Learning and representation of declarative memories by single neurons in the human brain

Thesis by

Ueli Rutishauser

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2008

(Defended May 8th, 2008)

Acknowledgments

The research presented in this thesis would not have been possible without the help, advice, and encouragement of a large group of people. While this thesis bears my name, this is partially their work too. I feel exceedingly fortunate in having such a great group of friends, colleagues, and mentors, all of whom contributed to this work.

First, I would like to thank my advisors Erin Schuman, Christof Koch, and Adam Mamelak. Their continued support, guidance, and intellectual freedom allowed me to pursue my interests. The work presented in this thesis was primarily done under the mentorship of Erin Schuman and Adam Mamelak. While initially intimidating, I'm grateful that I was given the opportunity to build, from scratch, a human single-unit recording operation. This was a large endowment of trust, given that I had little experience in neuroscience, let alone electrophysiology or medicine. It was a formative experience, which could not have succeeded without the continued guidance and support of Erin Schuman. The members of her laboratory created a nurturing environment, despite me working on a topic far removed from the Lab's mainstream. Alana Rathbun provided much needed assistance with numerous administrative tasks.

While the work I produced under the mentorship of Christof Koch is not part of this thesis, it had an unquestionable impact on my current thinking as reflected in this thesis. I admire his deep and intense love of science and the search for truth. I would like to thank Christof for all his support, encouragement, enthusiasm, and the nurturing environment that "K-lab" provided for me.

I also received advice from the other members of my committee: John O'Doherty and Gilles Laurent. In addition, I had the pleasure of interacting with Ralph Adolphs, John Allman, Henry Lester, Thanos Siapas and Pietro Perona. My first home at Caltech was Pietro Perona's lab, of which I have fond memories. I'm grateful for the freedom he gave me to explore my interests, which eventually led to this work.

Much of the experimental work I performed took place at Huntington Memorial Hospital (HMH). It was a pleasure to interact with people at HMH, in particular the EBMP (Epilepsy&Brain Mapping) nurses, staff, and physicians. Their help is greatly appreciated—many of these experiments would not have been possible without the continued support. I appreciate help with the Neuropsychological evaluations by Linda Philpott. Most of all, I would like to thank the many patients (whom I not allowed to list by name) for their participation. I'm very grateful for their willingness to participate during a time of great distress and uncertainty in their lives.

Over the years, I have greatly benefited from discussions about my research with Costas Anastassiou, Moran Cerf, Tobi Delbrueck, Daniela Dieterich, Rodney Douglas, Wolfgang Einhaeuser, Michael Fink, Alan Hampton, Jonathan Harel, Alexander Huth, Constanze Hofstoetter, Alex Holub, Hiroshi Ito, Vivek Jayaraman, Andreas Kotowicz, Sally Kim, Gabriel Kreiman, Sotiris Masmanidis, Kevan Martin, Florian Mormann, Dirk Neumann, Zoltan Nadasdy, Kerstin Preuschoff, Johannes Schwarz, Bryan Smith, Michael Sutton, Chin-Yin Tai, and Nao Tsuchiya (in alphabetical order).

I had the opportunity to mentor and work with Caltech undergraduates, SURF students and master students Arjun Basal, Matt McKinley, Anthony Chong, Davhyd Wing and Andreas Kotowicz. It was a pleasure to work with them.

I would not be at Caltech today without the help and encouragement of my previous advisor, Rodney Douglas. I'm grateful for the opportunity he gave me by exposing me to the most interesting research problems. I would also like to thank the faculty members at my undergraduate institution HSR that have supported me in this endeavor—Josef Joller, Gabor Hardy and Andreas Mueller. Thanks also to Gaston Wolf of ZHW for advice in transfering to graduate school.

My way into graduate school was not as direct as it perhaps could have been—I spent several years in the software industry. Several people from this time have been influential in my personal development and I would not be where I'm today without them. Rene Berchten was my first mentor and taught me many of the fundamental software skills I still use today. Enzo Giannini always supported my efforts and was very helpful in the transition between business and academia. I would also like to thank my business partner Alain Schaefer of many years—It was a pleasure running easc together. The results in this thesis are the product of an estimated 60'000 lines of code (mostly Matlab). The skills I was fortunate to acquire have thus been tremendously useful.

Big thanks go to my friends, both here at Caltech and in Switzerland. I'm grateful for their friendship, support, and time spent away from the lab; much of this work could not have succeeded without them: Maik Berchten, Oliver Bandle, Moritz Delbrueck, Nicolas Foong, Patrick Gaus, Kasia Gora, Jonathan Harel, Alexander Huth, Prashant Joshi, Rene Lavanchy, Christof Moser, Tony Oertli, Bobby Rohrkemper, Bryan Smith, Alain Schaefer, and Aaaron Wittmann (in alphabetical order). Thanks to my fellow class-of-2008 members: Anusha Narayan, Hilary Gliden, Matthew Nelson, Dirk Neumann, Ming Gu, Casimir Wierzynski, Will Ford, and Jason Rolfe.

The California Institute of Technology and the Computation & Neural Systems (CNS) program has been an ideal home institution, providing the best possible environment to explore my diverse research interest. Much of my academic thinking has been shaped here. I'm grateful to have been part of Caltech. The group of people that make up the Caltech community is truly remarkable. I specifically would also like to thank International Student Programs, the Office of Technology Transfer, the Center for Neuromorphic Engineering, and the Library for help over the course of my stay. It was very rewarding to be at a place where intense dedication to science and technology is the norm rather than exceptional.

Last, but not least, I would like to thank my family who has continually supported me unconditionally, despite the long distance to Switzerland.

Abstract

Episodic memories allow us to remember not only that we have seen an item before but also where and when we have seen it (context). Neurons in the medial temporal lobe (MTL) are critically involved in the acquisition of such memories. Since events happen only once, the ability to distinguish novel from familiar stimuli is crucial in order to rapidly encode such events after a single exposure. Theoretically, this is a hard learning problem (single-trial learning). Yet, successful detection of novelty is necessary for many types of learning. During retrieval, we can sometimes confidently report that we have seen something (familiarity) but cannot recollect where or when it was seen. Thus episodic memories have several components which can be recalled selectively. We recorded single neurons and local field potentials in the human hippocampus, amygdala, and anterior cingulate cortex while subjects remembered, and later retrieved, the identity and location of pictures shown. We describe two classes of neurons that exhibit such single-trial learning: novelty and familiarity detectors, which show a selective increase in firing for new and old stimuli, respectively. The neurons retain memory for the stimulus for at least 24 h. During retrieval, these neurons distinguish stimuli that will be successfully recollected from stimuli that will not be recollected. Similarly, they distinguish between failed and successful recognition. Pictures which were forgotten by the patient still evoked a non-zero response. Thus, their response can be different from the decision of the patient. Also, we demonstrate that listening to these neurons (during retrieval) enables a simple decoder to outperform the patient (i.e., it forgets fewer pictures). These data support a continuous strength of memory model of MTL function: the stronger the neuronal response, the better the memory (as opposed to a dual-process model). I also describe specific power increases in specific frequencies of the local field potential that are predictive of later retrieval success. These neural signatures, recorded during learning, thus indicate whether plasticity was successful or not.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACC	anterior cingulate cortex
AMPA	alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic
AED	anti-epileptic drug
BOLD	blood-oxygen-level-dependent (see fMRI)
BDNF	brain-derived neurotrophic factor
Ca^{2+}	calcium
CA1-3	Cornu ammonis fields (of the hippocampus)
CT	computerized tomography
Cl	chloride
CS	conditioned stimulus
CR	conditioned response
CTA	conditioned taste aversion
CV	coefficient of variation
CMA	cingulate motor area
DA	dopamine
DG	dentate gyrus
DLPFC	dorsolateral prefrotal cortex
EC	entorhinal cortex
EEG	electroencephalography
EPSP	excitatory postsynaptic potential
EPSC	excitatory postsynaptic current
ERP	event-related potential
ERN	error-related negativity

FDR	false discovery rate
FEF	frontal eye fields
fMRI	functional magnetic reasonance imaging
Glu	glutamate (neurotransmitter)
GABA	γ-aminobutyric acid (neurotransmitter)
IPSP	inhibitory postsynaptic potential
IPSC	inhibitory postsyanptic current
IT	inferotemporal cortex (monkey)
ISI	interspike interval
kstest	Komogorov-Smirnof goodness-of-fit test
K^+	potassium
LFP	local field potential
LGN	lateral genculate nuclei
LTP	long-term potentiation
LTD	long-term depression
LIP	lateral intraparietal area
MRI	magnetic reasonance imaging
MEG	magneto-encephalographic
MTL	medial temporal lobe
MLREG	multiple linear regression
mPFC	medial prefrontal cortex
MT	middle temporal area of the cortex
MUA	multi-unit activity
MLE	maximum likelihood estimate

NMDA	N-methyl-D-aspartic acid (an amino acid)
Na^+	sodium
OCD	obsessive-compulsive disorder
OFC	orbitofrontal cortex
PCA	principal component analysis
PET	positron emission tomography
PFC	prefrontal cortex
PPF	paired-pulse facilitation
RLSC	regularized least-square classifier
RT	reaction time
ROC	receiver operator characteristic
RMS	root-mean-square
SNR	signal-to-noise ratio (but also see below)
SNr	nigra pars reticulata, substantia nigra
SVM	support vector machine
STDP	short-time dependent plasticity
SUA	single-unit activity
STD	standard deviation
SE	standard error
STN	subthalamic nucleus
TE	anterior part of inferior temporal cortex (monkey)
TEO	posterior part of inferior temporal cortex (monkey)
TLE	temporal lobe epilepsy
UR	unconditioned response

US	unconditioned stimulus
V1	primary visual cortex, Brodman area 17.
V2	part of the extrastriate visual cortex , Brodman area 17.
V3	part of the extrastriate visual cortex , Brodman area 17.
V4	part of the extrastriate visual cortex , Brodman area 17.
VTA	ventral tegmental area