Prof. Abraham Nixon (Examiner)

1. Author reports that different operant conditioning including stimulus detection tasks, Delayed Non-Match to Sample (DNMS), as well as Go/No-Go tasks were tried as part of optimization. Author concluded these tasks as failed ones. However, clear reasoning for the failure is not explained. It would be useful if author explains this with details.

New section added: "Operant conditioning experiments failed to match requirements" to Chapter 2 – "Behaviour".

<u>*** ADDED (to lines ~2431-2458):</u>
Operant conditioning tasks have been extensively and successfully modeled in a variety of laboratories. For our specific experiments, we required a task that could be learnt within 1-2 weeks. This was because we were not confident on how many simultaneous days of chronic imaging, we could achieve with the *in vivo* chronic 2-photon calcium imaging preparation (see Chapter 3 – "Imaging"). Additionally, it takes typically 3-4 weeks for rodents to acquire sufficient expertise in behavioural performance for Operant Conditioning tasks (Guo et al., 2014).

One alternative that we could have tried was to train the animals to expert levels of performance, and subsequently performed the hippocampus prep. The issue(s) with this is that,
a) We wanted to study the hippocampal CA1 network during the learning or acquisition phase of behavioural training, as a distinct experiment from those published in literature.
b) In such a protocol, we would require two separate surgeries, viz., i) Head-bar implant, and, ii) hippocampus preparations. This, we believed would increase the technical difficulty of the overall experiment and could be more stressful for the experiment animals.
c) We suspected that there could be unknown effects on behaivoural performance, post-surgery, complicating the analysis and insights we aimed to study.

Across all Operant Conditioning protocols attempted, we could not observe behavioural discrimination (not licking to incorrect phases) within 1-2 weeks of training. We considered the experiments as failures, in line with this reasoning.

2. Page 75, line no. 1681 & Page 79, line no. 1755: Author used only 2-4 mice for the stimulus detection tasks and concluded as the failed ones. Making conclusions from this low number of experimental animals is not convincing. What was the reason for using only 2 mice for certain protocols and 4 mice for the other ones?

*** ADDED (to lines ~2253-2256):
Each animal was presented with 600 training trials/day (1 session/day). The inability to behaviourally discriminate or withhold incorrect or unrewarded licks even for 7-14 sessions, was considered and the task was ultimately unsuccessful.

New section added: "Operant conditioning experiments failed to match requirements" to Chapter 2 – "Behaviour".

*** ADDED (to lines ~2430-2460):
Operant conditioning tasks have been extensively and successfully modeled in a variety of laboratories. For our specific experiments, we required a task that could be learnt within 1-2 weeks. This was because we were not confident on how many simultaneous days of chronic imaging, we could achieve with the *in vivo* chronic 2-photon calcium imaging preparation (see Chapter 3 – "Imaging"). Additionally, it takes typically 3-4 weeks for rodents to acquire sufficient expertise in behavioural performance for Operant Conditioning tasks (Guo et al., 2014).

One alternative that we could have tried was to train the animals to expert levels of performance, and subsequently performed the hippocampus prep. The issue(s) with this is that,
a) We wanted to study the hippocampal CA1 network during the learning or acquisition phase of behavioural training, as a distinct experiment from those published in literature.
b) In such a protocol, we would require two separate surgeries, viz., i) Head-bar implant, and, ii) hippocampus preparations. This, we believed would increase the technical difficulty of the overall experiment and could be more stressful for the experiment animals.
c) We suspected that there could be unknown effects on behaivoural performance, post-surgery, complicating the analysis and insights we aimed to study.

Across all Operant Conditioning protocols attempted, we could not observe behavioural discrimination (not licking to incorrect phases) within 1-2 weeks of training. This was cited as the reason why we considered the experiments as failures.

3. Page 79, line no. 1771: What was the reasoning for selecting specific frequencies in the audible range? Please explain.

We attempted to follow Jaramillo & Zador (2014) for the specific values used.

<u>* ADDED (to lines ~2343-2351):</u>**
We referenced previously published protocols (Jaramillo and Zador, 2014) for their selection of auditory tone frequencies (3 kHz - 16.3 kHz) to select frequencies in the audible range. We had designed the experiment in such a way that the animal could behaviourally express a response to the perception of the various stimuli presented, as well as having learnt the overall task. We tried to incorporate more tones, in the hope that this may improve the chances of the animals focusing on the task specifics, instead of producing licks to just any stimulus.

4. Page 80, line no. 1789: Please explain the reasons for the failure of DNMS task.

The reasoning to abandon the Operant Conditioning paradigm was the same for DNMS, Go/No-Go, or Stimulus detection, viz., lack of behavioural discrimination within 1-2 weeks of training.
Response included in new added text. Please see response to comment 1.

<u>* ADDED (to lines ~2430-2460):</u>**
Operant conditioning tasks have been extensively and successfully modeled in a variety of laboratories. For our specific experiments, we required a task that could be learnt within 1-2 weeks. This was because we were not confident on how many simultaneous days of chronic imaging, we could achieve with the *in vivo* chronic 2-photon calcium imaging preparation (see Chapter 3 – "Imaging"). Additionally, it takes typically 3-4 weeks for rodents to acquire sufficient expertise in behavioural performance for Operant Conditioning tasks (Guo et al., 2014).

One alternative that we could have tried was to train the animals to expert levels of performance, and subsequently performed the hippocampus prep. The issue(s) with this is that,

a) We wanted to study the hippocampal CA1 network during the learning or acquisition phase of behavioural training, as a distinct experiment from those published in literature.

b) In such a protocol, we would require two separate surgeries, viz., i) Head-bar implant, and, ii) hippocampus preparations. This, we believed would increase the technical difficulty of the overall experiment and could be more stressful for the experiment animals.

c) We suspected that there could be unknown effects on behaivoural performance, post-surgery, complicating the analysis and insights we aimed to study.

Across all Operant Conditioning protocols attempted, we could not observe behavioural discrimination (not licking to incorrect phases) within 1-2 weeks of training. We considered the experiments as failures, in line with this reasoning.

5. Are there any alternative methods with better temporal resolution to study the time cells' representation under various behavioral contexts? Please discuss.

New section added: "Better temporal resolution requires new techniques" to Chapter 5 – "Discussion".

*** ADDED (to lines ~3698-3728):

There are many other techniques that experimenters in the field have employed to record activity. Many of these techniques do, in fact, achieve much better temporal resolution. Here are some examples:

1) Resonant Scanning based 2p calcium imaging can achieve even up to 30 Hz for 4x larger fields of view, or more frame rates for smaller fields of view (Leybaert et al., 2005; Nguyen et al., 2001; Bonin et al., 2011; Rochefort et al., 2009). At the time when we started the experiments for the thesis, Resonant scanning microscopes required a lot of additional, expensive components to be purchased. Towards this, we co-wrote a sanctioned DBT grant application (BT/PR12255/MED/122/8/2016) and began setting up the new microscope. However, we did not have this technology available for experiments before 2020.

2) High-density tetrodes can be used to perform electrical recordings at >=20 kHz, as compared to ~14.5 Hz for our galvo-scanning 2p calcium imaging experiments. This

technique typically achieves yields of ~40 cells for hippocampal recordings, and we argued that we could achieve a higher yield (>100 cells) with galvo-scanning 2p calcium imaging. The relative sparsity of the hippocampal neural code in terms of cells participating in any engram, mandates high-yield recordings to identify the full temporal sequence of CA1 activations (Foster 2017).

3) Neuropixels (Jun et al., 2017) can be used to perform electrical recordings at >=20 kHz. At the time when we started the experiments for the thesis, these sorts of electrical probes had yet to be successfully deployed in published literature.

We discuss all of these techniques while comparing electrical vs imaging based recording strategies in Chapter 1 – "Introduction". Fundamentally, given the technological constraints at the time, we had devised combined behaviour with galvo-scaning 2p calcium imaging as the principle for the experiments described in this thesis.

Prof. Arvind Kumar (Examiner)

Introduction:

This chapter is very poorly written. Most of the text feels like it is notes that a student might use to construct the introduction. This is by no means a scientifically meaningful summary of the research area in which the student is working. A number of paras are either irrelevant or not integrated with the test. Perhaps more crucially it never becomes clear why and what needs to be standardized in both behavioural experiments and imaging of brain activity. The student also has done a rather poor job in integrating his work with the literature. So I strong recommend a complete rewrite of the text in which various argues are logically linked.

Major revisions with added text and reworked section headings now incorporated.

- I think the section title 'Project and overall goals' is not needed. Its presence makes a rather awkward start to the thesis.

We had chosen to state the various projects early in the Introduction with the idea to help define scope.

- The section 'Neural systems and behavior' hardly says anything about neural systems or behavior.

We agree. Section heading changed to "A toolkit to study time cells: Thesis Objectives", along with additional text.

<u>* ADDED (to lines ~449-466):</u>**

It has been an important goal to study memory and the neural code in terms of finer temporal order, *viz.*, behavioural time scales (~ms to s). Combining

- stable, adaptable trace eye-blink conditioning behaviour, and,
- cellular resolution 2-photon calcium imaging of hippocampal CA1, *in vivo*,
- with the goal to study time cell physiology,

was the core objective of the toolkit and the thesis as a whole.

Development in circuit manipulation tools using light-mediated activation or suppression of neuronal excitability (Luo et al., 2018), afford experimenters the ability to control circuits at ms time scales. Concomitant progress in effective physiological models of network activity during bouts of recall of the learnt behavioural trace require standardized behaviour and recording. For us, this mandated the design of a relatively low-cost, end-to-end configurable, combined behavioural and recording technology, to reliably study the neural code at the ms time scale, *in vivo*.

- Individual paragraphs are not linked or ordered. It really feels like notes one starts with. For example, lines 186-200 talk about engrams, then comes brief mention of cellular signaling and from line 211 we again return to tracing of engrams. And para from 219 talks about neural excitability without much links to the previous arguments.

Added linking paragraph(s) for this section. Please kindly refer to lines ~449-466 in Chapter 1 – "Introduction" of the thesis, or the reply to the previous comment.

- Similarly, while TEC is central to the thesis work, barely 7 lines are used to introduce this and that too completely out of blue.

Some more insights from the TEC literature have been added.
*** ADDED (to lines ~592-655):
We developed and standardized a multi-session, adaptable Trace Eye-Blink Conditioning (TEC) paradigm in which head-fixed mice learn to form associations between neutral and high-valence stimuli. We describe associative learning in later sections as well in Chapter 2 – "Behaviour". TEC has been previously observed to elicit CA1 activity sequences even in a single session of training (Modi et al., 2014). The functional involvement of the hippocampus in the acquisition of Conditioned Responses (CRs) has been studied and implicated by studying acquisition rates to multiple trace intervals. It was found that memory load, inferred in terms of task difficulty with longer trace intervals (300 ms vs 500 ms), was crucial to observing an effect of hippocampal lesions on the behavioural expression of CRs (Moyer et al., 1990).

Several stimulus modalities have been used as the Unconditioned Stimulus (US) such as periorbital air-puffs and electrical shocks (see Weiss and Disterhoft, 2016 for review). Throughout all our TEC experiments, we chose to use the mildly aversive air-puff to elicit

the Unconditioned Response (UR) to the US. Further, we use a flash of a blue LED as the Conditioned Stimulus (CS), expecting to observe reliable Conditioned Responses (CRs) to the CS within 3-7 days, for trace intervals of ~250 ms (Seigel et al., 2015). CRs are observed as a preemptive eye-blink response elicited reliably before the presentation of the US – a reproducible attempt to avoid the discomfort of the aversive air-puff. This is in accordance with the Rescorla-Wagner model of Classical Conditioning, which assumes that association of the CS and US based on repeated pairing depends on how well the presence of the CS predicts the future occurrence of the US, along with other variables such as the relative intensities and modalities of the presented stimuli (Rescorla and Wagner, 1972). Extensions to this model have suggested that there could be negative effects to the associative learning when other CS (CS1, CS2, etc.) are also paired together in within-compound-association tasks such as "backward blocking" (Hamme and Wasserman, 1994). For our experiments we used only one CS for any training session, typically a 50 ms flash of a blue LED. However, our behavioural setup allows for multiple CS types, *e.g.*, CS1 = LED flash and CS2 = auditory tone, to be presented based on the experiment.

Transient increases in CA1 excitability post acquisition of the task were described up to 4-5 days (Moyer et al., 1996) and could be important to the forging of the task specific spatiotemporal sequences during learning. In an *in vitro* assay, coronal sections of the hippocampus (Figure 1) were stimulated at the Perforant Path to the cells of the Dentate Gyrus in patterns that could be mnemonically mapped to stereotypic, temporal sequences of Excitatory Postsynaptic Potentials (EPSPs) read out at the hilar mossy cell layer ~400-500 ms later (Hyde and Strowbridge, 2013). This suggested the presence of temporal sequences even at the Dentate Granule cell layer, many synapses before the hippocampal CA1.

On a longer timescale, hippocampal lesion based experiments on mice have been used to describe the role of the hippocampus to within 4 weeks of TEC, with deficits in Conditioned Responses (CRs) as a readout of the effect of the lesion (Takehara et al., 2002). We aimed to examine the processes that underlie this time-dependent role of the hippocampus by chronically tracking the same cohort of hippocampal CA1 cells across the sessions of TEC, at cellular resolution, using galvo-scanning 2-photon calcium imaging. We were specifically interested in studying the emergence and long-term activity dynamics of time cells, touted to be the behavioural time scale (~$10-10^3$ ms) expression of the

memory engram during associative learning, as described in later sections. Preliminary results and additional details on our TEC paradigm may be found in Chapter 2 – "Behaviour".

- Theories of hippocampal function: Again these feel like notes. The four theories are described in a way too superficial manner. There is no proper definition of this idea (episodic memory is not defined, contextual mapping is not defined) or what is the current status of that idea (e.g. do we still believe that Hippocampus is needed for response inhibition). Are these the only theories? What about new ideas which implicate hippocampus in reinforcement learning, the Tolman-Eichenbaum Machine or the idea that hippocampus is needed for imagining the future. And several others. So this section really needs a good amount of rewrite or essentially the student should convert these notes into a meaningful text.

Major revisions and additional background information now added to this whole section ("Theories on the function of the hippocampus", lines ~659-746).

- Lines 365-371: do not seem to relate to the section or to the next para.

Major revisions and additional background information now added to this whole section ("Theories on the function of the hippocampus", lines ~659-746).

Specific to each theory/function, additional description has been added.

<u>* ADDED (to lines ~671-674)</u>:**
Studies have now implicated the hippocampus in the facilitation of correct responses and inhibition of incorrect responses during contextual memory tasks, though not for visual discrimination of contexts (Kim & Lee, 2012).

<u>* ADDED (to lines ~676-680)</u>:**
a form of declarative or explicit memory that refers to the ability and mechanistic paradigms that allow for the behavioural recall of a collection of past personal experiences, occurring at particular places and times to the subject. The term was coined by Endel Tulving in 1972 (see Clayton et al., 2007)

is the ability of experimental animals to locate and ascribe valence to points in space, during navigation of environments.

Pattern separation and conjunctive representation of the combined mutli-modal experience in the hippocampus, has been implicated in reinforcement learning (Ballard et al., 2019). Contextual mapping considers the hippocampal-entorhinal as a Tolman-Eichenbaum machine (TEM) with the medial entorhinal cortex (MEC) cells thought to describe important aspects of past experience and the hippocampal cells implicated in binding the current sensory experience with prior experience (Whittington et al., 2020, Eichenbaum 2007), with the goal to develop a functional model of the subjective experience of animals and flexibly selecting appropriate responses.

Episodic memory forms the central function of study for all of the hippocampus related experiments described in the thesis. Trace Eye-Blink Conditioning (TEC) is an example of a task used to study episodic recall (Thompson 2004). The behavioural acquisition and expression of CRs correlate well with the neuronal expression of spatiotemporal sequences of time cells (Modi et al., 2014). Our experiments aimed to describe finer details such as how the animals assign valence to the neutral Conditioned Stimulus (CS) and how time cell population codes adapt to changes in the trace interval. Some characteristic features of the engram at behavioural time scales, *viz.*, time cell activation sequences, have been described as preliminary results (Chapter 3 – "Imaging"), under the behavioural context of TEC.

- Line 376-378: By this definition every subcortical region ends up being in a privileged position!

Appropriate references and rationale added.

The hippocampal circuit is anatomically >3-4 synapses away from the peripheral nervous system, and information typically arrives after many layers of intervening processing and computation. It is thus suggested that the hippocampus holds a privileged position in the brain, receiving the outcomes of the computation of the brain's various modules, and

relating to them (Baudry and Lynch, 1981; Poppenk et al., 2013, Moscovitch et al., 2016; Ekstrom and Ranganath, 2017; Tao et al., 2021).

Line 547: I think it is fair to mention dimensionality reduction methods here. Bayesian decoding and information theory are not just used for multiple single unit data but also for single neuron-multi-trial data.

Additional material added to text under heading "Dimensionality reduction in the analysis of physiology".

*** <u>ADDED (to lines ~1080-1103)</u>:

Bayesian and Information theoretic approaches as well as methods like Principal Component Analysis (PCA) afford the experimenter a variety of mathematical procedures to examine dimensionality reduction, viz., the transformation of high-dimensional neurophysiological activity from recorded cells into low-dimensional representations that still retain meaningful properties of the original data (Maaten et al., 2009; Pudil and Novovicova, 1998). Claude Shannon is often credited with development of Information Theory (Shannon 1948) yet the field has evolved to also describing how relevance is assigned to a signal, based on statistical associations between multiple signals or stimuli in a sensory experience (Bialek et al., 1996; Tishby et al., 1999; Bialek et al., 2001; Chigirev and Bialek, 2004).

In the first *bioRxiv* preprint of our paper – "Synthetic Data Resource and Benchmarks for Time Cell Analysis and Detection Algorithms" (https://www.biorxiv.org/content/10.1101/2022.01.01.474717v1?versioned=true), we had compared several dimensionality reduction techniques such as Principal Component Analysis, Temporal Information, trial-averaging, etc.). In the final version of the paper (Ananthamurthy and Bhalla, 2023; attached as Chapter 4 - "Analysis"), we reoriented focus on the most popular of the algorithms, and provide very well performing Python/C++ implementations, especially of the Ridge-to-background ratio (Modi et al., 2014) and Temporal Information (Mau et al., 2018) calculations.

Line 577: This is a very limited perspective on the use of Info theory. I understand that the thesis is about hippocampus but the discussion in this section is more general. So it is important to recall the work of Bialek, Rieke etc.

Additional material on Information Theory added to text under heading "Dimensionality reduction and synthetic benchmarks for analysis of physiology".

* ADDED (to lines ~1080-1103):**

Bayesian and Information theoretic approaches as well as methods like Principal Component Analysis (PCA) afford the experimenter a variety of mathematical procedures to examine dimensionality reduction, viz., the transformation of high-dimensional neurophysiological activity from recorded cells into low-dimensional representations that still retain meaningful properties of the original data (Maaten et al., 2009; Pudil and Novovicova, 1998). Claude Shannon is often credited with development of Information Theory (Shannon 1948) yet the field has evolved to also describing how relevance is assigned to a signal, based on statistical associations between multiple signals or stimuli in a sensory experience (Bialek et al., 1996; Tishby et al., 1999; Bialek et al., 2001; Chigirev and Bialek, 2004).

In the first _bioRxiv_ preprint of our paper – "Synthetic Data Resource and Benchmarks for Time Cell Analysis and Detection Algorithms" (https://www.biorxiv.org/content/10.1101/2022.01.01.474717v1?versioned=true), we had compared several dimensionality reduction techniques such as Principal Component Analysis, Temporal Information, trial-averaging, etc.). In the final version of the paper (Ananthamurthy and Bhalla, 2023; attached as Chapter 4 - "Analysis"), we reoriented focus on the most popular of the algorithms, and provide very well performing Python/C++ implementations, especially of the Ridge-to-background ratio (Modi et al., 2014) and Temporal Information (Mau et al., 2018) calculations.

- Lines 591: I do not think the mention of synthetic data makes sense here. It is related to the thesis topic so it should come when you introduce your specific research problems.

Justification for the use of synthetic data added to text under new, reworked, heading "Synthetic benchmarks for pre-hoc development of analytical procedures"

* ADDED (to lines ~1121-1130):**

We resorted to the use of such surrogate synthetic datasets acting as a large cohort of user-configurable test datasets to benchmark and characterize the predictive performance of a variety of time cell detection algorithms (Ananthamurthy and Bhalla, 2023; attached as

Chapter 4 – "Analysis"). Analysis on real physiology datasets where categorical labels must be assessed is expected to benefit from this comparative analysis. Our Python/C++ implementations were rigorously tested and developed to the extent of excellent predictive performance. These algorithmic procedures may be used to study the nuances of time cell sequences with more statistical confidence.

- The whole text about correlation analysis should be put under a different heading. It is a different topic or at least does not fit the way the student has set up this section on multi-unit-single trial and single-unit-multi-trial.

Section moved under heading "Synthetic benchmarks for pre-hoc development of analytical procedures" at lines ~1104-1181. Section now describes the value of benchmarks along with typical issues with untested algorithms, citing synfire chains.

- And I am wondering why the synfire chains show up here? What is the relevance?

Justification added.
*** ADDED (to lines ~1173-1181):

The use of properly developed and benchmarked analytical procedures, tested and verified on physiology-equivalent test datasets (synthetic), is expected to help alleviate potential doubts in published physiology results (real recordings). For proper testing of our time cell detection algorithms, we incorporated many important user-controllable parameters for the generation of synthetic datasets, such as (but not limited to) Noise, Event Widths, Hit Trial Ratio, Trial-pair Imprecision, and Background Activity.

- The section at line 655 is supposed to be a comparison of the eletrophys methods and 2-photon calcium imaging. But I don't think the section does this job. There is no mention of the time resolution and other issues associated with Ca imagine and those of spike sorting with extracellular electrophysiology. But in the wake of the next section anything between lines 655-696 is redundant.

Additional details added.
*** ADDED (to lines ~1203-1214):

Electrical recordings may be sampled even >20 kHz, while imaging based techniques typically cannot be used to record wide fields of view, sampled at >100 Hz (1-2 orders of magnitude, comparing sampling rates). Many spike sorting algorithms (see Buccino et al., 2022 for review) as well as automated ROI detection (see Robbins et al., 2021 for review) have been suggested for automated cell source identification, yet challenges remain, such as,
1) scalability - more sources to identify from ever larger datasets force longer analysis runs.
2) reproducibility – pre-processing analytical steps require manual curation to clean up the raw datasets, and this often affects the final result.

- And then there is associate learning out of the blue? Ideally it should have been at a place where TEC was first introduced.

<u>*** ADDED (to lines ~756-764):</u>
Associative learning is the overall process by which animals develop behavioural valence to neutral stimuli that occur in temporal conjunction to other potent, behaviour eliciting stimuli. We wished to study the network level responses in the hippocampus, *in vivo*, especially during early learning of associated stimuli, *i.e.,* behavioural acquisition, combining stable, adaptable associative learning paradigms such as Trace Eye-Blink Conditioning (TEC) with cellular resolution, behaviour time scale, high-yield recordings of neurophysiology.

- Line 872: This para is out of context. You should mention why it is relevant to study CA1 in associate learning - what is the current state of literature. This should lay the foundation for the next chapter. Associate learning has been studied for such a long time -- I find it hard to believe that the experimental protocols have not been standardised yet. Please elaborate on that.

Moved "A brief introduction to associative learning" section to before "Space and time in the hippocampus" (to lines ~748-830).

Added link between previously published associative learning paradigms the specific requirements for standardization, in brief.

Further details added to Chapter 5 – "Discussion", under section "Standardizing combined behaviour and recoding experiments".

ADDED (to lines ~ 3619-3651):

We have previously stated that studying the hippocampal CA1 time cells fire in reliable sequences under the TEC paradigm (Modi et al., 2014), and wished to develop the paradigm to more fully study time cells, especially during the early or acquisition phase of training (sessions 1-7). It was not considered trivial to bundle behaviour and recording in a non-interfering way. For instance, we needed to study time cells longitudinally or chronically, and this is likely achieved by ensuring that the experimental animals were not overtly stressed, but rather were reasonably compliant to the experiment in terms of motivation. Towards this,

1) We focussed on performing only one surgery, viz., head-bar implant and hippocampus to minimize surgery-induced trauma, rather than multiple surgery strategies.

2) We incorporated a treadmill for the animals to run on during the experiments, at the potential cost of observing z-axis drift in the imaging.

3) Imaging requires that the sample (experimental animals) be illuminated only by the excitation laser and that the sensor systems for the emitted photons receive only the photons from the excited sample. We considered and designed the filter sets before our photomultiplier tube (PMT) in the emission path, to reject all IR and partially red frequencies, not just to protect the sensor from the excitation 2p laser, but also the red/short IR illumination on the animal's eye for the beaviour camera.

Through our experiments, we were able to provide some evidence that somatosensory stimuli but not other neutral stimuli could trigger CA1 responses but the effect of behavioural training results in the development of CA1 responses to the CS, now triggering a whole spatiotemporal sequence of activation. Altogether, we were able to observe preliminary results regarding the tuning, de-tuning and re-tuning of time cells to temporal fields during learning, as described in Chapter 3 – "Imaging".

- Line 897: It is worth reminding the reader of these lacunae and then mention how you fill these gaps.

*** ADDED (to lines ~1370-1374):

There are not many published physiological readouts from large CA1 populations, *in vivo*, under behavioural contexts for time cell activation, during the early, acquisition phases

(Sessions 1-7, etc.) in associative learning. It is unclear what the spatiotemporal pattern of activity would reveal with a systematic, longitudinal survey.

- Line 935: Briefly remind the reader why these measures are needed to better describe TEC

*** ADDED (to lines ~1411-1413):
The degree of variability in time cell responses under these conditions is likely to inform mechanistic models of spatiotemporal sequences as observed in the hippocampus.

- Be consistent with the fonts.
We will keep this in mind, as an important checkpoint before the final submission.

Chapter 2

In this chapter the student has described various experimental methods. It is quite nice to see that experiments that have not worked are also included. But there is hardly any discussion of why the experiments may have failed and is it consistent with vast experimental literature that already exists. This chapter needs a major revision.

Major revisions with added text and reworked section headings now incorporated.

Specific comments:

Line 1477: Why is this of interest? Unless you specify that it is not easy to understand why you are designing a certain task.

*** ADDED (to lines ~2038-2042):
Multi-modal stimulus integration is typical of the sensory experience. We felt it interesting to study how animals learnt to associate each stimulus modality, individually, given no clear *a priori* reason to assume that the physiological response would be identical for each.

Line 1481: The definition of Assoc. Learning should have come much earlier in the I section starting at 819.
*** ADDED (to lines ~756-764):

Associative learning is the overall process by which animals develop behavioural valence to neutral stimuli that occur in temporal conjunction to other potent, behaviour eliciting stimuli. We wished to study the network level responses in the hippocampus, *in vivo*, especially during early learning of associated stimuli, *i.e.,* behavioural acquisition, combining stable, adaptable associative learning paradigms such as Trace Eye-Blink Conditioning (TEC) with cellular resolution, behaviour time scale, high-yield recordings of neurophysiology.

Line 1493: This is a very strange selection to highlight the importance of anaesthetised recordings.

The para 1491-1500 seems out of context and redundant

Paragraph meant to emphasize that anaesthetized animals were not studied.
* ADDED (to lines ~2063-2065):**
We chose to avoid anaesthetized animals for the experiments described in this thesis. It is unclear if deep states of anaesthesia are also conducive to time cell sequence activation.

- Line 1529: What is the state of literature on this question? This has not been addressed in the introduction. It should have been in the section starting at line 819

* ADDED (to lines ~2095-2101)**:

We aimed to more systematically study time cells with the granularity of well-defined TEC behaviour and ms resolution CA1 activity using 2-photon imaging. We believe this is why there is still insufficient clarity on the network level reponses in the CA1, *in vivo*, especially during early learning of associated stimuli, *i.e.,* behavioural acquisition. We outline our set of core objectives required as features for the behavioural task:

- Line 1552: I suppose a section title is missing here?

This was a typographical error.

* EDIT (to line ~2122):**
In the following sections, ...

- Line 1665: When was the aversive stimulus given? Immediately after the wrong lick or only in the reward phase?

We only used aversive stimulus based punishment in some protocols, not others. We have attempted to now make this clear in the text.

*** ADDED (to line ~2229):
No aversive stimuli presented in this particular protocol.

- Line 1680: It is not too small a cohort to claim a failure?

*** ADDED (to lines ~2252-2255):
Each animal was presented with 600 training trials/day (1 session/day). The inability to behaviourally discriminate or withhold incorrect or unrewarded licks on for 7-14 sessions, was considered and the task was ultimately unsuccessful.

- Protocol 3: DNMS and Go/NoGo are classical tasks and numerous animals have been trained for these tasks. So I do not understand why the protocol has failed? What is the difference between your version of these protocols and the ones that have been reported in previous papers.

New section added: "Operant conditioning experiments failed to match requirements" to Chapter 2 – "Behaviour".

*** ADDED (to lines ~2430-2460):
Operant conditioning tasks have been extensively and successfully modeled in a variety of laboratories. For our specific experiments, we required a task that could be learnt within 1-2 weeks. This was because we were not confident on how many simultaneous days of chronic imaging, we could achieve with the *in vivo* chronic 2-photon calcium imaging preparation (see Chapter 3 – "Imaging"). Additionally, it takes typically 3-4 weeks for rodents to acquire sufficient expertise in behavioural performance for Operant Conditioning tasks (Guo et al., 2014).

One alternative that we could have tried was to train the animals to expert levels of performance, and subsequently performed the hippocampus prep. The issue(s) with this is that,

a) We wanted to study the hippocampal CA1 network during the learning or acquisition phase of behavioural training, as a distinct experiment from those published in literature.

b) In such a protocol, we would require two separate surgeries, viz., i) Head-bar implant, and, ii) hippocampus preparations. This, we believed would increase the technical difficulty of the overall experiment and could be more stressful for the experiment animals.

c) We suspected that there could be unknown effects on behaivoural performance, post-surgery, complicating the analysis and insights we aimed to study.

Across all Operant Conditioning protocols attempted, we could not observe behavioural discrimination (not licking to incorrect phases) within 1-2 weeks of training. We considered the experiments as failures, in line with this reasoning.

- Project-I is presented without much reference to literature. These tasks are frequently used. In what way the protocols you have used are different from those mentioned in the literature? And why your version of the protocols failed deserves some discussion.

New section added: "Operant conditioning experiments failed to match requirements" to Chapter 2 – "Behaviour".

<u>* ADDED (to lines ~2430-2460):</u>**

Operant conditioning tasks have been extensively and successfully modeled in a variety of laboratories. For our specific experiments, we required a task that could be learnt within 1-2 weeks. This was because we were not confident on how many simultaneous days of chronic imaging, we could achieve with the *in vivo* chronic 2-photon calcium imaging preparation (see Chapter 3 – "Imaging"). Additionally, it takes typically 3-4 weeks for rodents to acquire sufficient expertise in behavioural performance for Operant Conditioning tasks (Guo et al., 2014).

One alternative that we could have tried was to train the animals to expert levels of performance, and subsequently performed the hippocampus prep. The issue(s) with this is that,

a) We wanted to study the hippocampal CA1 network during the learning or acquisition phase of behavioural training, as a distinct experiment from those published in literature.

b) In such a protocol, we would require two separate surgeries, viz., i) Head-bar implant, and, ii) hippocampus preparations. This, we believed would increase the technical difficulty of the overall experiment and could be more stressful for the experiment animals.
c) We suspected that there could be unknown effects on behaivoural performance, post-surgery, complicating the analysis and insights we aimed to study.

Across all Operant Conditioning protocols attempted, we could not observe behavioural discrimination (not licking to incorrect phases) within 1-2 weeks of training. We considered the experiments as failures, in line with this reasoning.

Project II

- Please include reference to previous studies on TEBC. There are several statements in the lines 1856-1879 that need citations. But in general for the different ISIs, weak and strong learners there must be previous data on how quickly animals acquire the task. Please compare that with your data.

Some more insights from the TEC literature have been added to Chapter 1 – "Introduction.

<u>* ADDED (to lines ~2469-2458):</u>**
We have introduced the Trace Eye-Blink conditioning paradigm in Chapter 1 – "Introduction", but some key definitions and results require mention.

<u>* ADDED (to lines ~592-655):</u>**
We developed and standardized a multi-session, adaptable Trace Eye-Blink Conditioning (TEC) paradigm in which head-fixed mice learn to form associations between neutral and high-valence stimuli. We describe associative learning in later sections as well in Chapter 2 – "Behaviour". TEC has been previously observed to elicit CA1 activity sequences even in a single session of training (Modi et al., 2014). The functional involvement of the hippocampus in the acquisition of Conditioned Responses (CRs) has been studied and implicated by studying acquisition rates to multiple trace intervals. It was found that memory load, inferred in terms of task difficulty with longer trace intervals (300 ms vs 500 ms), was crucial to observing an effect of hippocampal lesions on the behavioural expression of CRs (Moyer et al., 1990).

Several stimulus modalities have been used as the Unconditioned Stimulus (US) such as periorbital air-puffs and electrical shocks (see Weiss and Disterhoft, 2016 for review). Throughout all our TEC experiments, we chose to use the mildly aversive air-puff to elicit the Unconditioned Response (UR) to the US. Further, we use a flash of a blue LED as the Conditioned Stimulus (CS), expecting to observe reliable Conditioned Responses (CRs) to the CS within 3-7 days, for trace intervals of ~250 ms (Seigel et al., 2015). CRs are observed as a preemptive eye-blink response elicited reliably before the presentation of the US – a reproducible attempt to avoid the discomfort of the aversive air-puff. This is in accordance with the Rescorla-Wagner model of Classical Conditioning, which assumes that association of the CS and US based on repeated pairing depends on how well the presence of the CS predicts the future occurrence of the US, along with other variables such as the relative intensities and modalities of the presented stimuli (Rescorla and Wagner, 1972). Extensions to this model have suggested that there could be negative effects to the associative learning when other CS (CS1, CS2, etc.) are also paired together in within-compound-association tasks such as "backward blocking" (Hamme and Wasserman, 1994). For our experiments we used only one CS for any training session, typically a 50 ms flash of a blue LED. However, our behavioural setup allows for multiple CS types, *e.g.*, CS1 = LED flash and CS2 = auditory tone, to be presented based on the experiment.

Transient increases in CA1 excitability post acquisition of the task were described up to 4-5 days (Moyer et al., 1996) and could be important to the forging of the task specific spatiotemporal sequences during learning. In an *in vitro* assay, coronal sections of the hippocampus (Figure 1) were stimulated at the Perforant Path to the cells of the Dentate Gyrus in patterns that could be mnemonically mapped to stereotypic, temporal sequences of Excitatory Postsynaptic Potentials (EPSPs) read out at the hilar mossy cell layer ~400-500 ms later (Hyde and Strowbridge, 2013). This suggested the presence of temporal sequences even at the Dentate Granule cell layer, many synapses before the hippocampal CA1.

On a longer timescale, hippocampal lesion based experiments on mice have been used to describe the role of the hippocampus to within 4 weeks of TEC, with deficits in Conditioned Responses (CRs) as a readout of the effect of the lesion (Takehara et al., 2002). We aimed to examine the processes that underlie this time-dependent role of the hippocampus by chronically tracking the same cohort of hippocampal CA1 cells across the

sessions of TEC, at cellular resolution, using galvo-scanning 2-photon calcium imaging. We were specifically interested in studying the emergence and long-term activity dynamics of time cells, touted to be the behavioural time scale (~10-10^3 ms) expression of the memory engram during associative learning, as described in later sections. Preliminary results and additional details on our TEC paradigm may be found in Chapter 2 – "Behaviour".

- Line 1919: This section comes very abruptly. It does not give any reason why suddenly the reader should worry about treadmill and track speed? You must explain why Trace eyeblink conditioning should be done in animals running in a treadmill.

*** ADDED (to lines ~2537-3541):
… during experiments. We considered treadmill tracking as a relevant variable to keep track of, despite the potential complications this could provide to imaging, viz., z-axis drift owing to relative motion between the brain tissue and microscope objective as the head-fixed animals run.

- Figure 13 left column: what is plotted here? I am wondering how you can get average speed distribution over a single trial? I am assuming that this is data from a single session of a single animal.

Left: Trial-by-trial running speed (cm/s) over a 50 ms window-sized bin for any frame. We move the averaging window across the trial points.

*** ADDED (to line ~2553):
... over 50 ms sized bins, ...

- Figure 21: Define what is weak and strong learner?"

We classified strong learners as animals with a performance of >60% hit trials/session. Weak learners have a performance of 30-60% hit trials/session.

*** ADDED (to line ~2645-2648):"
We additionally set a criterion that a performance of >60% be considered "strong learning", 30-60% be considered "weak learning", and "0-30%" be considered "non-learning".

We also added the same information to the Figure 18 legend:

Figure 18: ... "Here M2 is a strong learner (>70% hit trials/session) and M5 is a weak learner (30-60 % hit trials/session). M1, M3, and M4 did not learn the task."

Chapter 3

In this chapter the student has described main results about neural activity during TEC tasks. The major limitation is that most of the results are presented in an anecdotal manner. The main focus of the work is 'time cells' but they are very poorly described and defined. This chapter requires a major revision.

Specific comments:

- Parts of the somatosensory cortex were removed for imaging. Did it result in some deficits in the animals over time since recording continued for 2-3 weeks?

<u>* ADDED (to lines ~2894-2896):</u>**

We kept track of animals that showcase unusual gait or low/no mobility and avoid their use altogether for production datasets and experiments, in accordance with previously published protocols (Dombeck et al., 2010).

- Very little information is provided about the figures 29AB have been constructed. Why was clustering done? What were the clustering algorithm parameters? At least from the figure, there is not much indication of clustering. Later the student is comparing correlation structure in spontaneous and evoked activity. But again too little information is provided how the correlations were calculated in the evoked activity which barely lasts for a few frames. So what duration of activity was taken around the stimulus presentation to calculate the evoked correlation matrix. Since there are way too many cells responding to the stimulus, how was this corrected for. Overall this is an interesting result as it is, it looks pretty much anecdotal -- I think all that has been shown is data from one animal.

We wanted to essentially look at the spatial organization of cells with high correlation in activity, since imaging as a recording method provides an unabiguous anatomical map of each of the recorded cells.

We set the initial seed to k = 3, with 1000 bootstrap iterations (see Modi et al., 2014 for sensitivity analysis), distance as specified by pair-wise correlation coefficient (Figure 29A; Figure 31A; Figure 32A), and a threshold of 80% to bundle meta-groups.

Correlations were calculated for all frames post the whisker-puff (stimulus), which corresponds to ~8 s of network activity. It is across this period that the clustering was performed.

The correlation analysis was performed to check for consistency across previous recordings in the lab (Modi et al., 2014; Ashesh Dhawale's Thesis). The idea here is to be able to provide proof-of-principle results to confirm if the technique was working. Additionally, although the results are from 1 animal, the session consists of 60 trials for each of the >100 cells recorded. This very preliminary result has been studied and described in detail, previously (Modi et al., 2014).

- Line 2322: For imaging, parts of the somatosensory cortex were removed. So ideally such stimuli should not be tried. But I am wondering if there was any temporal structure to how the response to such stimuli evolved over time (across days) as the cortex recovered from the load of tissue.

The somatosensory cortex damage inflicted was only on one hemisphere. We presented somatosensory cortex to the ipsilateral whiskers, information that is expected to be processed on the contralateral hemisphere. In any case, we observed responses at the CA1 timed to the presentation of the whisker-puff, unlike trials with tone. We describe tuning modulations across days for the chronically tracked time cells in the last section of Chapter 3 – "Imaging". It is very difficult to isolate and therefore comment on the effect of tissue recovery, since we did not directly measure this variable.

- Line 2520: Definition of time cell is not clear. Frame rate is 10/15 frames per sec. The time between CS and US is 300 or 500ms. That means that there are only 2-4 frames between CS and US. Just because cells are locked to CS or US they do not become time cells.

Figure 33: what does time = 0 correspond to?

Here, time at 0 ms is the presentation of the CS.

*ADDED (to Figure 33 legend):**

The conditioned stimulus (CS; blue LED) is presented at time 0 ms.
"

- Line 2532: Please specify the analysis and checkpoints.

95

Please see section "Preliminary analysis to identify time cells" – lines ~2764-2780.

* ADDED (to lines ~3210-3214):**

We define time cells as cells with a higher probability of eliciting activation (tuning fields) to specific temporal landmarks across trials, rather than uniformly over the whole trial. For the results described in this chapter (Chapter 3 – "Imaging"), our temporal information calculation was used as a functional definition for time cells.

- Line 2544: This line is just reporting the figure 36 -- no discussion of what is actually in the figure and what it means

* ADDED (to lines ~3265-3277):**

It has been observed in time cell literature as well as in our real physiology recordings that time cells with tuning to later time points in a trial, tend to have wider tuning curves (B. J. Kraus et al., 2015; MacDonald et al., 2011, 2013; Mau et al., 2018; Pastalkova et al., 2008). This is why no obvious trend in a graph of Temporal Information vs time of peak sorted time cells, seemed curious enough to mention in the thesis. However, while we report the observation, we made no attempts to delineate further detail, **given a limited number of observations**. This legacy adaptation of the temporal information calculation had not been benchmarked and thus the result was not certain, at the time. Subsequently,

we have observed similar results in other lab colleague's recordings, and avoided further discussion, with regard to the scope of this thesis.

- Line 2559-2561: This expansion in the time cells seems to be reported for one animal (Fig 37-38). The student should report group statistics. Also discuss how this expansion relates to training and/or learning.

The results presented are towards the extent of proving proof of principle for the technologies developed. We have cited these results as "preliminary". Other lab colleagues have taken up these experiments and their results are expected to be published outside of the scope of this thesis.

- Figure 42: There is no explanation for choosing the threshold of 0.2 to suggest stability of response. It would be good to show any example where the cell pair correlation is 0.2. In any case there is no information about how many cells from how many animals were included in this analysis.

*** ADDED (to line ~3316):
(~60% of all session-wise cells tracked; ~80 cells)

*** ADDED (to lines ~3318-3320):
Typically, we observed only at best a correlation coefficient of ~0.6-0.8 across cell pairs (please kindly refer to Figures 29 and 31). Accordingly, we decided to choose 0.2 as the threshold.

- Line 2602: Width of tuning curve is first time mentioned here. How was this defined? Please define it and show the results.

*** ADDED (to line ~3237-3241):
Peri-stimulus time histograms (PSTH) or event time histograms (ETH) were created by summing the number of threshold crossing activity events per bin (bin size = 3 frames/bin or ~200 ms/bin at 14.5 Hz) across trials. Different cells showcase different widths for ETH or tuning curves (Figure 35A; Figure 35B).

Additionally, please kindly refer to Figure 35.

- Line 2617: Why is it surprising? What is the criteria to determine if the response persistence is expected or not?

*** ADDED (to lines ~3349-3353):

Single cell measurements tend to define responses up to 200-1000 ms post US. Our own correlation based clustering analysis was across the w6hole post US phase (vs pre-stim phase), which corresponded to ~8s. It was surprising to us, to find the number of time cells even many frames after the US (>400 ms or >6 frames at 14.5 Hz).

- This chapter ends very abruptly -- without any conclusion. What is the significance of these findings about time cells?

*** ADDED (to lines ~3363-3366):

Our preliminary imaging and behaviour results describe neuronal sequence activations based on the emergence or re-tuning dynamics of temporal tuning by time cells, during early phases of behavioural acquisition, in a chronically tracked fashion.

Chapter 4

This is a peer-reviewed manuscript and there are no further edits needed for this.

Chapter 5

This chapter provides a brief summary of the results. Since this is a Phd thesis one would expect much more discussion about the methods and results. But the chapter is pretty devoid of any such discussion. I think this chapter should be expanded quite a bit.

We have added major sections to Chapter 5 – "Discussion" to address this, supported with appropriate changes to Chapters 1-3. However, we did follow administrative guidelines to write short summaries in this final section.

Specific comments:

- Line 2819: I am not really sure what has been standardised and in what specific way. This is one of the main claims of the thesis and this should be made very clear. First in the

introduction - what needs to the standardized and the in the discussion what has been standardised.

Please kindly refer to Chapter 1 – "Introduction" (lines ~170-187) or the section entitled "Required features" in Chapter 2 – "Behaviour" (lines ~1736-1742).
In brief, we have standardized TEC to be coupled with 2p Imaging, in concert, with the goal to study time cells. It is non-trivial to combine behaviour and imaging for longitudinal studies.

- Line 2822: What are the 'several types of behavioural adaptation ...'? The first claim is well known [Line 2824]. The third claim is not tried -- it's just an expectation [line 2828].

Rewrote this section to add emphasis on main results with TEC standardization.

ADDED (to line ~3548-3564):

... (data not shown). TEC experiments with multiple CS, viz., blue LED and Tone, have been tried in the lab. They form the primary behavioural modulation being investigated by a lab colleague. These experiments are now directly possible due to the standardization efforts.
4. In our experiments where we extended the Trace Duration, animals show retention of previously learnt CR times (Figures 19-21), showcase complex blinks (Figure 19) without change to CR onset (Figure 20).
5. We could also train animals on very long Trace durations (550 ms, and 750 ms), which have previously not been reported for head-fixed mice.
6. Across all the single interval training experiments, the animals only produce one conditioned response, with time of peak adjusted relative to the timing of the US.
7. Complex blinks were only observed in animals trained to more than one trace interval.

- Line 2839: Again here it is important to clearly mention what has been standardized for 2-photon imaging -- both in the intro and in the discussion

Added a new section titled "Standardizing combined behaviour and recording experiments" (to lines ~3206-3232) to help bring out the standardization.

*** ADDED (to lines ~3619-3651):

We have previously stated that studying the hippocampal CA1 time cells fire in reliable sequences under the TEC paradigm (Modi et al., 2014), and wished to develop the paradigm to more fully study time cells, especially during the early or acquisition phase of training (sessions 1-7). It was not considered trivial to bundle behaviour and recording in a non-interfering way. For instance, we needed to study time cells longitudinally or chronically, and this is likely achieved by ensuring that the experimental animals were not overtly stressed, but rather were reasonably compliant to the experiment in terms of motivation. Towards this,

1) We focused on performing only one surgery, viz., head-bar implant and hippocampus to minimize surgery-induced trauma, rather than multiple surgery strategies.

2) We incorporated a treadmill for the animals to run on during the experiments, at the potential cost of observing z-axis drift in the imaging.

3) Imaging requires that the sample (experimental animals) be illuminated only by the excitation laser and that the sensor systems for the emitted photons receive only the photons from the excited sample. We considered and designed the filter sets before our photomultiplier tube (PMT) in the emission path, to reject all IR and partially red frequencies, not just to protect the sensor from the excitation 2-photon laser, but also the red/short IR illumination on the animal's eye for the behaviour camera.

Through our experiments, we were able to provide some evidence that somatosensory stimuli but not other neutral stimuli could trigger CA1 responses but the effect of behavioural training results in the development of CA1 responses to the CS, now triggering a whole spatiotemporal sequence of activation. Altogether, we were able to observe preliminary results regarding the tuning, de-tuning and re-tuning of time cells to temporal fields during learning, as described in Chapter 3 – "Imaging".

- Line 2924: This section comes very abruptly. I am not sure how it relates to the thesis work. The first two paragraphs the author has tried to build some narrative but it does not serve the purpose.

*** ADDED (to lines ~3730-3734):
An important directive to neuroscience research is to understand the brain and nervous system, in how these structures allow animals to interact meaningfully with their environment. More humbly, however, the ultimate goal of this thesis was to help provide a multi-disciplinary toolkit to study time cells in the hippocampus.

Additional sections added to Chapter 5 – "Discussion".

We use this section to emphasize that the brain continues to predict across different states and experiences, in an on-going effort. We use this narrative thread to link the essential summary of what the thesis has tried to achieve, *viz.*,

1) Standardized Behaviour – Trace Eye-Blink Conditioning (TEC) in head-fixed mice,

2) Standardized Imaging – Chronic 2p Calcium Imaging of hippocampal CA1, *in vivo*, and,

3) Benchmarked Analysis – Multiple analysis algorithms for detecting time cells in physiological datasets.