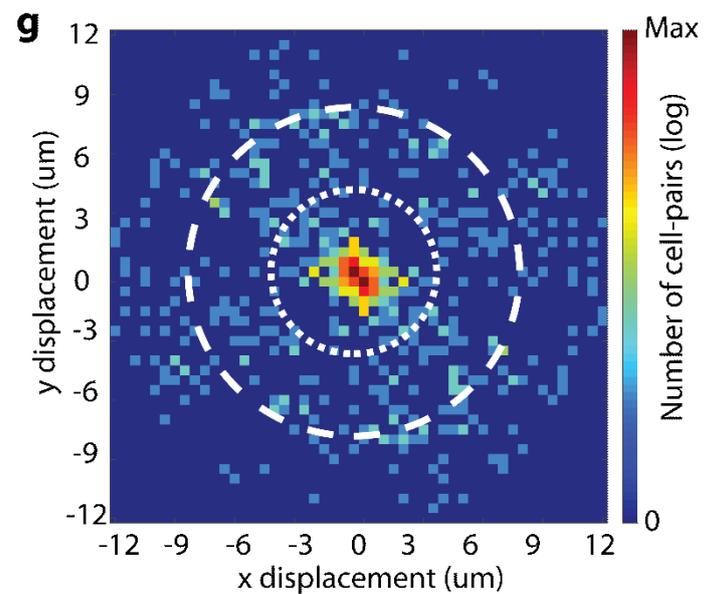
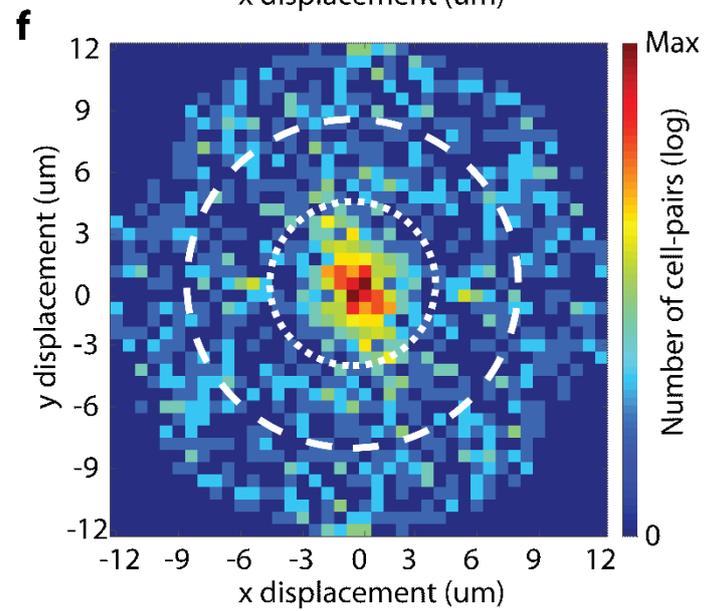
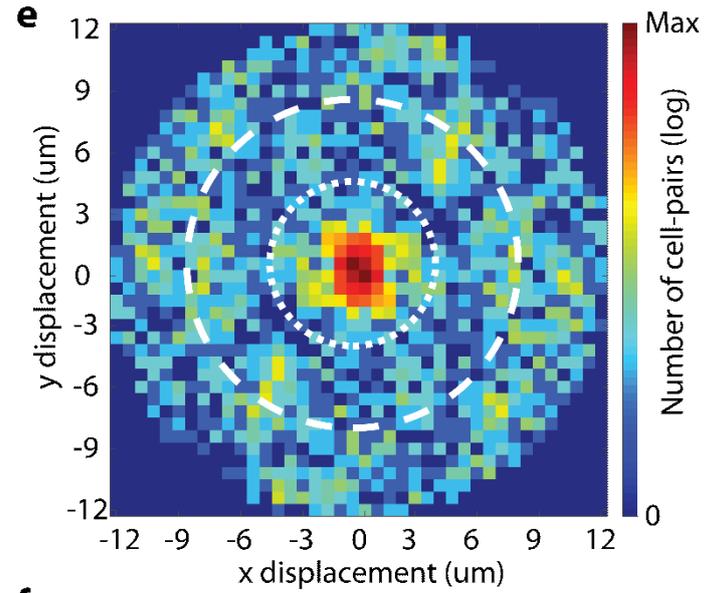
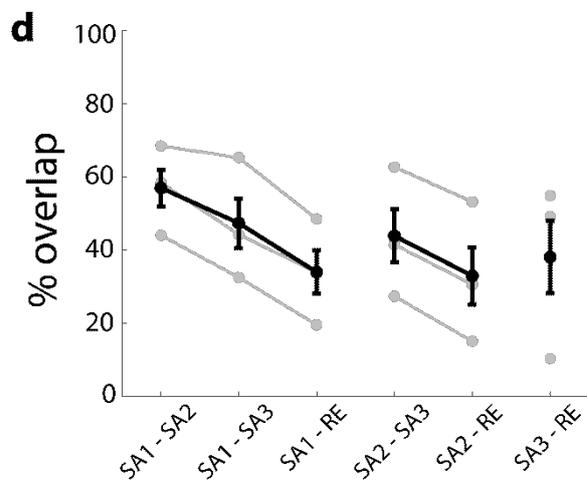
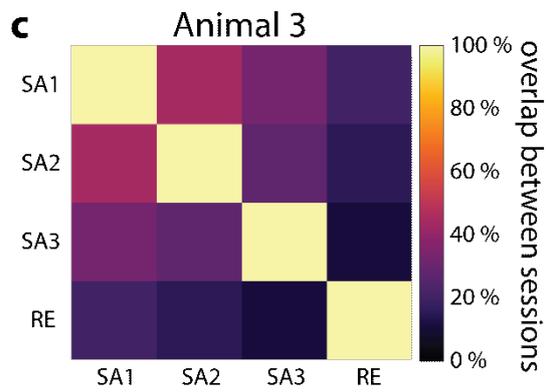
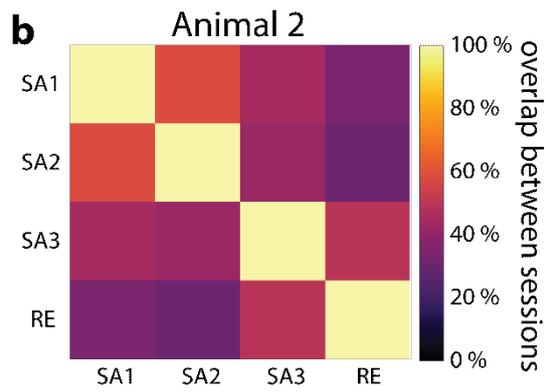
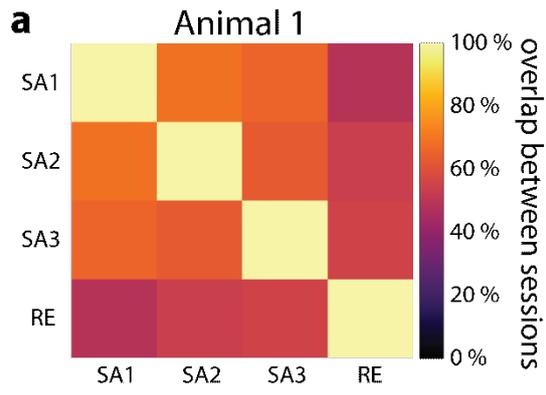


Figure 18: Percentage of cells that could be registered throughout sessions and the X/Y displacement accuracy. **a-c** Cross correlation matrices of the percentage of cells that could be matched in the three SA and the RE session for each animal. Each panel shows the data for an individual animal, with purple colour coding indicating low percentages of the number of overlapping cells between sessions and yellow high a percentage, respectively. **d** Percentage of overlap between individual pairs of sessions. Individual animals are shown in grey; the mean value is shown in black. **e-g** X/Y displacement colour coded of all cell pairs for each animal. High values in the centre indicate successful matching.

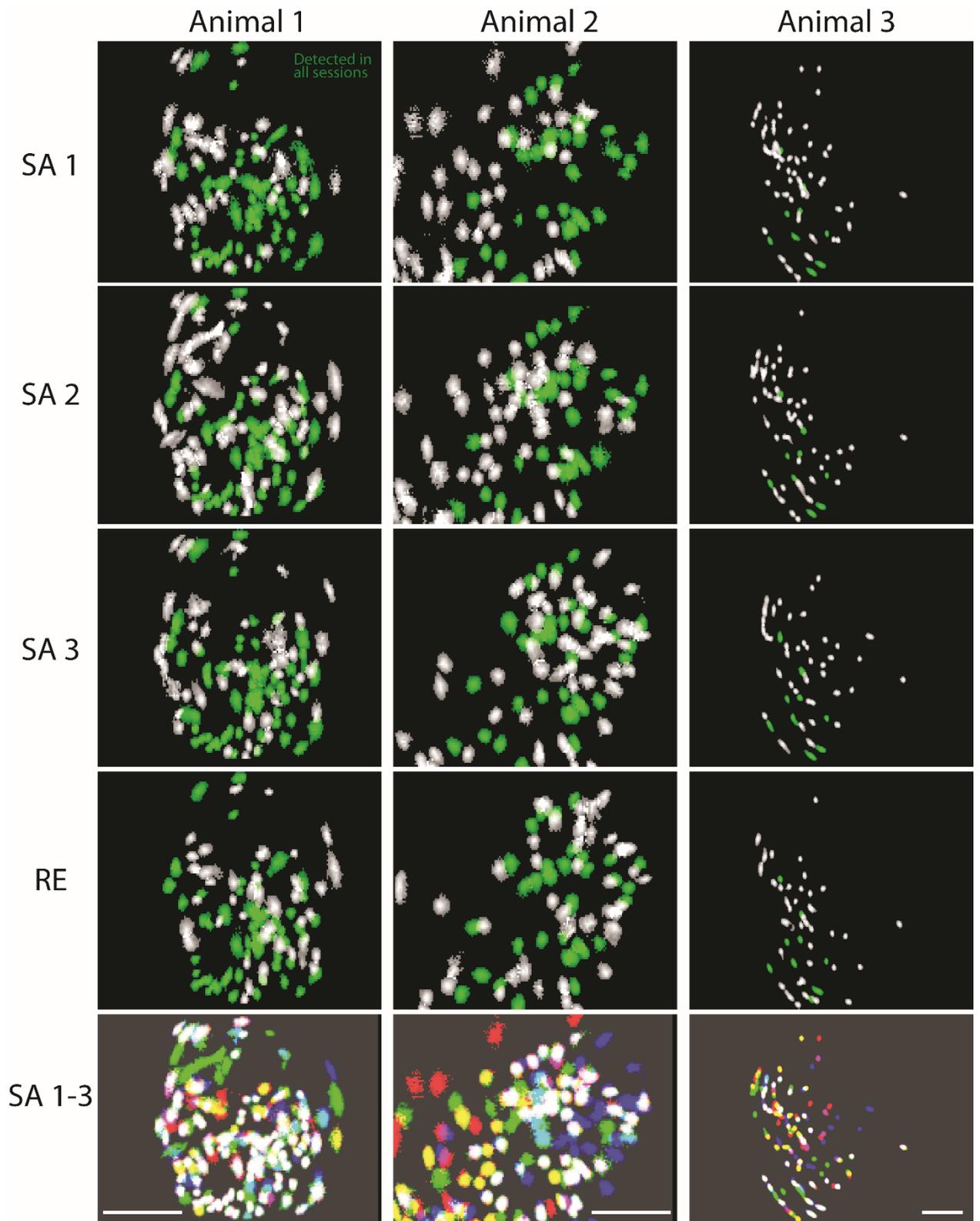


Figure 19: Fields of view for all individually matched sessions and registration of SA sessions. The columns show individual animals and the rows individual sessions. In the top four rows the footprints of the individual sessions are depicted. Green colour in SA 1-3 and RE rows indicates that the cell was matched across all four sessions. The lowest row shows the

registration of all three SA sessions. The neuronal footprints in the FOV are colour coded by session (red for SA 1, green for SA 2 and blue for SA 3). White areas indicate a consensus for all three sessions; yellow (SA 1 on SA 2), purple (SA 1 on SA 3) and turquoise (SA 2 on SA 3) areas for two sessions. White scale bar represents 100 μm .

3.7. Infralimbic network dynamics during operant self-administration

As shown before (Figure 18 and Figure 19), recorded neurons could be tracked during multiple sessions. In addition, some neurons are active during specific actions of the animal during the reward seeking (Figure 15), also referred to as tuning, and this tuning can be classified into three different phases based on the behavioural data (Figure 16).

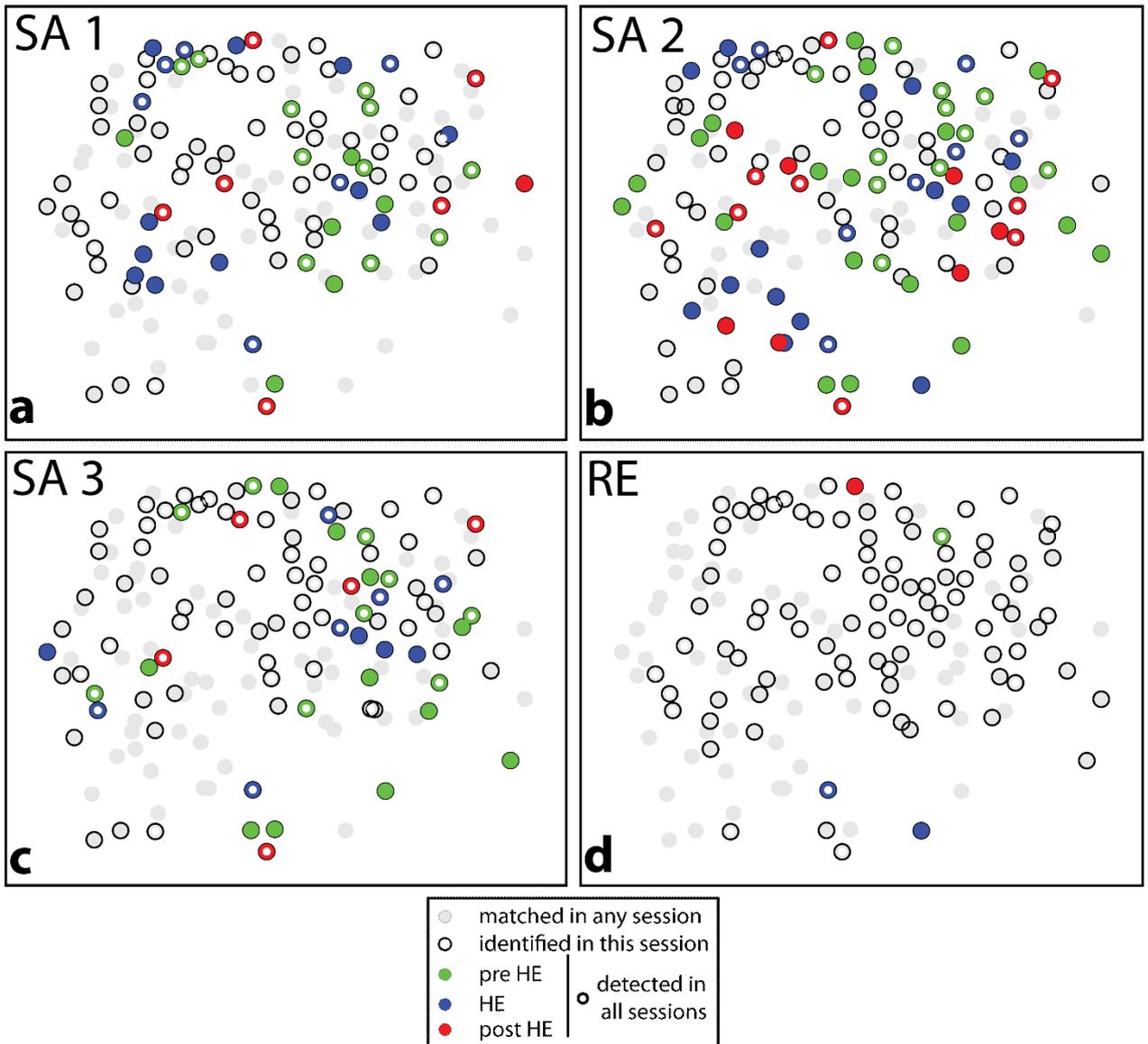
In order to better understand how the network changes its activity patterns, the matched FOVs were analysed in depth to provide information about the tuning of the neurons (Figure 20). A slight modification of Figure 20 is shown in Figure 21, which reduced the information to only the neurons which showed the same tuning in every single SA session. Evidently this stable ensemble is relatively small with 13 neurons for animal 1 and 3 and 6 neurons for animal 2 and 3 respectively. The word tuning hereby describes both a tuning to the self-administration session itself, as well as a tuning to a particular phase during individual trials.

Evidently, patterns of the neuronal ensembles change across sessions. Most cells are not active across all recorded sessions. Notably, if they do show tuning in multiple consecutive sessions, the specific phase to which they are tuned appears relatively stable, although some cells may change their tuning (Figure 22).

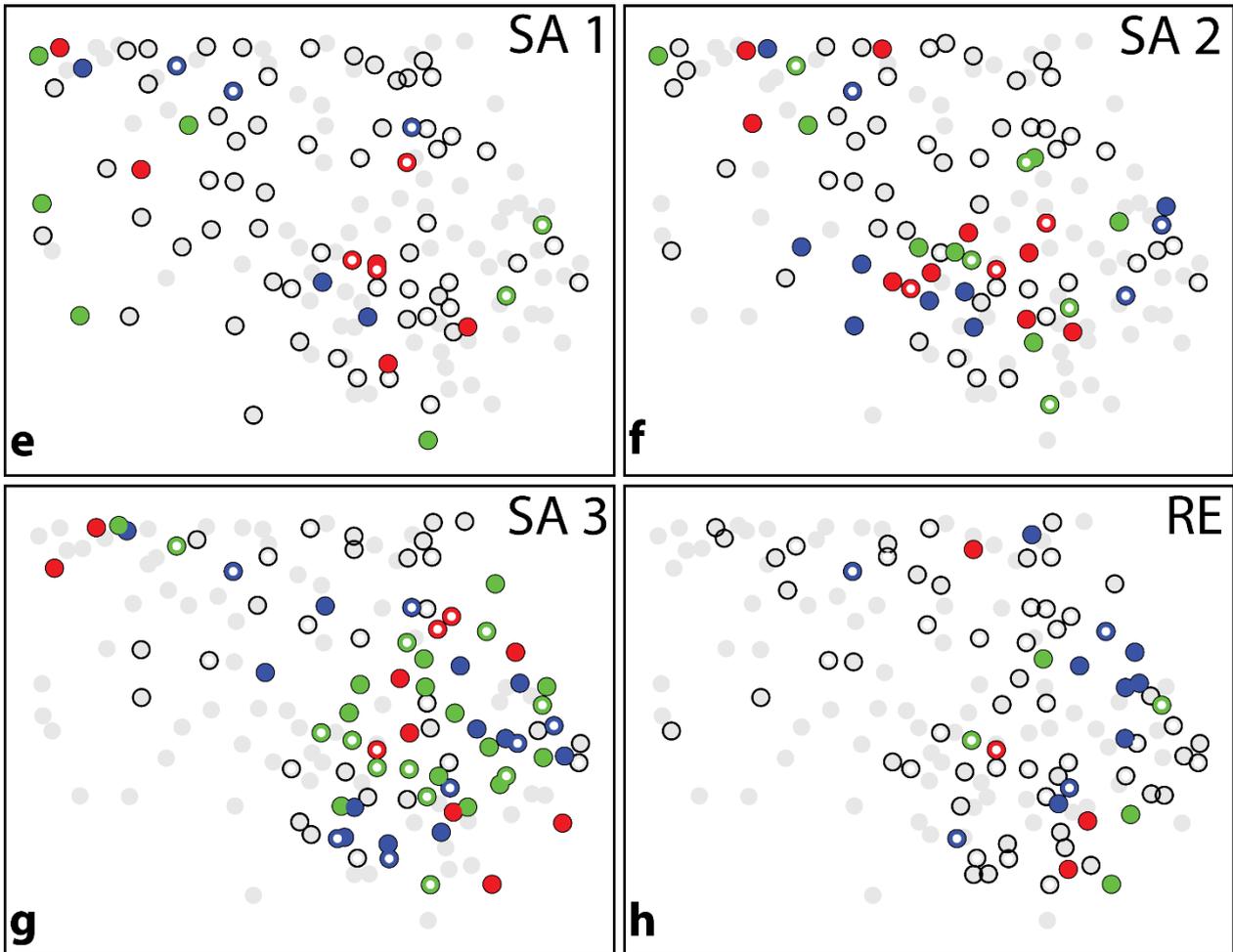
It is apparent, that the number of cells that reach the bootstrap criterion in the RE sessions is much lower than in the SA sessions. The overall number of cells that were detected in the RE sessions, regardless of their tuning, however did not change (Figure 11). This effect, however, can in part be attributed to the lower number of LPs which are available to the analysis (Figure 15) as explained in chapter 3.4.

In conclusion, analysis of the FOV activation patterns (Figure 20) indicate that the network itself constantly changes its configuration by recruiting different neurons to essentially perform the same task. On the other hand, the number of neurons active per sessions seems stable (Figure 17). These fluid like changes of the ensemble configuration are referred to as 'fluidity' for the remaining text.

animal 1



animal 2



animal 3

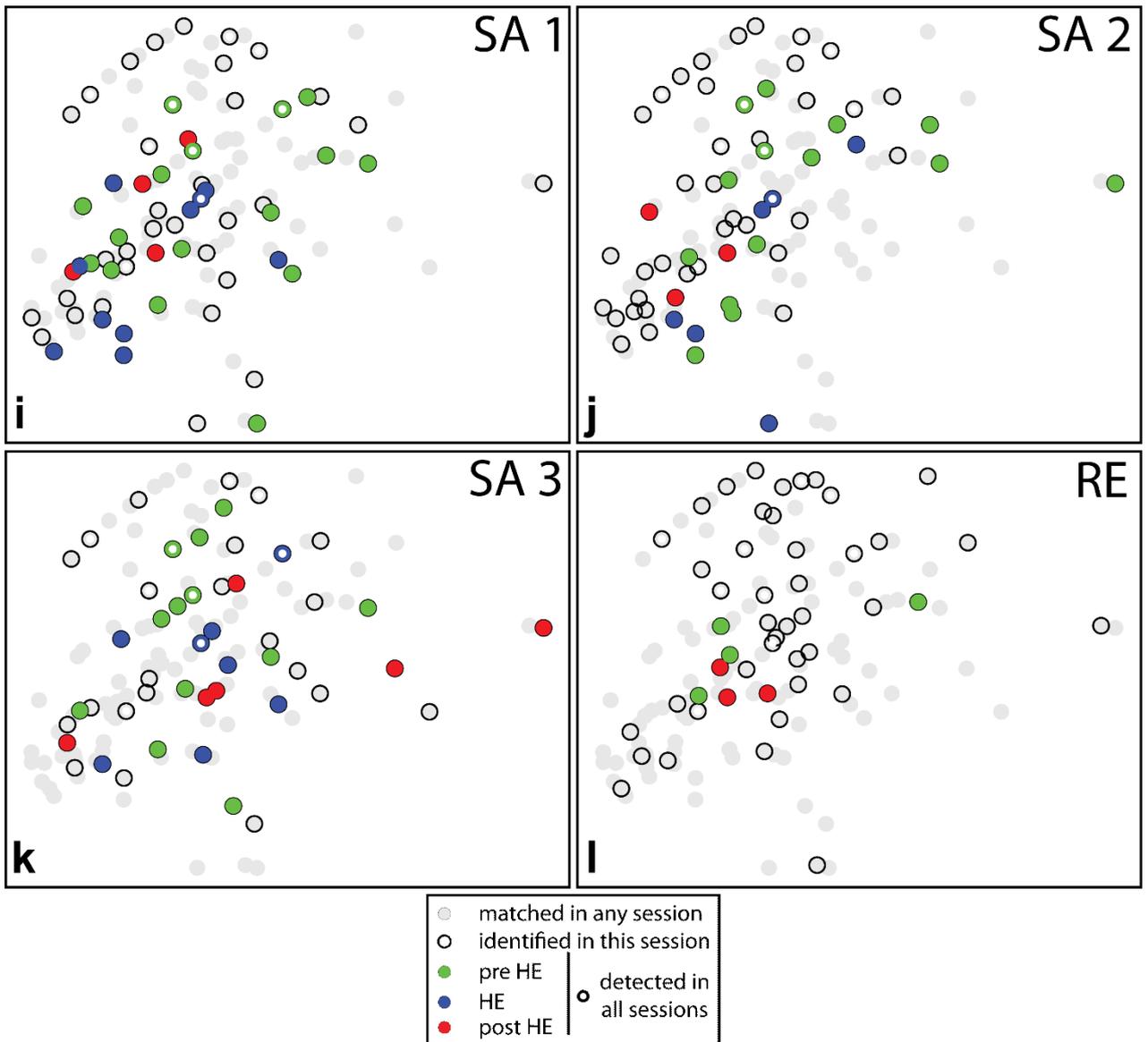
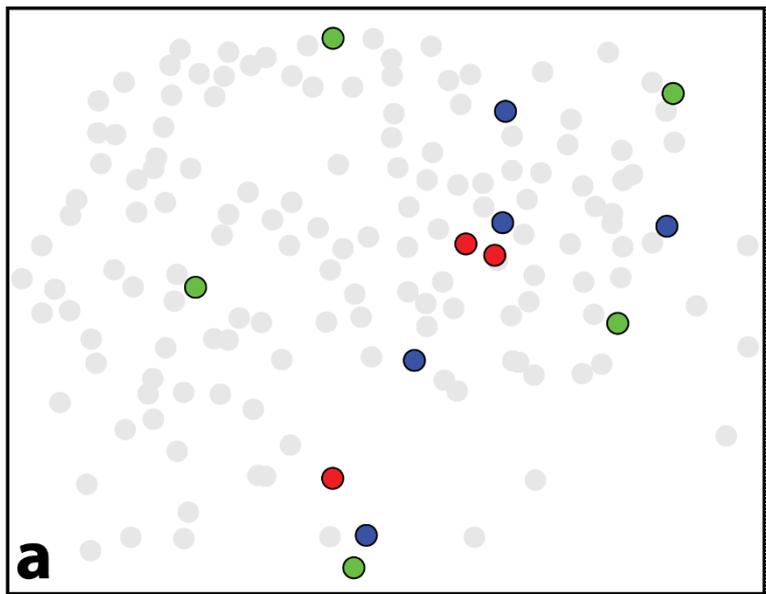
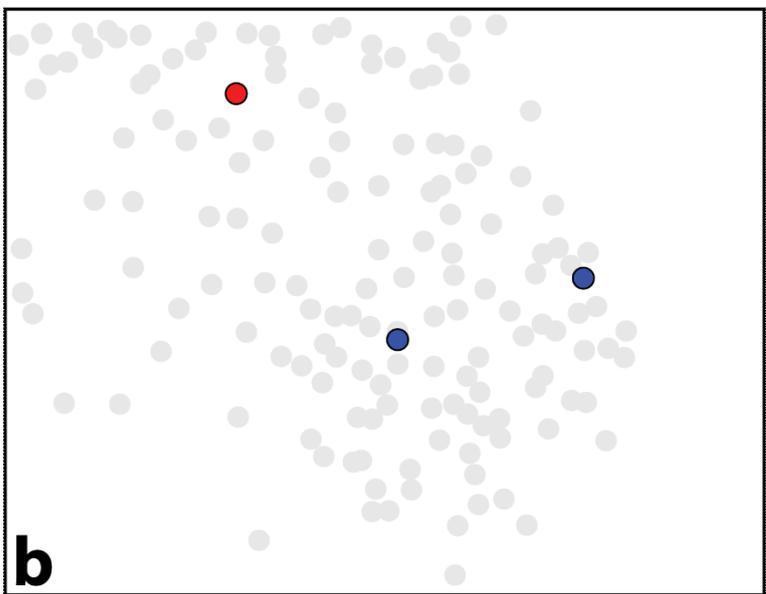


Figure 20: Field of view plots of all individual sessions/animals, colour coded for inter-session stability and phase tuning. (animal 1: a-d; animal 2: e-h; animal 3: i-l). Each dot represents the centroid of a cell and represents the best common matched position across all four sessions. The grey dots show all of the centroids matched between all sessions. Dots which are outlined in black or coloured were active in the session depicted. The colour indicates the phase of this cell for this session (pre-HE: green; HE: blue; post-HE: red). Cells that contain a small white circle in their centre showed calcium transients in all four sessions. This does not mean they also showed tuning.

animal 1



animal 2



animal 3

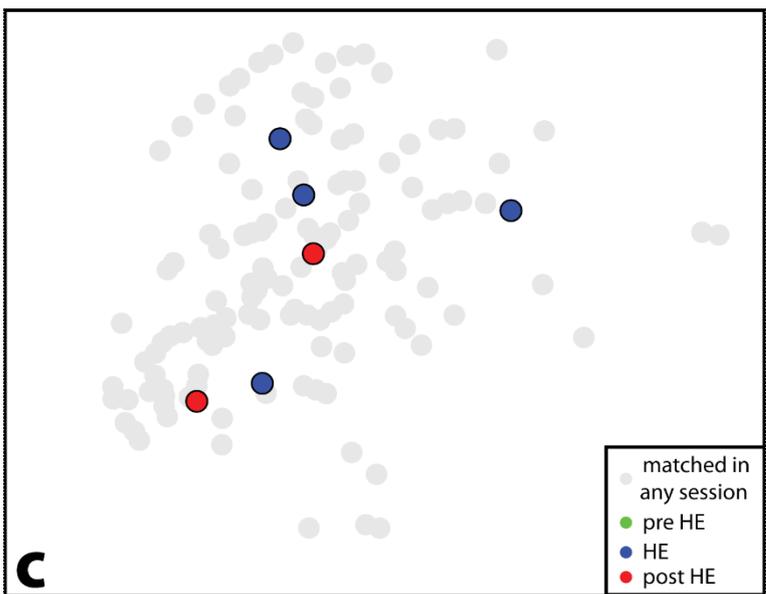


Figure 21: Neurons with stable phase tuning in all SA sessions. a-b show animals 1-3 respectively. Each dot represents the centroid of a cell and represents the best common matched position across all four sessions. Dots which are coloured were active and tuned to the same phase in all 3 SA sessions. The colour indicates the phase.

3.8. Recruitment and stability of neuronal tuning during the operant self-administration paradigm

It is evident from the data that the configuration of the network shows fluidity, between different sessions (Figure 18 and Figure 20). Even though the animals perform the same tasks in similar ways, albeit with an arguable degree of variance (Figure 11), and the overall number of cells that could be captured in the FOV (Figure 11, Figure 17 and Figure 18) are similar across the sessions, the specific cells which are active fluctuate, and so does their tuning to the self-administration task.

The previous analysis was based on the categorization of neuronal tuning and on the predominant response time point of the averaged $[Ca^{2+}]_i$ sequences relative to the behavioral signal. The transients were pooled from all available SA sessions. This behavioural categorization was only done for cells which showed overall tuning to the self-administration behaviour, as determined by the bootstrapping approach (Figure 15). To investigate how stable the phase tuning is between the sessions, this calculation was repeated for the averaged $[Ca^{2+}]_i$ transients of individual sessions (Figure 22). Cells that showed tuning in at least two sessions were then scatter plotted with one axis showing the earliest and the latest time point during which the neuron showed its maximum averaged $[Ca^{2+}]_i$ peak during the individual sessions (Figure 22). It is important to note, that the time displayed on the axes does not follow the same alignment for all of the time points. Seconds 0 - 2 show a time window which was fixed to the rLP (pre-HE phase), seconds 2 - 5 show a time window which was fixed to the beginning of the HE (HE-phase) and seconds 5 - 7 show a time window which was fixed to the end of the HE (post-HE phase). However, in order to better visualize the interaction and possible fluctuation of phases for individual neurons, the time axis were concatenated in this way. It has to be noted, it is not correct to assume that the time points chosen for the alignment are fully capturing the possible range of neuronal tunings the network could exhibit. This is due to the limited amount of information (LPs, HE beginning and end) about the behavioral state. The color coded quadrants in Figure 22 mark the overlap of the three different phases on the two axis (time). Hence, if a neuron's marker lies within one of the colored quadrants, or close to the midline in general, it kept its phase. If it lies in one of the white quadrants it changed its phase between the sessions.

The majority of the cells are within the colored quadrants, with many of them positioned close to the midline, which means they did not change their phase tuning from one session to another. Taking a closer look at the pre-HE phase cells (based on the global marker; grey dot), 8 are positioned within the pre-HE phase quadrant, 1 is close to its border and 3 are deviating into the HE- (1), or the post-HE (2) phase. With one exception, they all show their averaged $[Ca^{2+}]_i$ peak after 1 second. Since the pre-HE phase analysis window was set to 1 second prior to the rLP, this means that only a single cell had its averaged $[Ca^{2+}]_i$ peak before the lever was pressed by the animal. All other neurons show the averaged $[Ca^{2+}]_i$ peak after the rLP, suggesting that no neurons that are involved in a preparatory action of the LP were observed.

In the HE-phase a majority of the neurons are scattered in the lower halves of the quadrant and most of them are placed quite close to the intersection between the pre-HE and the HE phase. Some neurons, are positioned very close to the borders of the phases, especially the pre-HE and HE phase (10). This is likely an artefact of the way the phases were defined. Because there is only incomplete information about the behavioural state of the animal, this analysis is limited to only three signals, namely the LP and the start and end of the HE. Hence even though these cells objectively switched their phase, it is arguable whether this provides meaningful information about their purpose within the network. The lack of cells at the far ends of the phase quadrant (top right corner of the phase quadrant) could be caused by the alignment to the beginning of the HE. Because of this, as the signals occur later in time, they appear smeared because the end of the HE (or next behavioural cue) is an independent event from the start of the HE.

As for the post-HE phase cells, while a large number of them are aligned close to the midline, and a total of 15 are set in the post-HE quadrant, 7 lie within the HE-phase quadrant and three even in the pre-HE quadrant. This is surprising, given that the heat map of the averaged $[Ca^{2+}]_i$ transients for post-HE alignment cells showed signs of an inhibition during the HE phase (Figure 16). This discrepancy could be caused, again, by the way the time window was selected for the analysis which is limited by the recorded information on the animals' behavioural state.

In summary, a majority of the cells keep their tuning to the phases of the self-administration sequence. This is the case, even though the classification relied on limited information about the behavioural state of the animal. So far, it was shown, that the network which is active during the self-administration sequences fluctuates between sessions, meaning that different subsets of neurons from a much larger pool are active throughout the consecutive sessions (Figure 18). It was also shown, that the cells show stability when it comes to time points (phases) within the

self-administration sequence during which they are active (Figure 22). Hence this indicates a fluid network, which retains a stable set of properties.

To better understand how the stability of the neuronal phase preference fits in with the finding of ensemble fluidity, multiple groups were defined in regard to the stability of the phase tuning of individual neurons between sessions (Figure 22 b and c). The first group are neurons which never showed any tuning towards the self-administration and hence could not be assigned a phase. They account for 40% to 50% of the neurons (44% average between animals). The second group are neurons which only showed tuning during a single session. These neurons account for another 30% to 40% (37% average between animals). Of the remaining fraction, 15% to 25% percent (19% average between animals) showed tuning in more than a single session. This fraction was split up into two additional groups; neurons which were assigned the same phase in all session in which they showed tuning, and those which showed tuning to multiple phases. Of the latter, 5% to 10% showed tuning for multiple phases (6% average between animals), while 10% to 20% percent showed stability in their tuning (13% average between animals).

Roughly half the cells showed tuning to the self-administration at some point during the three SA sessions. The majority of those only show tuning a single time, and of those which do show tuning multiple times, the majority retained their phase.

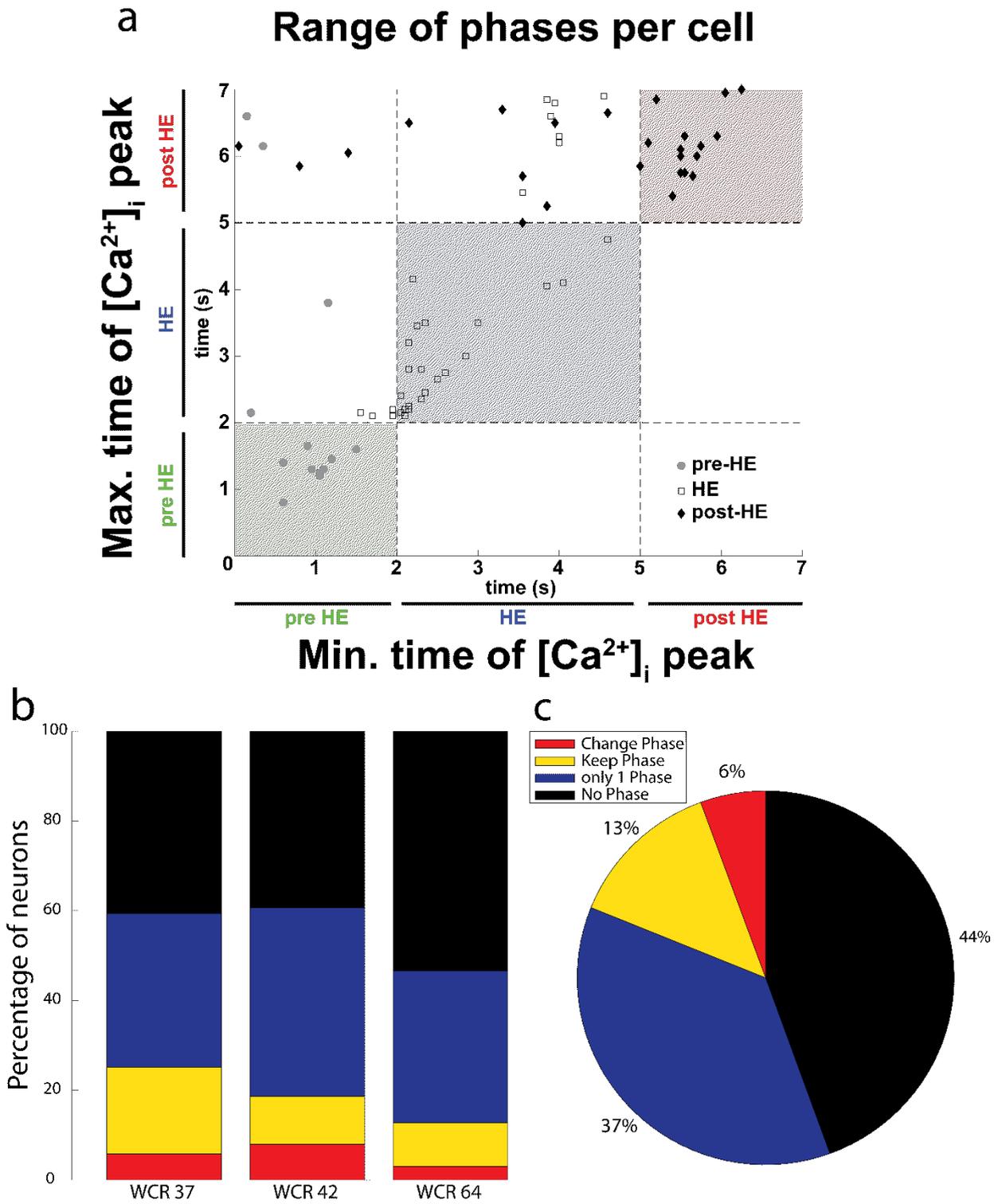


Figure 22: Neuronal tuning across sessions and fluidity of the tuned network. **a** different phase tunings for individual cells for the earliest and latest phase of the three SA sessions, data points are shown as minimum (earliest) and maximum (latest) time points of the $[Ca^{2+}]_i$ peak signal, based on the time window of the respective phase (pre-HE phase: based on the rLP alignment at 1 s; HE phase: based on the alignment on the beginning of the HE at 2 s; post-HE phase: based on an alignment on the post-HE phase at 5 s). Each individual marker shows a cell

that was matched for at least two SA sessions and showed tuning in at least two of the SA sessions with the symbol indicating the phase of the averaged signal of all self-administration sessions (global average). pre-HE phase: grey dot, HE phase: rectangle and post-HE: black diamond). **b** Percentage of neurons within four categories with each bar representing one of the animals. The categories are whether the cells did not show a phase at all (black), whether they only showed a phase a single time (green), whether they kept their phase for all sessions in which they did show tuning (SA and RE, or whether they changed their phases (SA and RE. **c** Percentages for the four categories shown in b as mean of all animals.

4. Discussion

4.1. Summary of findings

This study investigated properties of neuronal ensembles emerging in the IL of the mPFC during operant self-administration. Using microendoscopic imaging of IL neurons expressing a genetically encoded indicator of activity, 50-120 neurons were monitored in each rat over extended periods of time.

Subsets of neurons showed tuning towards reward seeking and were separated into three different phases (pre-, during-, post-head entry). Within individual sessions, neuronal tuning remained stable, forming an ensemble representing the temporal sequence of the reward seeking. Between sessions, the neuronal ensemble reformed from a different subset of neurons, but showed similar compositions of the tuning of the involved neurons. Also tuning of individual neurons across sessions remained stable for most cases, but neurons would not show tuning or even activity in all sessions. Interestingly, a small number of neurons remained stably tuned and active in all sessions examined.

In conclusion, this study revealed neurons in the IL that represent the sequence of behavioural phases associated with learned reward seeking. Distinct ensembles of neurons represent these behavioural phases and remain stable within a session, indicating that the ensembles active in each session act as a cognitive map of the reward seeking behaviour. These ensembles however are formed from different subsets of neurons within each session, reflecting a shift in the neural representation of the reward seeking on the level of ensembles, in spite of identical reward seeking episodes. Single neurons in contrast, often showed a conserved phase preference and few even remained stable in their phase tuning throughout all episodes and may be speculated to conduct the episode-specific ensemble formation based on the current network state. Therefore, IL ensembles are not stable over time despite a highly monotonous action sequence, but regroup for each new session. This represents a novel concept of fluid ensembles in the IL, which therefore could act as frame of reference to assess and potentially correct the outcome of the behaviour.

4.1.1. Neurons are tuned to specific time points of the self-administration sequence

A fraction of the neurons showed a preference to be active at similar time points during the behavioural sequence (Figure 13 and Figure 16), consistent with previous observations of stimulus tuning, including place and grid cells (Hafting, Fyhn, Molden, Moser, & Moser, 2005; O'Keefe & Dostrovsky, 1971), direction-selective cells in V1 cortex or the retina (Cruz-Martín et al., 2014; Hubel & Wiesel, 1959). A difference to these sensory ensembles may be that the IL

ensembles described here are more likely ensembles involved in decision making or initiation of behaviour, as it is shown that they closely follow actions during the reward seeking that cannot simply be reduced to a single sensory modality. E.g. pressing the lever and then entering the HE port was always represented by the activation of neurons that would be specifically activated within certain time points of that sequence of actions. An example for neurons that represent inputs of higher complexity are 'grandmother'-cells (Gorban, Makarov, & Tyukin, 2019). Because of the complexity of the presumably encoded stimulus, it becomes harder to gather enough data to properly correlate the recorded neuronal activity with the complex external stimuli.

An example of this, showing the limitations of the behavioural dataset that was acquired in this study, is the lack of precise information on the time point of reward collection. While the HE information gives a useful indication, it does not inform about whether the animals actually collected the reward. While it appears unlikely that the animal will not collect the reward when they enter the HE port, the information about when precisely the reward was collected is lost as well. Because of this, the tuning could only be coarsely assigned to one of three phases of the behavioural sequence (pre-HE, HE and post-HE, Figure 16). To infer differences in neuronal tuning regarding the association with the HE port versus the reward itself, future studies should gather information about reward collection. This could be implemented through the replacement of the dipper cup (2.6.1) with a port containing a device to register licking.

With one exception, all neurons showed their peak averaged $[Ca^{2+}]_i$ transients after the occurrence of the LP (Figure 22, single data point showing its minimum and maximum $[Ca^{2+}]_i$ peak before 1 s). This is in line with reports suggesting that the PL is crucial in the acquisition of reward seeking behaviours and the IL is involved in its suppression during states of altered reward outcome, because a strong preparatory signal should be expected in an area initiating the behaviour (e.g. extinction, reinstatement, Ashwell & Ito, 2014; Morgan et al., 1993; Quirk et al., 2000). Under the assumption that the IL is involved in the evaluation and subsequent suppression of reward seeking behaviour under unfavourable circumstances (e.g. no reward), it is reasonable to only expect marginal neuronal tuning in the IL prior to the beginning of the behaviour.

Interestingly, during the reinstatement sessions, the composition of cells that are tuned to a particular phase appeared to remain the same (Figure 19). This again partially fits with the proposed role of the IL during reward seeking behaviours, because for the evaluation of the behaviour, tuning to particular events should be expected and the only difference in the behaviour is the collection of the reward, with the reward missing (Ashwell & Ito, 2014). This leads to an unexpected observation, because if the reward is missing it is reasonable to expect changes in the tuning around the HE phase during which time the reward would be 'evaluated'. There is a number of reasons why no change in the neuronal tuning during HE port interactions

might be observed. First, because neurons were still tuned to the HE-, or post-HE phase, this does not mean that the presence or absence of the reward was not encoded. The presence, identity and rating of the reward could be encoded both through individual neurons, as well as ensemble activity, or a combination of both. In addition, given that there is no information about the precise timing of the reward collection, the element of the ensemble which 'encodes' information about the reward could be missing, or is hidden in the imprecision of the alignments used. The $[Ca^{2+}]_i$ transients observed in relation to the exit of the HE port could e.g. simply encode an association to the behavioural task. Also, there may be neurons tuned to the receipt, or the absence of the reward. In this case more data would be needed to determine which case, or both, is true. For both cases the sequence of activations and the composition of the neuronal ensemble would look the same to the observer in the form that it was collected and presented in this study. To solve these issues, two measures can be taken. First, in order to better understand how the reward is encoded in the ensemble activity during each particular session, future studies need to collect information about the precise timing of the reward collection. And second, in order to better understand how precisely the presence, identity and value of a reward is encoded, variations to the behavioural paradigm need to be introduced, such as switching of multiple rewards under the same conditioned stimulus or by increasing the number of LPs which are needed to receive a reward.

4.2. Neuronal tuning and fluid network dynamics during reward seeking

As mentioned previously, studies have found evidence for shared neuronal populations that are active during the reinstatement in studies comparing different types of rewards (Pfarr et al., 2018). However, it has also been shown that neurons in the prefrontal cortex change their firing patterns in response to behavioural changes during rule learning, e.g. the firing patterns change if the animal adapts its strategy to obtain a reward in response to changes in the environment (Durstewitz, Vittoz, Floresco, & Seamans, 2010; Rich & Shapiro, 2009). These findings open a new interpretation, in that instead of the rewards being represented in a small neuronal subpopulation, only changes of the reward type are encoded in individual neurons, or even, that individual neurons do not inform about the type of the reward over longer periods of time and that reward representations which have previously been found are transient.

The prefrontal cortex is crucial in the control of decision making (Kepecs, Uchida, Zariwala, & Mainen, 2008; Sul, Kim, Huh, Lee, & Jung, 2010), exertion of reward encoding (Kvitsiani et al., 2013; Peters et al., 2005) and others, and firing states of the neuronal population in the PL and ACC change, based on the rule of the current behavioural paradigm in a strategy switching task (Malagon-Vina 2018, Rich & Shapiro, 2009). If the animals are required to apply multiple

strategies to obtain a reward, the firing state will change along with the strategy. If a rule that was presented before, is presented a second time, the firing state will not revert back to its state during the first period of that rule, but form a new state (Malagon-Vina et al., 2018). In addition, two separate neuronal populations were observed. The first group positively correlated to the performance while dynamically changing its firing patterns during the behaviour, whereas the other group showed a negative correlation (Malagon-Vina et al., 2018).

The flexible state of the prefrontal network could be a desired property in order to allow the flexible evaluation of constantly changing environments with each different situation requiring a different balance in the inhibition or execution of behaviours (Asaad, Rainer, & Miller, 2000; Malagon-Vina et al., 2018). This requires the network to flexibly adapt to changes in the environment, while retaining enough stability to exclude irrelevant information (Asaad et al., 2000; van Schouwenburg, O'Shea, Mars, Rushworth, & Cools, 2012).

Contrary, in parts, to the studies discussed above is work from Liberti 3rd et al., 2016, which investigated two different neuronal populations within a songbird pre-motor nucleus (HVC). They show that excitatory ensembles undergo drastic shifts, while inhibitory ones show persistent activity.

The mechanisms for the emergence of cortical fluidity are not yet understood, however the evidence for this constant change of the neurons which are active has to be considered for the investigation of the role of the mPFC in addiction. For example, as mentioned above, if two different neuronal ensembles are found for two different rewards, this could reflect an adaptation of the network to a change in the environment in general, rather than the specific encoding of a second reward. Or simply it encodes the change of the reward, and if the change is reversed back a third neuronal population could emerge.

Different neuronal ensembles are present and involved in the self-administration behaviour during different recordings sessions (Figure 18 and Figure 20). A similar type of fluidity has been described in previous studies (Malagon-Vina et al., 2018) during strategy switching tasks. Interestingly, within this fluidity, the 'composition' of tunings to individually defined phases appeared to remain relatively stable, while the neurons which engaged in these ensembles remained the same, as does the preference of individual neurons to be active during a particular phase of the self-administration (see 4.3).

This indicates that the representation of reward seeking behaviour is encoded in a population vector, rather than in a fixed and reliably activated neuronal ensemble. Many questions remain, for example it is unclear why many of the neurons are inactive during whole sessions, or stop to show tuning. Similarly, if they did show tuning, why are cells a lot more likely to show similar

tuning to previous sessions? As discussed in chapter 1.3, neuronal tuning is always dependent on the preceding network, hence for neurons which return to their tuning must likely have been activated through similar sets of cells than before. A remaining possibility which could not be addressed in this study, is that there could be a subpopulation of neurons which guide the network into a particular state, similar to the stable inhibitory ensembles found in the HVC of songbirds (Liberti et al., 2016). In case of such a sub-network comprised of excitatory cells, these would be always activated and then drive the activation of the rest of the population. In order to show this, however, longer series of operant self-administration sessions would have to be recorded to prove the stability of some neurons. As for a potentially inhibitory ensemble, which drives the principal neurons, additional experiments using targeted expression of GCaMP in inhibitory neurons will be necessary.

4.3. Neurons tuned in multiple sessions are likely to keep the initial tuning

The neuronal ensemble involved during the self-administration sequences strongly changed its composition in between the individual sessions. On a level of the individual neurons which comprise this ensemble, the majority of the neurons only participated during a single session (Figure 22b). Of those which showed tuning in multiple sessions however (19%), two thirds (13%) kept their tuning preference and hence remained stable, even though the overall ensemble changed.

In addition, this 'unstable' third of the neurons which responded in multiple sessions could in part be attributed to the way the phase tuning was defined. As can be seen in Figure 22, some of the neurons are close to the border between phases and will hence easily be assigned to different phases in consecutive sessions. This switching of phases is then an artefact of the way the phases were defined, and therefore a limitation of the amount of information that is available about the behavioural state, rather than a property of the neurons. It is unlikely that a 'complete' set of information about the behavioural state of the animal will be available, hence there will always be ambiguity in the definition of behavioural phases and neuronal tuning to them. This suggests that of the remaining neurons which change their phases, there is at least a subset that remained stable, however, the relevant behavioural information was missing in this analysis.

Another source of error could be the way tuning was defined here. In this study increases in the signal of the averaged $[Ca^{2+}]_i$ transients were used to identify relevant neurons. This makes sense in such that $[Ca^{2+}]_i$ indicators are designed to the presence of $[Ca^{2+}]_i$ rather than its absence. If cells are inhibited, this can only be indirectly shown. As can be seen in Figure 16g (post-HE alignment), many of the cells do indeed appear to show an increase in the $[Ca^{2+}]_i$ levels before and after the

HE, and a decrease during the HE. This can lead to neurons being assigned to for example the pre-HE phase in one session and the post-HE phase in the next, even though the more intriguing property of that cell did not show activity during the behavioural sequence, or a particular part of it. Unfortunately, identifying these cell directly through the bootstrap is not possible if the variance is used as a metric. An approach to achieve this could for example be a measurement of the activity of the cell and the subsequent sorting into very high and very low confidence intervals of the bootstrapped distribution. Neurons in the high intervals would have shown a higher than expected activity, while neurons which fall in the lower confidence intervals would have been less active than expected.

In conclusion, the stability of the phase tuning goes in contrast to the fluidity of the overall tuned neuronal ensemble which is active during the self-administration. The way the phase tuning was determined, was limited by the amount of information that was available on the behavioural state of the animal.

4.4. Lack of topological organisation of tuned neurons

No topological organisation of neurons tuned to the behaviour in multiple sessions or during a particular phase could be identified. While this may not be surprising, the FOV provided through the GRIN lenses may obstruct any potentially meaningful arrangement of neurons. In addition, we have no information how the implantation changes the local circuitry. Given that we can still detect tuning of neurons and therefore would likely be able to make meaningful inferences from the $[Ca^{2+}]_i$ transients about the behavioural states, it is unlikely that the GRIN implant strongly disrupted the microanatomy of the IL.

4.5. Technical limitations of microendoscopic recordings

Calcium imaging is a powerful tool for the identification and tracking of neuronal ensembles over longer time periods. The optical access allows separation of individual neurons and even their potential differentiation by projection targets (Cameron, Murugan, Choi, Engel, & Witten, 2019) or genetic background (Krabbe et al., 2019). The temporal resolution is superior to approaches using IEGs, however at the expense of spatial resolution. Compared to electrode recordings however, the temporal resolution imposes constraints on the experimental design and the conclusions that can be drawn from the data. For example, the rise time of GCaMP6f for 10 action potentials is about 80ms, with a decay time of 335ms. While it was shown in neuronal cultures that single action potentials can be measured using GCaMPs, this is only possible under specific

and tightly controlled conditions. Because of this, it is hard to make precise statements about the start and even maximum of $[Ca^{2+}]_i$ signals in relation to behavioural cues. A benefit over electrophysiological measurements is the long term stability of *in-vivo* applications, which allow access to the same cells over longer periods of time and even the manipulation of individual cells based on prior recordings of the same group of neurons (Packer, Russell, Dagleish, & Häusser, 2015). While this is also feasible with electrode recordings (Liberti et al., 2016; Lima et al., 2009), optical access remains the more feasible option due to the easier implementation and flexibility of the methods.

4.5.1. Issues during microendoscopic imaging in rats

Most published endoscopic recordings to date were done in mice and only one in rats (Cameron et al., 2019). A serious limitation encountered in this work concerns the very low yield of analysable experiments.

Regarding the implantation of GRIN lenses and subsequent recordings using UCLA Miniscopes, multiple potential improvements could be made. While multiple issues of the surgeries itself were solved to increase the stability of the implants (2.3 and 3), there remained the problem of placing the viral injections in an optimal position to the lens. The exact reason is not clear and could be due to any number of reasons, from the size of the lens (1 mm) causing tissue compression, the lens holder being misaligned to the injection cannula, up to potential inaccuracies of the stereotactic frames that were used. Since troubleshooting such a complex procedure tends to be very time intensive, especially if the time period between experiments can be multiple weeks, the remaining issue of the lens placement in relation to the injection could be alleviated by coating of AAV particles to the bottom of the GRIN lens using fibroin protein (Jackman et al., 2018). This has been tested in multiple areas (cortex, hippocampus and thalamus) using GRIN implants in mice (data not shown), and the results showed great promise. Calcium transients could be extracted from numbers of cells that were comparable to those in the study presented here (scaling in number to the size of the implanted GRIN lens). Z-score amplitudes were comparable to those reached in the current study (estimated by the brightness of cells given a certain LED excitation power), however the background signal appeared to be reduced. This could likely be an effect of the overall reduced number of neurons that can be virally transduced, with most of them being in close range to the bottom of the implanted lens, hence reducing the amount of signal coming from the large excitation volume outside of the focal plane (1.4.2). Another effect observed was a denser labelling of cells at the edges of the lens compared to the centre. This is likely a problem

of the coating technique of applying multiple drops of fibroin AAV mixture to the lens, which causes larger amounts of 'coating' to dry at the edges.

An additional issue was that rats frequently demolished the miniaturized microscopes. This was largely caused by the fact that the cable was equally interesting to some specimen as the lever, leading to broken soldering connections, or even chewed up cables. The exact number and nature of incidents for individual animals was not recorded, however, most of the rats damaged the cable connecting the Miniscopes to the DAQ board at least once. In order to not disturb the other recordings, every animal was hence assigned an individual Miniscope and spare parts were kept ready to ensure that the recordings could continue immediately or at least for the next scheduled recording session. Using recently released versions of the UCLA Miniscopes (Aharoni, 2016a) that use a wireless connection and locally store the data may lead to less destruction of the material. This however would mean that the synchronisation strategy has to be adapted.

In summary, we successfully established calcium imaging with miniaturized epifluorescence microscopes in combination with an operant self-administration paradigm. Despite the many setbacks, I believe that this approach is suited and feasible for a broader application. The combination of relatively cheap materials which were used (Cai et al., 2016) and a model system which is well developed for a large number of different behavioural approaches is a powerful and versatile tool to better understand how neuronal networks are formed and shaped over time during reward seeking.

4.6. Caveats in the identification of tuned neurons through bootstrapping

The number of neurons that can be defined as tuned, based on the results of the bootstrapping approach, is dependent on two parameters (Figure 15). The percentile which is used to divide tuned from un-tuned neurons can be set by the experimenter and in this study was set rather conservatively at the 100th percentile or above (Figure 14). Because the number of neurons that are identified to be tuned can be massively influenced by the experimenter, it is advisable to not simply rely on comparisons between conditions for a single percentile. Obviously, it is necessary for some of the analyses to pick a percentile, however, the information of how the choice of percentile influenced the number of tuned neurons identified should always be included.

The same goes for the dependence of the number of tuned neurons by the number of LPs that were available for the analysis. Because the number of LPs can influence the resulting number of tuned neurons, sessions which have drastically different numbers of LPs, for example for comparisons between SA and RE sessions, either have to be compared using fixed parameters for

both conditions, e.g. a defined percentile and the largest common number of LPs available, or instead of the direct comparison of the number of tuned neurons, the dependency of tuned neurons on the LP number (Figure 15).

4.6.1. Alternative metrics for bootstrapping

This study used the variance of averaged $[Ca^{2+}]_i$ transients, aligned to the rLPs, to determine if a neuron is tuned to the self-administration. This metric was useful here because it does not rely on the alignment which is used to perfectly represent the tuning to a particular action during the behavioural sequence. Neurons which would show a broader peak in the averaged $[Ca^{2+}]_i$ transients will still have a high variance. Because of this, the initial identification of tuned neurons was solely based on the LP alignment.

This approach of course will not identify neurons showing overall higher activity levels without a preference to be active during a specific time of the behavioural sequence. To identify these neurons and also to refine the identification of neurons tuned to the sequence, different metrics could be used. For example, metrics like the number of peaks in the $[Ca^{2+}]_i$ transients could be applied to estimate general activity levels.

In addition, in order to avoid the caveats of the bootstrapping approach, completely different methods could be used. An alternative approach would be to employ a k-means clustering of the neuronal responses to the available behavioural information, similar to what has been done by (Gründemann et al., 2019). Another interesting approach would be to calculate the mutual information between the $[Ca^{2+}]_i$ transients and the sequence of behavioural events (Timme & Lapish, 2018). However, this is complicated by the independence of the behavioural events and the tuning of the neurons to one particular event, something to which one may or may not have the behavioural correlate available. Hence the number of neurons identified to contain more information about the behavioural state, could be highly dependent on the amount of information available about this state.

It is also feasible to remove the focus onto the short sequences of reward seeking behaviour and instead use the 'whole' recording session for the analysis. For example, the level of synchrony in the network and its individual participants could be calculated (Hamm et al., 2017).

4.6.2. Concluding remarks on the bootstrapping

Using bootstrapping of the variance of averaged $[Ca^{2+}]_i$ transients aligned to the rLPs revealed that simple comparisons between the SA and RE sessions may be biased by the number of LPs the animals performed. While this is to some degree an effect of the high variance of the behaviour data (e.g. number of rLPs (Figure 11)), this can be monitored by observing the results for a parameter space (e.g. percentile and number of rLPs (Figure 15) rather than a fixed set of parameters.

The metrics underlying the bootstrapping approach can be adapted for different questions or improved in other ways. Hence it provides a powerful tool to identify neurons which are tuned to the behaviours for further investigation. It also provides the basis to investigate neural ensemble properties, without introducing an observer bias, e.g. through the application of simple thresholding.

4.7. Conclusions and outlook

The present study showed that neurons in the IL of the mPFC form ensembles which are tuned to an operant self-administration task. These ensembles fluctuate over the course of multiple sessions. This fluctuation however is only apparent in the specific identities of the neurons which are recruited into each transient ensemble. The number of neurons that are part of these ensembles, as well as the composition of its individually tuned neurons remained similar during all SA sessions (RE excluded due to low number of LPs). The tuning preference to a particular time point of individual neurons also showed stability (within the constraints of the methods used). I conclude from this that a larger, stable ensemble exists, of which interim ensembles are recruited within each session. Within the transient ensemble, the sequential progression of the reward seeking is encoded. It remains unclear, if individual neurons carry information about the type of reward or its value.

Specific responses to the rLP of individual neurons (increase in firing rate) have previously been described by Peters et al., 2005, through electrophysiological recordings in the mPFC (specific area not mentioned). They report three types of responses within 63% of the recorded neurons during self-administration of a water reward. Neurons either showed increases in their firing rate prior to the LP, or increases or decreases in their firing rate after the LP. While in this study only a single neuron appeared to show increases in $[Ca^{2+}]_i$ prior to the LP, similar increases after the LP were found and in Figure 16g a decrease in $[Ca^{2+}]_i$ during the head entry period is visible for

some neurons. These early findings are in line with the observation of phase tuning and complement the presented study.

The IL forms transient cognitive maps, specifically referencing behavioural events during reward seeking. This map however is transiently generated from a subset of neurons that show tuning towards specific behavioural events.

4.7.1. Future improvements potential extensions

While this study only focussed on a single reward type (saccharin), and only on the analysis of the SA and the RE sessions, many conclusions about the underlying IL network could be drawn. This however should only be the baseline to establish a set of key network principles (e.g. global stability vs. local fluidity).

Regarding the neuronal tuning, there is a range of experiments that are needed to further resolve the principles of how the reward and its evaluation are encoded in the IL. First, it remains unclear how many neurons remain stably tuned over more than the three sessions which were recorded in this study. Two scenarios are feasible, in the first, the percentage of neurons that keep their tuning between sessions will gradually tail out, indicating the constant 'turnover' of the component of the network which was here described as stable. This would suggest, that individual neurons only play a negligible role in the encoding, putting an emphasis on analysing the ensemble activity further. In the second scenario, a group of neurons would be identified that show stable tuning over a longer time period, suggesting that a small group of neurons acts as nucleators for the ensemble activity. Similar processes to both examples can be observed in the hippocampus where CA1-CA3 neurons show precise and highly context-specific activity which continuously changes and the dentate gyrus shows a stable spatial code with lower context specificity but high stability (Hainmueller & Bartos, 2018).

To further investigate how the nature of the reward and its evaluation is encoded, experiments which alternate rewards or present two different types of rewards (e.g. natural and drug reward) could be employed. In light of experiments showing ensemble changes in the mPFC after strategy changes in behavioural tasks, without a return to previous ensemble states during a previously demanded task, it will be interesting to see how the activated ensemble is shaped during only mildly changed conditions within a session, e.g. by switching the type of reward (Malagon-Vina et al., 2018).

Future studies will need to adapt the operant self-administration in order to be able to identify how, or if, the type of reward is encoded. The same goes for contextual information. The ability to

identify neurons that are preferentially active during defined events of the reward seeking, means that they can be used to infer how differences in the used paradigm are encoded, which in turn allows the inference of the function of the IL. Given that the network shows fluidity between sessions, it will be crucial to further define network parameters, for example the precise composition of neurons that are tuned to definable events, as well as their stability over longer time periods. In addition, in order to better understand how the type of reward may be encoded, it will be necessary to combine different rewards within individual sessions, or decrease the time between sessions to a minimum. This is necessary due to the fresh recruitment of the neuronal ensemble in each session.

Another interesting aspect will be to compare the stability of individual neuronal tunings and the overall composition of the transiently forming ensemble between natural reward seeking and animals which are used as models for addiction. For example, if the IL is crucial in the evaluation of the outcome of a particular behaviour, and if that behaviour is encoded in a neuronal ensemble that is transiently formed, models of addiction could show a range of effects that will cause reward seeking that is detrimental to the organism. Regarding the results of this study, there may be a decrease in the overall fluidity of the network, locking it into an unfavourable stable state. Another possibility is that there is an increase in fluidity and also a decrease in the number of neurons that are tuned towards events in the reward seeking, preventing the IL from reliably evaluating the outcome of a behaviour because the contextual information is missing.

This would be in line with previous findings which showed that the IL is necessary in the evaluation of previously acquired behaviours. Among these are lesion studies showing increased levels of fear, despite extinction training, without changes in the acquisition of the behaviour (Morgan et al., 1993; Quirk et al., 2000), studies of conditioned self-administration which ablated neurons in the IL that were active during conditioning, leading to an increase in reward seeking during the reinstatement (Pfarr et al., 2015), and lastly, increases in the activation of the IL (e.g. through optogenetic manipulation) which decreased fear responses during the presentation of the CS after extinction (Do-Monte et al., 2015). However, this remains speculative until further data was gathered, using more complex operant self-administration paradigms.

Bibliography

- Aharoni, D. (2016a). Miniscope.org. Retrieved from miniscope.org
- Aharoni, D. (2016b). UCLA Miniscope analysis. Retrieved from GitHub website:
https://github.com/daharoni/Miniscope_Analysis
- Aharoni, D. (2016c). UCLA Miniscope CMOS sensor. Retrieved from GitHub website:
https://github.com/daharoni/Miniscope_CMOS_Imaging_Sensor_PCB
- Aharoni, D. (2016d). UCLA Miniscope DAQ Board. Retrieved from GitHub website:
https://github.com/daharoni/Miniscope_DAQ_PCB
- Aharoni, D. (2016e). UCLA Miniscope DAQ board housing. Retrieved from GitHub website:
https://github.com/daharoni/Miniscope_DAQ_Box
- Aharoni, D. (2016f). UCLA Miniscope DAQ software. Retrieved from GitHub website:
https://github.com/daharoni/Miniscope_DAQ_Software
- Aharoni, D. (2016g). UCLA Miniscope machined parts. Retrieved from GitHub website:
https://github.com/daharoni/Miniscope_Machined_Parts
- Aharoni, D., & Hoogland, T. M. (2019). Circuit investigations with open-source miniaturized microscopes: Past, present and future. *Frontiers in Cellular Neuroscience*, 13(April), 1–12.
<https://doi.org/10.3389/fncel.2019.00141>
- Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences*, 13(7), 266–271. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1695401>
- Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, 84(1), 451–459. <https://doi.org/10.1152/jn.2000.84.1.451>
- Ashwell, R., & Ito, R. (2014). Excitotoxic lesions of the infralimbic, but not prelimbic cortex facilitate reversal of appetitive discriminative context conditioning: the role of the infralimbic cortex in context generalization. *Frontiers in Behavioral Neuroscience*, 8, 63.
<https://doi.org/10.3389/fnbeh.2014.00063>
- Baird, G. S., Zacharias, D. A., & Tsien, R. Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), 11241–11246. <https://doi.org/10.1073/pnas.96.20.11241>
- Beyeler, M., Rounds, E. L., Carlson, K. D., Dutt, N., & Krichmar, J. L. (2019). Neural correlates of sparse coding and dimensionality reduction. *PLOS Computational Biology*, 15(6), e1006908. Retrieved from <https://doi.org/10.1371/journal.pcbi.1006908>

- Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., ... others. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature*, *534*(7605), 115–118.
- Cameron, C. M., Murugan, M., Choi, J. Y., Engel, E. A., & Witten, I. B. (2019). Increased Cocaine Motivation Is Associated with Degraded Spatial and Temporal Representations in IL-NAc Neurons. *Neuron*, *103*(1), 80-91.e7. <https://doi.org/10.1016/j.neuron.2019.04.015>
- Carr, D. B., O'Donnell, P., Card, J. P., & Sesack, S. R. (1999). Dopamine terminals in the rat prefrontal cortex synapse on pyramidal cells that project to the nucleus accumbens. *Journal of Neuroscience*, *19*(24), 11049–11060. <https://doi.org/10.1523/jneurosci.19-24-11049.1999>
- Carrillo-Reid, L., Yang, W., Kang Miller, J., Peterka, D. S., & Yuste, R. (2017). Imaging and Optically Manipulating Neuronal Ensembles. *Annual Review of Biophysics*, *46*(1), 271–293. <https://doi.org/10.1146/annurev-biophys-070816-033647>
- Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., ... Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. <https://doi.org/10.1038/nature12354>
- Corbit, L. H., & Balleine, B. W. (2003). The role of prelimbic cortex in instrumental conditioning. *Behavioural Brain Research*, *146*(1–2), 145–157. <https://doi.org/10.1016/j.bbr.2003.09.023>
- Corder, G., Ahanonu, B., Grewe, B. F., Wang, D., Schnitzer, M. J., & Scherrer, G. (2019). An amygdalar neural ensemble that encodes the unpleasantness of pain. *Science*, *363*(6424), 276–281. <https://doi.org/10.1126/science.aap8586>
- Courtin, J., Chaudun, F., Rozeske, R. R., Karalis, N., Gonzalez-Campo, C., Wurtz, H., ... Herry, C. (2014). Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature*, *505*(7481), 92–96. <https://doi.org/10.1038/nature12755>
- Cruz-Martín, A., El-Danaf, R. N., Osakada, F., Sriram, B., Dhande, O. S., Nguyen, P. L., ... Huberman, A. D. (2014). A dedicated circuit links direction-selective retinal ganglion cells to the primary visual cortex. *Nature*, *507*, 358. Retrieved from <https://doi.org/10.1038/nature12989>
- Dana, H., Mohar, B., Sun, Y., Narayan, S., Gordus, A., Hasseman, J. P., ... Kim, D. S. (2016). Sensitive red protein calcium indicators for imaging neural activity. *eLife*, *5*(MARCH2016), 1–24. <https://doi.org/10.7554/eLife.12727>
- Dana, H., Sun, Y., Mohar, B., Hulse, B. K., Kerlin, A. M., Hasseman, J. P., ... Kim, D. S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods*, *16*(7), 649–657. <https://doi.org/10.1038/s41592-019-0435-6>
- Denk, W., Delaney, K. R., Gelperin, A., Kleinfeld, D., Strowbridge, B. W., Tank, D. W., & Yuste, R. (1994). Anatomical and functional imaging of neurons using 2-photon laser scanning microscopy. *Journal of Neuroscience Methods*, *54*(2), 151–162. [https://doi.org/10.1016/0165-0270\(94\)90189-9](https://doi.org/10.1016/0165-0270(94)90189-9)

- Denk, Winfried, Strickler, J. H., & Webb, W. W. (1990). Two-photon laser scanning fluorescence microscopy. *Science*, *248*(4951), 73–76. <https://doi.org/10.1126/science.2321027>
- Diego, F., & Hamprecht, F. A. (2014). Sparse space-time deconvolution for Calcium image analysis. *Advances in Neural Information Processing Systems*, *1*(January), 64–72.
- Do-Monte, F. H., Manzano-Nieves, G., Quiñones-Laracuente, K., Ramos-Medina, L., & Quirk, G. J. (2015). Revisiting the role of infralimbic cortex in fear extinction with optogenetics. *Journal of Neuroscience*, *35*(8), 3607–3615. <https://doi.org/10.1523/JNEUROSCI.3137-14.2015>
- Durstewitz, D., Vittoz, N. M., Floresco, S. B., & Seamans, J. K. (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron*, *66*(3), 438–448. <https://doi.org/10.1016/j.neuron.2010.03.029>
- Firmata. (2015). Firmata. Retrieved from GitHub website: <https://firmata/arduino>
- Gardiner, T. W., & Toth, L. A. (1999). Stereotactic Surgery and Long-Term Maintenance of Cranial Implants in Research Animals. *Contemporary Topics in Laboratory Animal Science*, *38*(1), 56–63.
- George, O., & Hope, B. T. (2017). Cortical and amygdalar neuronal ensembles in alcohol seeking, drinking and withdrawal. *Neuropharmacology*, *122*, 107–114. <https://doi.org/10.1016/j.neuropharm.2017.04.031>
- Ghosh, K. K., Burns, L. D., Cocker, E. D., Nimmerjahn, A., Ziv, Y., Gamal, A. El, & Schnitzer, M. J. (2011). Miniaturized integration of a fluorescence microscope. *Nature Methods*, *8*(10), 871–878. <https://doi.org/10.1038/nmeth.1694>
- Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., ... Pnevmatikakis, E. A. (2019). CalmAn an open source tool for scalable calcium imaging data analysis. *ELife*, *8*, 1–45. <https://doi.org/10.7554/eLife.38173>
- Gorban, A. N., Makarov, V. A., & Tyukin, I. Y. (2019). The unreasonable effectiveness of small neural ensembles in high-dimensional brain. *Physics of Life Reviews*, *29*, 55–88. <https://doi.org/https://doi.org/10.1016/j.plrev.2018.09.005>
- Gourley, S. L., Lee, A. S., Howell, J. L., Pittenger, C., & Taylor, J. R. (2010). Dissociable regulation of instrumental action within mouse prefrontal cortex. *The European Journal of Neuroscience*, *32*(10), 1726–1734. <https://doi.org/10.1111/j.1460-9568.2010.07438.x>
- Gourley, S. L., & Taylor, J. R. (2016). Going and stopping: Dichotomies in behavioral control by the prefrontal cortex. *Nature Neuroscience*, *19*(6), 656–664. <https://doi.org/10.1038/nn.4275>
- Greenberg, D. S., & Kerr, J. N. D. (2009). Automated correction of fast motion artifacts for two-photon imaging of awake animals. *Journal of Neuroscience Methods*, *176*(1), 1–15. <https://doi.org/10.1016/j.jneumeth.2008.08.020>
- Grewe, B. F., Gründemann, J., Kitch, L. J., Lecoq, J. A., Parker, J. G., Marshall, J. D., ... Schnitzer, M. J. (2017).

- Neural ensemble dynamics underlying a long-term associative memory. *Nature*, 543(7647), 670–675. <https://doi.org/10.1038/nature21682>
- Gründemann, J., Bitterman, Y., Lu, T., Krabbe, S., Grewe, B. F., Schnitzer, M. J., & Lüthi, A. (2019). Amygdala ensembles encode behavioral states. *Science*, 364(6437), eaav8736. <https://doi.org/10.1126/science.aav8736>
- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052), 801–806. <https://doi.org/10.1038/nature03721>
- Hainmueller, T., & Bartos, M. (2018). Parallel emergence of stable and dynamic memory engrams in the hippocampus. *Nature*, 558(7709), 292–296. <https://doi.org/10.1038/s41586-018-0191-2>
- Hamm, J. P., Peterka, D. S., Gogos, J. A., & Yuste, R. (2017). Altered Cortical Ensembles in Mouse Models of Schizophrenia. *Neuron*, 94(1), 153–167.e8. <https://doi.org/10.1016/j.neuron.2017.03.019>
- Hammond, S. L., Leek, A. N., Richman, E. H., & Tjalkens, R. B. (2017). Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection. *PLoS ONE*, 12(12), 1–22. <https://doi.org/10.1371/journal.pone.0188830>
- Hebb, D. O. (1949). The organization of behavior; a neuropsychological theory. In *The organization of behavior; a neuropsychological theory*. Oxford, England: Wiley.
- Heidbreder, C. A., & Groenewegen, H. J. (2003). The medial prefrontal cortex in the rat: Evidence for a dorso-ventral distinction based upon functional and anatomical characteristics. *Neuroscience and Biobehavioral Reviews*, 27(6), 555–579. <https://doi.org/10.1016/j.neubiorev.2003.09.003>
- Heimer, L., Alheid, G. F., De Olmos, J. S., Groenewegen, H. J., Haber, S. N., Harlan, R. E., & Zahm, D. S. (1997). The accumbens: Beyond the core-shell dichotomy. *Journal of Neuropsychiatry and Clinical Neurosciences*, 9(3), 354–381. <https://doi.org/10.1176/jnp.9.3.354>
- Helmchen, F., & Denk, W. (2005). Deep tissue two-photon microscopy. *Nature Methods*, 2(12). <https://doi.org/10.1038/NMETH818>
- Helmchen, F., & Sakmann, B. (1996). *Ca²⁺ + Buffering and Action Potential-Evoked Ca²⁺ + Signaling in Dendrites of Pyramidal Neurons*. 70(February).
- Hockel, K., Trabold, R., Schöller, K., Török, E., & Plesnila, N. (2012). Impact of anesthesia on pathophysiology and mortality following subarachnoid hemorrhage in rats. *Experimental and Translational Stroke Medicine*, 4(1), 5. <https://doi.org/10.1186/2040-7378-4-5>
- Hoover, W. B., & Vertes, R. P. (2007). Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. *Brain Structure and Function*, 212(2), 149–179. <https://doi.org/10.1007/s00429-007-0150-4>
- Horton, J. C., & Adams, D. L. (2005). The cortical column: a structure without a function. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360(1456), 837–862.

<https://doi.org/10.1098/rstb.2005.1623>

- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, *148*(3), 574–591. <https://doi.org/10.1113/jphysiol.1959.sp006308>
- Jackman, S. L., Chen, C. H., Chettih, S. N., Neufeld, S. Q., Drew, I. R., Agba, C. K., ... Regehr, W. G. (2018). Silk Fibroin Films Facilitate Single-Step Targeted Expression of Optogenetic Proteins. *Cell Reports*, *22*(12), 3351–3361. <https://doi.org/10.1016/j.celrep.2018.02.081>
- Jacob, A. D., Ramsaran, A. I., Mocle, A. J., Tran, L. M., Yan, C., Frankland, P. W., & Josselyn, S. A. (2018). A Compact Head-Mounted Endoscope for In Vivo Calcium Imaging in Freely Behaving Mice. *Current Protocols in Neuroscience*, *84*(1), 1–29. <https://doi.org/10.1002/cpns.51>
- Kalivas, P. W. (2009). The glutamate homeostasis hypothesis of addiction. *Nature Reviews Neuroscience*, *10*(8), 561–572. <https://doi.org/10.1038/nrn2515>
- Kalivas, P. W., & Nakamura, M. (1999). Neural systems for behavioral activation and reward. *Current Opinion in Neurobiology*, *9*(2), 223–227. [https://doi.org/10.1016/S0959-4388\(99\)80031-2](https://doi.org/10.1016/S0959-4388(99)80031-2)
- Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–231. <https://doi.org/10.1038/nature07200>
- Kirschbaum, E., Hausmann, M., Wolf, S., Sonntag, H., Schneider, J., Elzoheiry, S., ... Hamprecht, F. A. (2018). *LeMoNAde: Learned Motif and Neuronal Assembly Detection in calcium imaging videos*. 1–28. Retrieved from <http://arxiv.org/abs/1806.09963>
- Koester, H. J., & Sakmann, B. (2000). Calcium dynamics associated with action potentials in single nerve terminals of pyramidal cells in layer 2/3 of the young rat neocortex. *The Journal of Physiology*, *529 Pt 3*, 625–646. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2270226&tool=pmcentrez&rendertype=abstract>
- Koob, G. F., Ahmed, S. H., Boutrel, B., Chen, S. A., Kenny, P. J., Markou, A., ... Sanna, P. P. (2004). Neurobiological mechanisms in the transition from drug use to drug dependence. *Neuroscience and Biobehavioral Reviews*, *27*(8), 739–749. <https://doi.org/10.1016/j.neubiorev.2003.11.007>
- Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: a neurocircuitry analysis. *The Lancet Psychiatry*, *3*(8), 760–773. [https://doi.org/10.1016/S2215-0366\(16\)00104-8](https://doi.org/10.1016/S2215-0366(16)00104-8)
- Koob, G., & Simon, E. (2009). The neurobiology of addiction: Where we have been and where we are going. *Journal of Drug Issues*, *39*(1), 1–17. Retrieved from <http://jod.sagepub.com/content/39/1/115.short>
- Krabbe, S., Paradiso, E., d'Aquin, S., Bitterman, Y., Courtin, J., Xu, C., ... Lüthi, A. (2019). Adaptive disinhibitory gating by VIP interneurons permits associative learning. *Nature Neuroscience*, *22*(11),

1834–1843. <https://doi.org/10.1038/s41593-019-0508-y>

- Kvitsiani, D., Ranade, S., Hangya, B., Taniguchi, H., Huang, J. Z., & Kepecs, A. (2013). Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature*, *498*(7454), 363–366. <https://doi.org/10.1038/nature12176>
- Lasseter, H. C., Xie, X., Ramirez, D. R., & Fuchs, R. A. (2010). Prefrontal cortical regulation of drug seeking in animal models of drug relapse. *Current Topics in Behavioral Neurosciences*, *3*, 101–117. https://doi.org/10.1007/7854_2009_19
- Liberti, W. A., Markowitz, J. E., Perkins, L. N., Liberti, D. C., Leman, D. P., Guitchounts, G., ... Gardner, T. J. (2016). Unstable neurons underlie a stable learned behavior. *Nature Neuroscience*, *19*(12), 1665–1671. <https://doi.org/10.1038/nn.4405>
- Liberti, W. A., Perkins, L. N., Leman, D. P., & Gardner, T. J. (2017). An open source, wireless capable miniature microscope system. *Journal of Neural Engineering*, *14*(4), 45001. <https://doi.org/10.1088/1741-2552/aa6806>
- Lima, S. Q., Hromádka, T., Znamenskiy, P., & Zador, A. M. (2009). PINP: A New Method of Tagging Neuronal Populations for Identification during In Vivo Electrophysiological Recording. *PLOS ONE*, *4*(7), e6099. Retrieved from <https://doi.org/10.1371/journal.pone.0006099>
- Lopes, G., Bonacchi, N., Frazão, J., Neto, J. P., Atallah, B. V., Soares, S., ... Kampff, A. R. (2015). Bonsai: an event-based framework for processing and controlling data streams. *Frontiers in Neuroinformatics*, Vol. 9, p. 7. Retrieved from <https://www.frontiersin.org/article/10.3389/fninf.2015.00007>
- Lu, J. (2018). Min1pipe. Retrieved from GitHub website: <https://github.com/JinghaoLu/MIN1PIPE>
- Lu, J., Li, C., Singh-Alvarado, J., Zhou, Z. C., Fröhlich, F., Mooney, R., & Wang, F. (2018). MIN1PIPE: A Miniscope 1-Photon-Based Calcium Imaging Signal Extraction Pipeline. *Cell Reports*, *23*(12), 3673–3684. <https://doi.org/10.1016/j.celrep.2018.05.062>
- Lütcke, H., Murayama, M., Hahn, T., Margolis, D. J., Astori, S., Zum Alten Borgloh, S. M., ... Hasan, M. T. (2010). Optical recording of neuronal activity with a genetically-encoded calcium indicator in anesthetized and freely moving mice. *Frontiers in Neural Circuits*, *4*(April), 9. <https://doi.org/10.3389/fncir.2010.00009>
- Malagon-Vina, H., Ciochi, S., Passecker, J., Dorffner, G., & Klausberger, T. (2018). Fluid network dynamics in the prefrontal cortex during multiple strategy switching. *Nature Communications*, *9*(1), 1–13. <https://doi.org/10.1038/s41467-017-02764-x>
- Marlin Firmware. (2015). Retrieved from GitHub website: <https://github.com/MarlinFirmware/Marlin>
- Maruyama, R., Maeda, K., Moroda, H., Kato, I., Inoue, M., Miyakawa, H., & Aonishi, T. (2014). Detecting cells using non-negative matrix factorization on calcium imaging data. *Neural Networks*, *55*, 11–19. <https://doi.org/10.1016/j.neunet.2014.03.007>

- Maslowski-Cobuzzi, R. J., & Napier, T. C. (1994). Activation of dopaminergic neurons modulates ventral pallidal responses evoked by Amygdala stimulation. *Neuroscience*, *62*(4), 1103–1119. [https://doi.org/10.1016/0306-4522\(94\)90347-6](https://doi.org/10.1016/0306-4522(94)90347-6)
- McDonald, A. J., Mascagni, F., & Guo, L. (1996). Projections of the medial and lateral prefrontal cortices to the amygdala: a Phaseolus vulgaris leucoagglutinin study in the rat. *Neuroscience*, *71*(1), 55–75. [https://doi.org/https://doi.org/10.1016/0306-4522\(95\)00417-3](https://doi.org/https://doi.org/10.1016/0306-4522(95)00417-3)
- Morgan, M. A., Romanski, L. M., & LeDoux, J. E. (1993). Extinction of emotional learning: Contribution of medial prefrontal cortex. *Neuroscience Letters*, *163*(1), 109–113. [https://doi.org/10.1016/0304-3940\(93\)90241-C](https://doi.org/10.1016/0304-3940(93)90241-C)
- Mukamel, E. A., Nimmerjahn, A., & Schnitzer, M. J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron*, *63*(6), 747–760. <https://doi.org/10.1016/j.neuron.2009.08.009>
- Nagai, T., Sawano, A., Eun Sun Park, & Miyawaki, A. (2001). Circularly permuted green fluorescent proteins engineered to sense Ca²⁺. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(6), 3197–3202. <https://doi.org/10.1073/pnas.051636098>
- Nakai, J., Ohkura, M., & Imoto, K. (2001). A high signal-to-noise ca²⁺ probe composed of a single green fluorescent protein. *Nature Biotechnology*, *19*(2), 137–141. <https://doi.org/10.1038/84397>
- Neufeld, S. Q., Hen, R., Kheirbek, M. A., Kass, R. E., Pnevmatikakis, E. A., Rodriguez-Romaguera, J., ... Zhou, P. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *ELife*, *7*, 1–37. <https://doi.org/10.7554/elife.28728>
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. [https://doi.org/https://doi.org/10.1016/0006-8993\(71\)90358-1](https://doi.org/https://doi.org/10.1016/0006-8993(71)90358-1)
- Ongur, D., & Price, J. (2000). The Organization of Networks within the Orbital and Medial Prefrontal Cortex of Rats, Monkeys and Humans. *Cerebral Cortex*, *10*(3), 206–219. <https://doi.org/10.1093/cercor/10.3.206>
- Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., ... Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, *543*(7643), 103–107. <https://doi.org/10.1038/nature21376>
- Packer, A. M., Russell, L. E., Dagleish, H. W. P., & Häusser, M. (2015). Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. *Nature Methods*, *12*(2), 140–146. <https://doi.org/10.1038/nmeth.3217>
- Paredes, R. M., Etzler, J. C., Watts, L. T., & Lechleiter, J. D. (2009). *Chemical Calcium Indicators*. *46*(3), 143–151. <https://doi.org/10.1016/j.ymeth.2008.09.025.Chemical>

- Pavlov, I. P. (1927). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. In *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford, England: Oxford Univ. Press.
- Peters, Y. M., O'Donnell, P., & Carelli, R. M. (2005). Prefrontal cortical cell firing during maintenance, extinction, and reinstatement of goal-directed behavior for natural reward. *Synapse*, *56*(2), 74–83. <https://doi.org/10.1002/syn.20129>
- Pfarr, S., Meinhardt, M. W., Klee, M. L., Hansson, A. C., Vengeliene, V., Schönig, K., ... Sommer, W. H. (2015). Losing control: Excessive alcohol seeking after selective inactivation of cue-responsive neurons in the infralimbic cortex. *Journal of Neuroscience*, *35*(30), 10750–10761. <https://doi.org/10.1523/JNEUROSCI.0684-15.2015>
- Pfarr, S., Schaaf, L., Reinert, J. K., Paul, E., Herrmannsdörfer, F., Roßmanith, M., ... Sommer, W. H. (2018). Choice for drug or natural reward engages largely overlapping neuronal ensembles in the infralimbic prefrontal cortex. *Journal of Neuroscience*, *38*(14), 3507–3519. <https://doi.org/10.1523/JNEUROSCI.0026-18.2018>
- Pnevmatikakis, E. A. (2019). Analysis pipelines for calcium imaging data. *Current Opinion in Neurobiology*, *55*, 15–21. <https://doi.org/10.1016/j.conb.2018.11.004>
- Pnevmatikakis, E. A., & Giovannucci, A. (2017). NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods*, *291*, 83–94. <https://doi.org/10.1016/j.jneumeth.2017.07.031>
- Pomrenze, M., Baratta, M., Cadle, B., & Cooper, D. (2012). Cocaine self-administration in the mouse: A low-cost, chronic catheter preparation. *Nature Precedings*, 3–4. <https://doi.org/10.1038/npre.2012.7040.1>
- Printrun. (2015). Pronterface. Retrieved from GitHub website: <https://github.com/kliment/Printrun>
- Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *Journal of Neuroscience*, *20*(16), 6225–6231. <https://doi.org/10.1523/jneurosci.20-16-06225.2000>
- Rich, E. L., & Shapiro, M. (2009). Rat prefrontal cortical neurons selectively code strategy switches. *Journal of Neuroscience*, *29*(22), 7208–7219. <https://doi.org/10.1523/JNEUROSCI.6068-08.2009>
- Richmond, B. J. (2009). *Information Coding* (L. R. B. T.-E. of N. Squire, Ed.). <https://doi.org/https://doi.org/10.1016/B978-008045046-9.00220-5>
- Roelfsema, P. R., & de Lange, F. P. (2016). Early Visual Cortex as a Multiscale Cognitive Blackboard. *Annual Review of Vision Science*, *2*(1), 131–151. <https://doi.org/10.1146/annurev-vision-111815-114443>
- Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, *18*(1), 1–26.

<https://doi.org/10.1186/s12859-017-1934-z>

- Samson, H. H., Pfeffer, A. O., & Tolliver, G. A. (1988). Oral Ethanol Self-administration in Rats: Models of Alcohol-Seeking Behavior. *Alcoholism: Clinical and Experimental Research*, *12*(5), 591–598.
<https://doi.org/10.1111/j.1530-0277.1988.tb00248.x>
- Savard, C., Lema, P. P., Hélie, P., & Vachon, P. (2009). Effects of timing of dexamethasone treatment on the outcome of collagenase-induced intracerebral hematoma in rats. *Comparative Medicine*, *59*(5), 444–448.
- Scofield, M., & Kalivas, P. (2014). Astrocytic Dysfunction and Addiction: Consequences of Impaired Glutamate Homeostasis. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *20*. <https://doi.org/10.1177/1073858413520347>
- Scott, S. H. (2000). Population vectors and motor cortex: neural coding or epiphenomenon? . *Nature Neuroscience*, *3*(4), 307–308. <https://doi.org/10.1038/73859>
- Seabrook, T. A., Burbridge, T. J., Crair, M. C., & Huberman, A. D. (2017). Architecture, Function, and Assembly of the Mouse Visual System. *Annual Review of Neuroscience*, *40*(1), 499–538.
<https://doi.org/10.1146/annurev-neuro-071714-033842>
- Sesack, S. R., & Bunney, B. S. (1989). Pharmacological characterization of the receptor mediating electrophysiological responses to dopamine in the rat medial prefrontal cortex: A microiontophoretic study. *Journal of Pharmacology and Experimental Therapeutics*, *248*(3), 1323–1333.
- Sesack, Susan R., Carr, D. B., Omelchenko, N., & Pinto, A. (2003). Anatomical Substrates for Glutamate-Dopamine Interactions: Evidence for Specificity of Connections and Extrasynaptic Actions. *Annals of the New York Academy of Sciences*, *1003*, 36–52. <https://doi.org/10.1196/annals.1300.066>
- Sheintuch, L. (2017). CellReg. Retrieved from GitHub website: <https://github.com/zivlab/CellReg>
- Sheintuch, L., Rubin, A., Brande-Eilat, N., Geva, N., Sadeh, N., Pinchasof, O., & Ziv, Y. (2017). Tracking the Same Neurons across Multiple Days in Ca²⁺ Imaging Data. *Cell Reports*, *21*(4), 1102–1115.
<https://doi.org/10.1016/j.celrep.2017.10.013>
- Sheng, M. jun, Lu, D., Shen, Z. ming, & Poo, M. ming. (2019). Emergence of stable striatal D1R and D2R neuronal ensembles with distinct firing sequence during motor learning. *Proceedings of the National Academy of Sciences of the United States of America*, *166*(22), 11038–11047.
<https://doi.org/10.1073/pnas.1901712116>
- Sherrington, C. (1906). *The integrative action of the nervous system*. New Haven: CT Yale University Press.
- Skinner, B. F. (1938). The behavior of organisms: an experimental analysis. In *The behavior of organisms: an experimental analysis*. Oxford, England: Appleton-Century.
- Skocek, O., Nöbauer, T., Weilguny, L., Martínez Traub, F., Xia, C. N., Molodtsov, M. I., ... Vaziri, A. (2018).

- High-speed volumetric imaging of neuronal activity in freely moving rodents. *Nature Methods*, 15(6), 429–432. <https://doi.org/10.1038/s41592-018-0008-0>
- Spanagel, R. (2009). Alcoholism: A systems approach from molecular physiology to addictive behavior. *Physiological Reviews*, 89(2), 649–705. <https://doi.org/10.1152/physrev.00013.2008>
- Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, 66(3), 449–460. <https://doi.org/10.1016/j.neuron.2010.03.033>
- Swanson, L. W. (1982). The projections of the ventral tegmental area and adjacent regions: A combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Research Bulletin*, 9(1–6), 321–353. [https://doi.org/10.1016/0361-9230\(82\)90145-9](https://doi.org/10.1016/0361-9230(82)90145-9)
- Timme, N. M., & Lapish, C. (2018). A Tutorial for Information Theory in Neuroscience. *Eneuro*, 5(3), ENEURO.0052-18.2018. <https://doi.org/10.1523/ENEURO.0052-18.2018>
- Todd, T. P., Vurbic, D., & Bouton, M. E. (2014). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiology of Learning and Memory*, 108, 52–64. <https://doi.org/10.1016/j.nlm.2013.08.012>
- Tschechne, S., & Neumann, H. (2014). Hierarchical representation of shapes in visual cortex—from localized features to figural shape segregation. *Frontiers in Computational Neuroscience*, 8, 93. <https://doi.org/10.3389/fncom.2014.00093>
- van Schouwenburg, M. R., O’Shea, J., Mars, R. B., Rushworth, M. F. S., & Cools, R. (2012). Controlling human striatal cognitive function via the frontal cortex. *Journal of Neuroscience*, 32(16), 5631–5637. <https://doi.org/10.1523/jneurosci.6428-11.2012>
- Vertes, R. P. (2004). Differential Projections of the Infralimbic and Prelimbic Cortex in the Rat. *Synapse*, 51(1), 32–58. <https://doi.org/10.1002/syn.10279>
- Volkow, N. D., Koob, G. F., & McLellan, A. T. (2016). Neurobiologic advances from the brain disease model of addiction. *New England Journal of Medicine*, 374(4), 363–371. <https://doi.org/10.1056/NEJMra1511480>
- Wallis, J. D. (2018). Decoding Cognitive Processes from Neural Ensembles. *Trends in Cognitive Sciences*, 22(12), 1091–1102. <https://doi.org/10.1016/j.tics.2018.09.002>
- Walter, J., You, Q., Hagstrom, J. N., Sands, M., & High, K. A. (1996). Successful expression of human factor IX following repeat administration of an adenoviral vector in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 93(7), 3056–3061. <https://doi.org/10.1073/pnas.93.7.3056>
- Weeks, J. R. (1962). Experimental morphine addiction: Method for automatic intravenous injections in unrestrained rats. *Science*, 138(3537), 143–144. <https://doi.org/10.1126/science.138.3537.143>

Wise, R. A. (2002). Brain reward circuitry: Insights from unsensed incentives. *Neuron*, 36(2), 229–240.
[https://doi.org/10.1016/S0896-6273\(02\)00965-0](https://doi.org/10.1016/S0896-6273(02)00965-0)

Zhang, L., Liang, B., Barbera, G., Hawes, S., Zhang, Y., Stump, K., ... Lin, D. T. (2019). Miniscope GRIN Lens System for Calcium Imaging of Neuronal Activity from Deep Brain Structures in Behaving Animals. *Current Protocols in Neuroscience*, 86(1), 1–21. <https://doi.org/10.1002/cpns.56>

Zhou, P. (2016). CNMF-E. Retrieved from GitHub website: https://github.com/zhoup/CNMF_E

Acknowledgements

It would not have been possible to complete this PhD thesis without the support, encouragement and inspiration of many people, whom I want to thank here.

First of all, I want to sincerely thank my advisor Prof. Dr. Thomas Kuner for giving me the opportunity to work in his lab. He gave me guidance throughout the course of my PhD, fostered critical thinking and independence and most of all remained patient throughout many setbacks. I am also deeply thankful to Apl. Prof. Dr. Wolfgang Sommer for the fruitful collaboration and many discussions on the project. I also want to thank Prof. Dr. Rainer Spanagel, Dr. Claudio Acuna Goycolea and Dr. Amit Agarwal for being members in my committee and for participating in the evaluation of this work.

Next I want to thank the many members of the Department of Functional Neuroanatomy. I thank Michaela Kaiser and Claudia Kocksch for AAV preparations as well as Marion Schmitt for providing technical assistance and help with anything in the lab. Furthermore, I'd like to thank Dr. Janine Reinert for initiating the Tracer project with me, for enduring my chaos as an office mate and for keeping me hydrated after long days in the lab. I thank Dr. Christoph Körber for helping me deal with ethics proposals and Christoph and Juan Carlos Boffi for many discussions on the SFB1134 A04 project. I thank Ursula Lindenberger for always standing up for me, even during the most blatant bureaucratic faults I produced, for organizing the lab drinks with me, always lending an ear and most importantly for keeping everything running. Furthermore, I would like to thank Johannes Knabbe, Varun Venkatarami, Livia Asan, Frank Hermannsdörfer, Anna Steinmann, Dimitar Tanev, Christopher Strahle, Rafaela Stehle, Gerald Bendner, Areej Albariri, Helia Saber, Steffen Sass, Maja Klevanski, Heinz Horstmann, Simone Hoppe and Patric Pelzer for both many fruitful discussions as well as many discussions on everything else serving indispensable procrastination.

Many thanks also go to Dr. Sidney Cambridge for enduring my upbeat attitude and for introducing me to the "Waldpiraten". I needed both.

I am grateful to Andreas Draguhn, Susanne Bechtel and everyone else involved, for initiating the DFG grant "SFB1134" which provided funding for this project and my position, and provided the opportunity for many additions of the project. I also appreciate the in- and output from all the members of the calcium imaging club who participated in creating a calcium imaging network in Heidelberg neuroscience community. I also want to thank Simone Pfarr, Janet Barroso-Flores, Laura Schaaf and Rebecca Hoffmann for working with me on the SFB1134 B04 project and especially Ariel Iporre-Rivas for the countless discussions and help with the analysis.

Most importantly I wish to thank my family, especially Margarete, Andrea and Elsa. I feel deeply privileged to have grown up in such a warm and loving family and am delighted to have Paul (middle) as a brother in law. I also want to thank my 'Kiwi' family, Jessie, Paul (big), Kiri and Zoe who always supported me no matter where in the world they resided. Furthermore, I want to thank my in-law-family Ruth and Volker, Annika, Christoph, Paul (small) and Franz, as well as Sarah, Frank and Nora for your support through years of studies and for always providing a joyful refuge close to Heidelberg. This dissertation would not have been possible without my new-built family; my wife Hannah Sonntag. I thank you for loving me, believing in me and supporting me and for offering valuable and professional advice throughout the PhD, as well as teaching me that there is a life after the lab. Without your encouragement I would not even have been able to finish the Bachelor, let alone this thesis.

Lastly I would like to thank Rainer Sonntag, my father, Ulla Jug my friend and Shawn Mikula, without whom I wouldn't have been suited for a PhD. Their attitude towards life, friendship and science will always remain a part of me.