

# MVDMLAB MANUAL

Copyright © 2019 van der Meer lab

WWW.VANDERMEERLAB.ORG



Licensed under the Creative Commons BY-NC-SA License (the “License”), version 4.0; you may not use this file except in compliance with the License. Briefly, you are free to share and adapt this material, under the condition that you give appropriate credit, do not use the material for commercial purposes, and distribute your contributions under the same terms. You may obtain a copy of the full License at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

Inspired by similar lab manuals by others. Particularly helpful were the wonderful examples from the [Peelle](#) and [Aly](#) labs.

Typeset in Palatino Linotype using  $\text{\LaTeX}$  and the `tufte-book` class.

*Version 0.2, August 2019*

# Contents

<i>Welcome!</i>	5
<i>About this manual</i>	7
<i>Research Using Animals</i>	9
<i>Animal care and recordkeeping</i>	11
<i>Values and Expectations</i>	13
<i>Code of Conduct</i>	13
<i>Expectations: everyone</i>	14
<i>Principal Investigator</i>	15
<i>Graduate Students</i>	16
<i>Research Technicians</i>	17
<i>Undergraduates</i>	18
<i>Visitors to the lab</i>	18
<i>Doing good science</i>	19
<i>Failing fast and often, but learn from your mistakes</i>	19
<i>Documentation and taking good notes</i>	20
<i>Asking good questions</i>	22
<i>Reading papers</i>	23
<i>Communication</i>	23
<i>Lab space</i>	25
<i>General expectations</i>	25

<i>Experiment rooms</i>	26
<i>Fine Assembly Room (FAR)</i>	27
<i>Surgery and anteroom</i>	27
<i>Workshop</i>	28
<i>Shared facilities</i>	28
<i>Vivarium space</i>	28
<i>Offices</i>	28
 <i>Electronic resources</i>	 29
<i>GitHub</i>	29
<i>Slack</i>	29
<i>Wiki</i>	30
<i>Calendar</i>	30
<i>Mailing list</i>	30
<i>Lab server</i>	30
<i>Others</i>	31
 <i>Computing</i>	 33
<i>Lab computers</i>	33
<i>Printing</i>	35
<i>Skills</i>	35
<i>Software</i>	36
 <i>Data management</i>	 39
<i>Data storage</i>	40
<i>Data promotion</i>	41
<i>Data use cases</i>	41
 <i>Bibliography</i>	 43

# Welcome!

THE VAN DER MEER LAB brings together people who share an interest in how the brain works. We aim to better understand how learning, memory, and decision-making arise from the coordinated activity of neurons. In the pursuit of this goal, we perform brain surgery, design and construct eclectic mazes, painstakingly build small devices so we can read the minds of rats, solder tiny circuit boards to even tinier wires, write thousands of lines of computer code, collect beautiful data, produce even more beautiful plots, and wonder what it all means.

EACH OF US BRINGS their own particular blend of motivations for engaging in these lab activities, but there are a few common threads: curiosity about how the physical stuff of the brain gives rise to thought and to behavior, a love for animals and computers<sup>1</sup> alike, a desire to help solve some of the big mysteries, and contributing to human knowledge and ultimately a better world.

OUR EFFORTS ARE *collaborative*, not only within the lab, but also with other labs in the department of Psychological & Brain Sciences at Dartmouth. We share equipment, space, and interests with these labs, and collaborate not only on joint scientific projects, but also on creating a scientifically exciting and supportive culture where everyone can thrive. We also have joint projects with other labs around the world. We collaborate because the problems we are trying to solve are hard, and because learning from and working with others is one of the great joys of working in science!

WORKING WITH OTHERS brings not only joy, but also expectations. Coding and data analysis can be fun, but are full of pitfalls. Sharing space and expensive equipment with others demands care and respect. There is satisfaction in getting an experiment to work, but

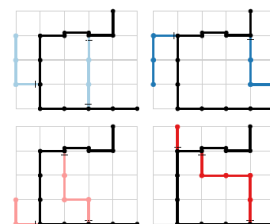


Figure 1: Example maze configurations. (From Emily Irvine's *shortcut* experiment.)

<sup>1</sup> Or, at least, a love of computable things.

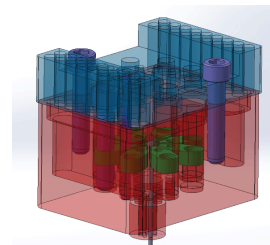


Figure 2: A small device for rat mind-reading (by Andrew Alvarenga).

it can be a long slog to get there. Perhaps most importantly of all, the use of animals in research comes with practical as well as moral responsibilities. *This manual is intended to provide guidance on how to navigate these issues, with particular focus on the van der Meer lab.*

IF YOU ARE NEW TO THE LAB, I would like you to read this manual through in its entirety. If you will not be working with animals directly, it may seem like chapter on Animal Management does not apply to you, but it will give you important insights into how the lab operates.

WELCOME. Let's do some great science together!

– MvdM

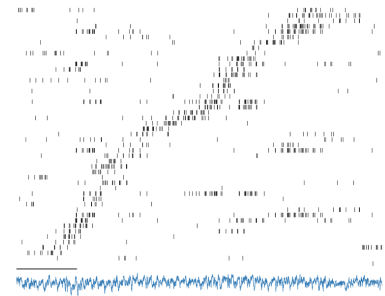


Figure 3: Some beautiful data, recorded from R050's hippocampus (by Alyssa Carey). Vertical tick marks indicate spikes (one row per neuron, sorted by place field location); horizontal axis indicates time; blue trace shows a local field potential. Scale bar is 1 s.

# About this manual

THIS DOCUMENT describes the principles that shape how the members of the lab interact and work. Rather than dealing with the nuts and bolts of how to get stuff done in the lab, the manual is about lab philosophy, expectations, and resources. It is an introduction to the lab. The lab manual is one of several classes (categories) of shared documents in the lab, which are:

- *Lab Manual*. You are looking at it. It contains information that does not change frequently. Only MvdM can change the lab manual, but I'd love to hear from you if you have thoughts of suggestions! Periodically, we may review the manual as a group at lab meeting to determine what needs updating.
- Private *GitHub repositories*<sup>2</sup> that contain protocols, i.e. documentation for experimental procedures<sup>3</sup>, and a separate repository for behavioral tasks and associated computer code<sup>4</sup>. This documentation is on GitHub so that we can say things like, "I used version 1.1 of the histology protocol", and so that we can track changes to those protocols as well as the reasons for those changes. If you perform procedures in the lab, you are expected to follow and to contribute to these protocols; see the [Documentation](#) section for more detailed information.
- A *lab wiki*<sup>5</sup> that contains tutorials, guides, lists of useful links, et cetera. The wiki is a more dynamic, easier to edit resource for content that doesn't rely as much on version control. The wiki currently contains the lab's MATLAB data analysis tutorials.
- We also have a lab *Slack* team<sup>6</sup> which is used for day to day communication. It hosts shared documents best described as "everything else", something that isn't a protocol, task, or tutorial.

These different venues reflect an overall organization ranging from content that is easy to create and easy to change (low overhead, quick; Slack, wiki), to content that is moderately stable and takes a

<sup>2</sup> [www.github.com/vandermeerlab](https://www.github.com/vandermeerlab).

Because these repositories are private, you will not be able to see them unless you are logged in and a member of the vandermeerlab organization.

<sup>3</sup> Examples include electrode plating, histology, drive building, et cetera: see the [mvdmlab-protocols](#) repository

<sup>4</sup> [mvdmlab-tasks](#)

<sup>5</sup> [rcweb.dartmouth.edu/~mvdM/wiki](https://rcweb.dartmouth.edu/~mvdM/wiki)

<sup>6</sup> [mvdmlab.slack.com](https://mvdmlab.slack.com). If you are new to the lab and you haven't received an invitation email to join it, e-mail MvdM.

bit more effort to change (but easier to track; GitHub), to principles that rarely change (lab manual). For instance, an idea for an experiment may be initially discussed on Slack, lead to a draft protocol shared there, and then get pushed to GitHub.

There are a number of other kinds of documents used in the lab that, unlike the above categories, are not typically consulted by multiple lab members. Most important among these are (1) animal records, described in more detail in the Animal Recordkeeping chapter, and (2) lab notebooks/notes, described in the [Keeping Good Notes](#) section.

**Important:** I assume the lab manual and procedures on GitHub are accurate. This means that you should follow all of the policies and procedures contained in the manual and GitHub. If you notice something that seems to be wrong, please let me know (for the lab manual) or change it yourself (if on GitHub). If there is something in the lab manual or GitHub that you notice people aren't doing, please bring this up at lab meeting, or to me directly. Don't assume this is okay (it's not).



# Research Using Animals

WE USE ANIMALS in the lab, based on the conviction that doing so accelerates scientific progress and provides a net benefit to humans, and perhaps to animals, too. There are many examples of scientific discoveries that were directly enabled by research in animals<sup>7</sup>. However, past successes do not mean we can experiment freely on any animal we want. The use of animals for research carries a moral, scientific, and legal obligation to only perform experiments when the benefits can be reasonably expected to outweigh the costs, to do so only if no suitable alternatives are available, and to do so while caring for our animals to the fullest extent possible. These notions are often phrased as the “3 R’s” (reduction, replacement, refinement)<sup>8</sup>.

More generally, research in animals demands that we do the best job possible, and maximize the usefulness of the research outcomes.

This means that:

- *We take care of our animals.* If animals are comfortable and well cared for, they will perform the behavioral tasks we want, be easier to handle, and the resulting findings will be more likely to generalize.
- We strive to *design experiments such that the outcomes will be as informative as possible*. There are many resources on this topic, including the specific demands on rigor and reproducibility mandated by the NIH ([link](#)). Many experimental design decisions can be brought into focus by preparing a formal [preregistration](#) of your proposed experiment, which requires justifying how many subjects you plan to run, what analyses you will carry out, and the inferences supported by possible outcomes.
- We aim to collect the *highest quality data possible*. We often invest a tremendous amount of time in each animal through behavioral training, construction of chronic implants, surgery, and so on. If

<sup>7</sup> Some good examples are highlighted in the “Position statement of the Max Planck Society concerning the use of animals in experiments for basic research” ([link](#))

<sup>8</sup> Good references on this topic include [NC3Rs](#).



Figure 4: A visual allegory for synergy between humans and rats.

any one of these procedures is not performed diligently, then your investment in all the others may come to nothing.

- For the data to be interpretable, *good records* need to be kept. Recordkeeping includes informal written notes and the completion of highly organized templates produced as procedures are performed. Good recordkeeping also encompasses the process of translating these notes into computer-readable formats suitable for data analysis.
- *Data is valuable*. It needs to be protected against inadvertent loss, and its impact needs to be maximized by making it possible for a single data set to inform multiple questions, including those asked by different investigators within and outside the lab. This means it needs to be organized and annotated in a specific common format.

In sum, the fact we are working with animals shapes both our lab philosophy and many practical aspects of working in the lab. Much of the content that follows in this manual is essentially unpacking the above points in more detail.

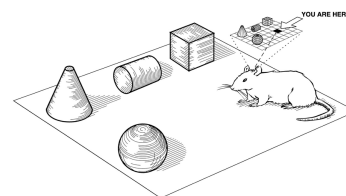


Figure 5: A cartoon rat with some objects (from Eichenbaum et al. 1999).

ANIMAL RESEARCH AT DARTMOUTH is overseen by the Institutional Animal Care and Use Committee (IACUC). All work with animals needs to receive prior approval from the IACUC in the form of *Animal Use Protocols* (AUPs)<sup>9</sup>. Only *procedures*<sup>10</sup> that have been approved by the IACUC may be performed on animals, and only by approved personnel (listed on the protocol). We are **required** to log all procedures; details on this process can be found in the Animal recordkeeping chapter. Invasive procedures, such as electrode implant surgery, require follow-up procedures providing post-operative care (administration of analgesia, monitoring).

To be formally approved to perform any of the procedures described in an AUP, a number of on-line and in-person training modules need to be completed. To initiate the training process, e-mail MvdM, and please forward copies of any confirmation of training completion. The current version of all AUPs can be found in the [#animalcareandrecords](#) channel on the [lab Slack](#).

<sup>9</sup> *Animal Use Protocol*: document outlining the scientific rationale, aims, and a set of proposed experiments, including specific animal numbers for each. The total number of approved animals may not be exceeded.

<sup>10</sup> *Procedure*: Any action or change beyond picking an animal up in the colony room (e.g. for changing cages or weighing) counts as a procedure. This includes, among others, placing an animal on a restricted diet, running a behavioral task, and injecting an analgesic.

**Important:** Violations of the above requirements, such as performing a procedure not previously approved, or neglecting to log a procedure that was performed, are serious. Beyond potentially violating our moral responsibility towards our animal subjects, the lab, the department, and Dartmouth risks losing federal approval to perform animal research if found delinquent. Systems are in place to minimize the likelihood of mistakes from happening; however, we are all human and mistakes do happen. If you notice any slips, please take action: rectify if you can, consider how to prevent this happening again, and let MvdM know. Transparency and honesty about what occurred is always the best course of action.

## *Animal care and recordkeeping*

KEEPING DETAILED AND ACCURATE RECORDS on each of our animals is an IACUC requirement (and indirectly a federal issue), as well as scientifically important. The main principles are that for any animal that is currently alive, there must be a binder (or a section in a binder) with that animal that contains an up to date record of all procedures<sup>11</sup> performed on that animal, as well as records of their weight.

ANIMAL HUSBANDRY<sup>12</sup> AT DARTMOUTH is provided by the Center for Comparative Medicine Research (CCMR). CCMR maintains a satellite animal facility in the basement of Moore Hall, where our animals are housed and cared for by CCMR staff. CCMR also employs veterinary staff, who are “on call” to provide support and advice (and in unusual cases, can mandate an animal be euthanized). Up-to-date contact information for CCMR staff and vets can be found in the [#animalcareandrecords](#) channel. When performing invasive procedures (surgery) and when administering drugs, CCMR requests that these are reported using a special cage card<sup>13</sup>. Writing procedures on these cards is *not* sufficient logging: you also need to log the procedure(s) in the animal’s binder.

By default, CCMR staff will feed, water, and change cages for all our animals. We pay them a *per diem* fee to do this. Unless requested otherwise, all animals receive *ad lib*<sup>14</sup> food and water.

We can also request CCMR to feed rats 18g/day. This is a useful amount that prevents rats from getting obese. To request this, use the relevant cage card<sup>15</sup>. For any other food/water regimens, you’ll need to use the “Experimenter will feed/water” cage card. Doing so

<sup>11</sup> What counts as a procedure: anything beyond what’s required to change a cage or weigh an animal. If in doubt, log it anyway.

<sup>12</sup> That is, animal care not consisting of procedures, such as changing cages and providing food and water.

<sup>13</sup> Picture of procedure cage card.

<sup>14</sup> *Ad libitem* is the Latin phrase for, literally, “to your libido”, i.e. as much as desired. Incidentally, *i.e.* is Latin for *id est*, “it is”.

<sup>15</sup> This one.

implies that the experimenter listed takes ownership<sup>16</sup> of that animal.

When animals first arrive, they are typically group-housed (multiple animals per cage). They are named<sup>17</sup> and start out with communal status. Communal animals do not have an owner yet. They are weighed weekly by the Communal Animal Caretaker<sup>18</sup>. When communal animals are below 400g in weight, the Caretaker also handles them weekly (about 5 minutes per animal appears to be the sweet spot, although you may want to take ownership of an animal that will be implanted with a hyperdrive early on so that it can be handled more). When communal animals reach 400g, the Caretaker separates them into individual cages so that CCMR can feed them 18g/day.

Once you euthanize an animal (or request it to be euthanized), collect all weight and procedure logs, along with any additional information (notes from behavior, for instance), scan them into a single PDF<sup>19</sup> and upload to the data vault (see the section on Promoting Data).

If you are working with animals, read the [Animal Recordkeeping Protocol](#), which provides detailed procedures for the above.

<sup>16</sup> As the owner of an animal, you are responsible for logging its weight and all procedures. If the animal is food and/or water-restricted, you are responsible for supplying those things, and logging that you have done so.

<sup>17</sup> Rats are named Rxxx, where xxx is a 3-digit counter that is incremented with each new animal. Mice are named Mxxx.

<sup>18</sup> Note that to track weights of group-housed animals, some kind of mark needs to be used. The lab convention is to write numbers 1, 2, 3 etc. on the base of the tail with a Sharpie. 1 indicates the animal with the lowest number. Mouse tails are too small to write numbers; stripes or rings can be used instead.

<sup>19</sup> The departmental copier/scanner is great for this, you can feed it a pile of documents and it can email you a PDF.

# *Values and Expectations*

THE LAB VALUES personal and scientific integrity, respect for each other, for animals, for equipment, and for data. We care about shaping and maintaining an environment that fosters personal development, curiosity, exploration, and the joy of discovery. We choose to work on hard problems that require dedication, persistence, resourcefulness, resilience, and humility. We support each other in when things get tough, and celebrate each other's successes. We value opportunities to learn from each other, set high standards but do not judge.

HOLDING THESE VALUES implies the following:

## *Code of Conduct*

THE LAB IS COMMITTED to providing a safe, friendly, and accepting environment for everybody. We will not tolerate any verbal or physical harassment or discrimination on the basis of gender, gender identity and expression, sexual orientation, disability, physical appearance, body size, country of origin, native language, race, or religion. We will not tolerate intimidation, stalking, following, unwanted photography or video recording, sustained disruption of talks or other events, inappropriate physical contact, and unwelcome sexual attention. To my knowledge, none of these have ever occurred in the lab – let's keep it that way.

Nevertheless, if you notice or experience any of these, please tell MvdM. If you have an issue with a member of the lab, tell MvdM. If you have an issue with MvdM that you are not comfortable discussing, the graduate student representative, department chair, and other faculty are available, as well as the resources listed below.

Please be assured that your concerns will be taken seriously, that you will be listened to, and that you can share them without fear of retaliation. Note the distinction between *private resources* and *confidential resources*<sup>20</sup>.

Before taking pictures or videos with a lab member in them, ask their consent first. Ask consent again before posting to social media. If you'd like to take pictures or videos of an experimental subject, ask permission from MvdM first.

#### IMPORTANT RESOURCES:

- Dartmouth policies: [undegraduate](#) and [graduate](#) code of conduct
- PBS [diversity](#) page.
- Confidential resources on campus: [WISE](#)<sup>21</sup>, [counselors](#) at Dick's House, religious leaders at the [Tucker Center](#)
- Private resources: MvdM, PBS department chair, PBS graduate chair, other faculty you trust
- Dartmouth [Title IX resources](#)

*Expectations: everyone*

#### BIG PICTURE:

- *Community.* Behave respectfully and professionally towards others in the lab, department, and beyond. Doing this well requires considering how your actions (or lack of them) affect others<sup>22</sup>. We value behaviors that enable everyone to flourish.
- *Collaborate.* Share your expertise. Make arrangements with others to take care of their animals one weekend and reciprocate another. Offer to do code reviews, read each other's manuscript, offer feedback on practive talks.
- *Contribute* to community resources and space. If you notice something that can be improved, fix it, mention it on Slack, or open a GitHub issue.
- *Communicate.* Tell others what you are working on, what is working well and what could be improved. Post updates on Slack frequently – sharing when you are stuck and what you have tried to get unstuck is just as important as sharing successes.

<sup>20</sup> Private resources are legally obligated to share a disclosure of sexual misconduct with the Title IX office, including all details that are known. This is to ensure adequate support for those who may need it. Confidential resources may not share any information (with very few exceptions).

<sup>21</sup> A WISE advocate is on campus one day a week; see the website for details.

<sup>22</sup> See the [Recurse guidelines](#) for some great examples.

- Take care of yourself. Know when it's time to push, and when it's time to take a walk in Pine Park<sup>23</sup>.
- Tell others when you notice they did something well, or when they made your life a little better today<sup>24</sup>.
- Be a little obsessive, and have high standards.
- Show up and keep at it. While brilliant ideas, creativity and skill certainly help, the best predictor of success is how consistently you put in the work. There will be days when it feels like you aren't making any progress and nothing is working. Don't underestimate the power of an iterative routine

<sup>23</sup> See [here](#) for a guide to some of Hanover's best trails, accessible from the lab in the space of a lunch break.

<sup>24</sup> We have the octothorp, awarded during lab meeting. But don't stop there!

#### SMALL PICTURE:

- Document. Anything you do more than once, or that you or someone else might have to do again, should be written up as a experimental protocol<sup>25</sup>. Anything done with an animal is a procedure and needs to be logged. Good recordkeeping is essential to doing good science. Keep track of what you did with experiments and analyses in your lab notebook.
- Treat your animal subjects with care, as laid out in the [Research with Animals](#) section.
- Participate in our weekly lab meetings.
- Be punctual and consistent if you are running experiments. Hours can be somewhat flexible, but in general all full-time lab members are expected to be in at work during at least the core hours of 11-4, so that we can interact. If you are going to be away from the lab for more than a day or two, enter it into the lab calendar.
- If you purchase anything with a chart string<sup>26</sup>, take a picture of the receipt and post it in the #admin channel on Slack.

<sup>25</sup> Not to be confused with an animal (IACUC) protocol.

<sup>26</sup> A Dartmouth-internal number that allows expenses to be charged to the correct account and category.

### *Principal Investigator*

#### I PROMISE TO:

- Provide an environment where great science can be done and people can thrive.
- Model the different aspects of the scientific process and different components of an academic career.

- Give my perspective on where the field is going, and how work in the lab will help shape that direction.
- Secure operating funds for the lab, which may include your stipend or salary.
- Be a mentor, supporter, and resource for you. This includes providing you with training in diverse aspects of science, helping you set priorities in the short and long term, promote your work and helping you identify opportunities.
- Set high standards and commit the time to help you meet them. In particular, I will hold weekly lab meetings and project one-on-ones with you. I will also give you timely feedback on project ideas, conference posters, talks, manuscripts, figures, grants.
- Support and monitor your trajectory in yearly feedback meetings.

### *Graduate Students*

- By the time you graduate, you are expected to make an original contribution to knowledge, in the form of a thesis that you will defend.
- In the process of reaching this goal, you will build a portfolio of published work as well as a set of skills. Become really good at at least one marketable technique or skill.
- When you first join the lab, you are expected to take on an existing or previously proposed project, so you can learn the ropes with plenty of interaction time with MvdM and senior lab members. As you progress, you are expected to take a more active role in helping to shape project directions within the scope of the lab's current and planned funding.
- Your productivity is expected to be low initially (1-2 years), but then ramp up. Expect to submit at least one paper per year in years 3+. Research is the clear priority – avoid spending too much time on coursework and teaching assistantships.
- Become the world expert on your thesis topic. Read superficially in a broad area, and deep in your chosen area<sup>27</sup>.
- Submit fellowship applications. Help MvdM write grants and review papers.
- Be aware of the program requirements, and everything else that

<sup>27</sup> See “How to read a book” book and MvdM’s slides for some pointers on different kinds of reading. Keep abreast of current developments by signing up for key journal table of contents, and use crowdsourcing for recommendations (e.g. Twitter).



is in the PBS grad guide. Be in charge of these program requirements. Coursework and milestones.

- Attend B4, and present regularly (this is a program requirement). Attend lab journal clubs, and take initiative in proposing a journal club topic or paper from time to time.
- Attend departmental colloquia, and go to lunch with the speaker a few times each year.
- Participate in Dartmouth research events (Neuroscience Day), regional meetings, and major conferences in the field such as the Society for Neuroscience annual meeting<sup>28</sup>, CoSyNe, Biological Psychiatry, or others.
- Participate in rotating lab assignments that change on a term-by-term basis (animal care, space, codebase, protocols, journal club chair, social chair).
- Attend a summer school such as Neural Systems & Behavior at MBL<sup>29</sup>, the Okinawa Computational Neuroscience Course<sup>30</sup>, and others.
- Be aware of what is required to take the next step of career paths you are considering<sup>31</sup>. If you are thinking about postdocs, identify potential advisors early on and take steps to get them to know about you and your work. Read some example letters of recommendation, and proactively consider what you would like yours to say.
- If you are going to be away from the lab for more than a day or two, let MvdM know and enter the dates in the [lab calendar](#).
- Take an active role in monitoring and promoting your mental health<sup>32</sup>.
- Don't forget that this is a special time in your life where you get to immerse yourself in a topic you are passionate about.

<sup>28</sup> Note that you need to be a student member to attend; see [www.sfn.org](http://www.sfn.org) for info.

<sup>29</sup> [www.mbl.edu/nsb](http://www.mbl.edu/nsb); Eric and Alyssa have TAd at it, MvdM and Youki were students, and MvdM has taught there for several years.

<sup>30</sup> Eric was a student, and MvdM has TAd at it.

<sup>31</sup> A great starting points for thinking about industry vs. academia is [this book](#); lab alumni are also a great resource.

<sup>32</sup> Some starting points for online resources include GoGrad, Dartmouth's graduate school [Mental Health Awareness program](#) and [Health & Wellness resource list](#).

### *Research Technicians*

- Model good lab practices such as recordkeeping, labeling, and organization of lab spaces.
- Work regular hours as defined by your contract. Flexibility is encouraged, but seek approval from MvdM first.

- cc MvdM on all PCARD transactions when you send them to the departmental administrator for reconciliation. Request permission first for anything unusual and anything over \$500.
- Discuss attending talks (B4, departmental colloquia) with MvdM first before going.

### *Undergraduates*

- Participating in research is different from taking a class. Coursework typically consists of specific assignments with instructions which you can follow. Research typically centers on a question or objective, and you will need to figure out how to make progress with it. Take an active role in identifying how to measure that progress, and how to tell when you are stuck. If you are stuck, discuss possible solutions on Slack and/or with other lab members.
- Especially if you do animal work, reliability and punctuality is crucial. Be realistic about the time you can commit when you have to balance exams and other activities, and plan ahead.
- Letters of recommendation. You can expect me to write letters of recommendation for you if you have worked in the lab for a year (three terms) or more<sup>33</sup>.
- Learn by observation: pay careful attention to how senior lab members do things, and copy them.
- Learn to code. This is one of the most useful skills you can learn – not only is it essential for doing research, but it is a highly transferable skill.
- Be proactive about identifying funding sources for research and travel, and be aware of the requirements of your funding<sup>34</sup>.

### *Visitors to the lab*

Do not bring visitors<sup>35</sup> into the Moore Hall basement without asking MvdM for permission first.

<sup>33</sup> When requesting letters of recommendation, please allow sufficient time (at least two weeks). If you are applying to multiple programs, provide a single document, such as a spreadsheet, listing the key information for each. When requesting letters, attach your personal statement, CV, and specific details you would like to be highlighted in the letter. Send a reminder email a few days before each deadline if you haven't yet received confirmation that a letter has been submitted.

<sup>34</sup> For instance, WISP requires you to present a poster at the end of your first two terms in the lab.

<sup>35</sup> Anyone who does not have their own swipe card access is considered a visitor.

## *Doing good science*

This chapter provides some general strategies that effective scientists use.

### *Failing fast and often, but learn from your mistakes*

SCIENCE INVOLVES A LOT OF TRIAL AND ERROR. You will be asking questions whose answers are often not yet known, using combinations of techniques and analyses that do not yet exist or have not been applied to your particular question. So you want to find ways to learn from errors as quickly as possible, and ways to avoid making them in the first place.

A general principle to accelerate learning is to find ways to get feedback often. For projects that are at the idea stage, talk with your peers about it, and solicit feedback from more senior colleagues. Find out what the likely pitfalls and challenging steps are, and who has already tried or even solved them before.

For projects that are experimental, the same idea applies: find a setup in which the feedback cycle is as short as it can be. For instance, if your project requires recording from a brain area that the lab has not recorded from before, chronically implanting an animal would take much longer until you find out if you hit, compared to doing an acute (non-recovery) surgery.

Experiments often consist of multiple steps or components that all need to work – failure in any component will make all the other steps in vain<sup>36</sup>. This implies that you should find ways to get feedback about whether each component is working correctly. For instance, inspect your tetrodes under a microscope before attaching them to the interface board.

<sup>36</sup> A typical example is that if you do not follow aseptic surgical procedures, an infection may cause your implant to detach. Having built perfect electrodes is then moot.

For analysis projects, a similar logic again applies: *formulate expectations about what the output of a given analysis step should look like*, and then verify. For all of these processes, apply a hacker/startup mindset that takes the quickest path towards a minimal working example or proof of principle. At that point, if the results look promising, you should clean up, improve and document things.

Often, things will not work. This is not always a problem – sometimes you just want to try something and only follow up if it looks like things are going to work. However, some failures actually block your way forward, and you will need to resolve them. Notes, as described next, will help you (and others) identify what the problem might be. Making the same mistake twice is not a good use of time and resources.

### *Documentation and taking good notes*

ONE WAY TO AVOID making the same mistake twice or to avoid making them in the first place is to have good documentation. Getting frequent feedback as described above is important, but given feedback you then need to solve the *credit assignment problem*: which of your actions do you need to change to improve things?<sup>37</sup> Lab notes are important in remembering what actions you took in the first place, so that you can better determine which to change. A related but different reason to care about documentation is that a desirable property of a scientific study is that it should be possible for others to replicate the result.

<sup>37</sup> The algorithmic and neural basis of the credit assignment problem is one of the lab's scientific interests, and core issue in the field of reinforcement learning.

### *Lab notes*

You are expected to keep notes in a lab notebook. This can be a physical notebook, or an electronic document. I keep notes in a binder for experiments, and in a Google Doc for data analysis<sup>38</sup>. Make a habit of writing notes every day that you perform experiments or analysis.

Some examples of things to keep notes about:

- Building a drive for implanting. Document the lengths of the pieces used, the height and weight of the drive, which tetrode wire you used, et cetera.
- Improvising a new way to glue an optical fiber to a silicon probe.

<sup>38</sup> Be aware that images pasted into any of the Google online documents are stored with lossy compression, so you will not be able to recover the original image!

- The filename of a new piece of code you used to try out a new analysis (that you're not ready to commit to GitHub<sup>39</sup> yet), some screenshots of the plots it produced, your interpretation of what you see, and the next step you plan to take the next day.

<sup>39</sup> See section XXX for an introduction to Git and GitHub.

Of course, any procedures performed with experimental subjects need to be logged in that subject's binder. Commonly performed procedures, such as histology and surgery, will have a protocol with fields that can be completed as you go along (discussed below). If you perform mostly well-established procedures, your notes binder will consist primarily of protocol sheets. Having your notes in a binder is great for this because you can mix protocol sheets and written notes.

### *Protocols*

PROTOCOLS ARE STEP-BY-STEP INSTRUCTIONS of procedures used in the lab. Example protocols include:

- Building a hyperdrive (electrode array) for chronic implantation.
- Staining brain sections for localizing electrodes.
- Daily protocol for preparing the "MotivationalT" task and running a rat on it.

ANYTHING YOU EXPECT TO DO AGAIN in the future is a candidate to be written up as a protocol. Writing really good protocols takes a substantial amount of work. Getting into the habit of writing up something (rather than nothing) and incrementally improving them is a great way of getting there. A good goal to aim for is that when you are ready to write up your methods in a paper or thesis for publication, you can refer to the protocols you have previously developed.

PROTOCOLS ARE HOSTED on the [mvdmlab-protocols repository](#) on GitHub<sup>40</sup>. This is done first, so that all protocols are in one place that is easy for everyone to find. Also, version control for protocols is helpful because we can track changes over time and the reasons for those changes, as well as refer to specific versions in our papers<sup>41</sup>. It also provides a place for comments and discussion related to specific protocols, and collaborative editing.

<sup>40</sup> See section XXX for a GitHub intro, and note that you need to have been added as a member of the vandermeerlab organization to see this private repository.

<sup>41</sup> e.g. "We used version XXX of the hyperdrive construction protocol," which is helpful for reproducibility.

WHAT MAKES A GOOD PROTOCOL? Use standard structure (purpose, ingredients, equipment). Step-by-step instructions: it should be a recipe, and algorithm. The main procedures should be clean and minimal so that it is easy to get an overview, but adding copious notes and explanations as footnotes/sidenotes/appendices are great. Steps should include checks/tests that can be used to tell if things are working as expected. Pictures can be very helpful here.

If you use a protocol, contribute to it regularly, even if it is only to say when it was last used successfully (and by whom).

## *Asking good questions*

*Scientific questions, hypotheses, and predictions.*

A MAJOR GOAL OF SCIENTIFIC INQUIRY is to iteratively develop working models of how the world works. Models of the brain can be used to (1) generate predictions about what an individual will do in the future based on brain activity now, (2) fix malfunctioning brains, and (3) build things like robots or computer programs that have abilities similar to brains. Working models of the brain, or some component of it like a specific circuit or neuron, are informed by experimental data. In general, we don't know what the best model is<sup>42</sup>, and so we are motivated to perform experiments to improve our models.

IN HYPOTHESIS-DRIVEN RESEARCH, the hypothesis is the model, and it generates predictions about what we expect to happen under various conditions<sup>43</sup>. Experiments can be performed to test those predictions. Your goal should not be to obtain a certain answer, but rather to obtain information that helps you decide among alternative models, refine existing models. You should formulate expectations about possible answers and consider if those would in fact inform the question you are asking; think about your experimental design, statistical power, and control experiments.

*Questions about computer code.*

IN PRACTICE, a major way you'll be asking scientific questions is not only by doing experiments, but also by writing computer code to analyze data. When troubleshooting computer code, aim to create a Minimal, Complete, and Verifiable example. In doing so,

<sup>42</sup> In fact, it seems clear that our current models of the brain aren't very good at all; another way of saying the same thing is that our understanding of how the brain works is far from complete.

<sup>43</sup> Note the distinction between a hypothesis and a prediction. Something like "I expect the treatment group to have impaired memory performance compared to the control group" is not a genuine hypothesis; it is a prediction that follows from some underlying understanding or model (the hypothesis). Try to make your hypothesis explicit. In this example, that could be something like, "Memory formation is dependent on the NMDA receptor."

you make it easier for others to replicate the problem you are experiencing, helping them help you. Moreover, often you'll find that in the process of creating such an example, you discover the source of the problem. See [this page](#) on Stack Overflow<sup>44</sup> for more detailed instructions.

<sup>44</sup> [www.stackoverflow.com](http://www.stackoverflow.com), an excellent resource for crowdsourced answers to programming issues.

### *Reading papers*

To be written. Inspectional, analytical, syntopical reading; bibliography management software; importance of finding the classic papers and intellectual roots of the question you are working on. Important resources.

### *Communication*

#### *Writing well*

To be written. Clear thesis; importance of making the structure of your argument clear; results packet, vomit draft, iterating often. Work to find the simplest way of making your point. Reading aloud helps identify problematic phrases. Cite the first, best, and most recent (this is a good rule of thumb to guide your reading, too.)

#### *Importance of visuals*

To be written. Simpler is better (Tufte design philosophy). Labels typically need to be bigger than you think. Use a consistent font (Helvetica is good).

#### *Giving talks*

To be written. Use “theory of mind” to simulate how what you say will appear to your audience.





## *Lab space*

Our lab space consists of the following:

Include map.

### *General expectations*

In the lab, we assemble precision devices that will be surgically implanted in live animals. We perform behavioral experiments that are sensitive to changes in many different conditions. The amount of time invested in these procedures, and the fact that we are doing this in live animals, demand careful attention. In addition, lab space is subject to general lab safety requirements ([link](#)) and IACUC inspections ([link](#)).

The principle that guides the use of shared lab space is that any lab member needs to be able to come in, find the items they need, and get things done to a high standard. In addition, the condition in which the lab spaces are kept should reflect pride taken in doing good work – it will be seen and interpreted by others viewing our space. Depending on the specific rooms this can mean somewhat different things, discussed in more detail below, but there are some common principles:

- *Clean up after yourself.* Use common sense about when to do this. If you're in the middle of something and stepping out to get some lunch, it makes no sense to clear your work area. However, if you've finished what you were working on, definitely clean up. Before going home for the day is usually a good time to at least clean a bit. See the specific room schedules for more specific cleaning requirements.
- *Organize where things are stored.* Any item in the lab has a home. Some individual items are sufficiently important that they have a dedicated place, such as "final cut scissors" (PICTURE). Some

items don't have a dedicated place but are a member of a more general category, such as "BNC interconnects". If you find an item that does not have an obvious home, check on Slack if anyone has suggestions, and/or create one for it. Creating a home for an item can be as simple as labeling a piece of a shelf somewhere. Make sure you tell everyone about it on Slack, in accordance with the Communication Principle!

- *Track stock levels.* If something is running low, don't just ignore it. Post about it on Slack. Store parts that go with a certain piece equipment with that equipment (for instance, by taping a ziploc baggie to it).
- *Label things.* Any liquid and food containers **MUST** be labeled with the contents and expiry date, if applicable. When you open a new package of something, write on it when you opened it. When a new thing comes in, write on it when it was received.
- *Treat tools with respect.* What may look like a cheap pair of scissors is likely to cost at least \$300. Use tools for their intended purpose, and use the correct tool for the job. For instance, don't use fine scissors to make rough cuts into hard materials. Don't use fine forceps to hold metal parts. <sup>45</sup>
- *Be safe.* The general lab safety training covers the basics, but specific spaces have hazards described below. First Aid kits are available in the FAR and in B101.

<sup>45</sup> Quick guide to forceps: "#55" and "#5 Biologie" are the finest.

## Experiment rooms

### AKA "running rooms"

These are the main spaces that our rats interact with. That has a few consequences:

- If you have a rat out, have the doors closed. Put a sign on the door saying, "Experiment in Progress".
- Don't make changes to the layout of the room while you are running an experiment. Changing the cues and landmarks available can change what strategy animals use to solve a behavioral task.
- Work to keep light levels, sound levels, and odor cues consistent.
- Any surface rats interact with needs to be cleanable. So, no cardboard, unsealed wood, and so on.
- Our IACUC protocols require cleaning apparatus at least weekly.

Cleanliness and organization is important, but don't let that hold you back in shaping your workspace to enable you do perform your experiments well and efficiently. Tape procedures you are using to the walls, print out the relevant atlas sections, make temporary storage spaces for the tools you use frequently.

### *Fine Assembly Room (FAR)*

The general directives about cleanliness and organization above all apply. In addition, the following are important:

- If you are working on something, it's okay to leave your work in progress on work surfaces; however, please be considerate in how much space you take up. Using a tray to hold you work is a great way to limit your footprint, and to make it easy for others to use the space if they need it.
- Be particularly on guard for fragile items such as drives and silicon probes.
- Close glass doors after you open them, and definitely when you leave for the day. They are there to protect sensitive equipment from dust and accidental touch.

### *Surgery and anteroom*

SURGERIES FROM WHICH ANIMALS RECOVER demand that these spaces are kept impeccably clean to minimize the risk of infection. In addition, during surgery sometimes time is of the essence, and finding what you are looking for quickly can make the difference between success and failure. Thus, it is especially important that B99 and B100 are organized well.

In particular:

- Clean surfaces before and after surgery with Clidox (disinfectant) and then 70% alcohol (solvent).
- Treat the stereotax, pump, anesthesia trolley, and microscope with utmost care – these are expensive, precision instruments.
- Be on the lookout for low supply levels. For instance, after finishing a surgery, post to Slack if the oxygen cylinder has  $\frac{1}{4}$  or less left (there should always be a spare available in surgery).

### *Workshop*

Moore B101 contains several important tools:

- Belt sander. Wear gloves and goggles when using it.
- Rotary tool. Wear gloves and goggles when using it.
- 3D printer.

### *Shared facilities*

Histology.

### *Vivarium space*

Our rats are in Moore B63, and our mice in Moore B61.

Vivarium rooms need to have binders containing weights and animal records.

### *Offices*

These are located on the ground floor of Moore, in the South wing (graduate students) or in the West wing (lab techs and postdocs). The grad offices house the lab library, consisting of books from MvdM's personal library. Please ask MvdM before taking any book out of Moore, and be vigilant about other labs borrowing but not returning them.

## *Electronic resources*

### *GitHub*

GitHub hosts multiple repositories containing important lab resources. These include:

- `mvdm-lab-protocols`: contains descriptions of experimental procedures.
- `mvdm-lab-tasks`: contains descriptions of behavioral tasks.
- `mvdm-lab-designs`: contains CAD designs for use with our 3D printer or printing services.
- The main lab codebase. MATLAB version and Python version
- Code and data associated with the lab's publications.
- Assorted project-specific repositories.

To access some of these repositories, you need to be a member of the [vandermeerlab organization](#). Ask MvdM to become a member. Although you can view files on the repositories through the GitHub website, you'll need to learn how to use `git` in order to use code stored there, and to contribute.

### *Slack*

[www.slack.com](https://www.slack.com) is a software platform for having group conversations in topic-specific channels. The lab Slack is [mvdm-lab.slack.com](https://mvdm-lab.slack.com). If you are joining the lab but haven't received an invite yet, please let MvdM know.

The lab Slack is used for everyday communication about all aspects of the lab's work, ranging from animal-related logistics and ordering supplies to logging and discussing ongoing project results and



Figure 6: Slack.

sharing grants. I suggest you join all channels initially – even channels on topics not necessarily related to your interests you should look at on occasion to see what is happening.

### *Wiki*

The lab [wiki](#) contains data analysis tutorials and other resources.

### *Calendar*

The lab has a Google account (username `mvdmlab`, password `ozomat11`) which is associated with a number of calendars. All experiment rooms and surgery have their own calendar, which you should use to schedule your planned use of these rooms. The main lab calendar contains labmeeting times, other scheduled events (e.g. journal clubs), and times different lab members will be away.

### *Mailing list*

The lab mailing list is used for announcements such as lab meetings and journal clubs. If you are joining the lab, MvdM will sign you up.

### *Lab server*

The lab server, physically located in MvdM's office, is the local storage point for all the lab's data. Information on how to access it can be found in the `#server` channel on Slack. It contains a number of folders, most importantly:

- **Promoted:** data folders that have been preprocessed<sup>46</sup> and can be used for data analysis.
- **InProcess:** data that is in the process of being promoted.
- **Incoming:** raw data. After acquiring data in an experiment, make sure to upload the data to the server; see the Computing section for details.
- **Tutorials:** example data sessions used for data analysis tutorials.
- **Dropoff:** a temporary folder that can be used for sharing files across the lab.

<sup>46</sup> See section XXX for an explanation of what data preprocessing entails.

*Others*

We use Dropbox<sup>47</sup> to facilitate sharing of files across computers. It should never be used for storage of data! If you see an opportunity to promote any files currently on Dropbox to the relevant GitHub repo, please do so (or create an issue first).

<sup>47</sup> Credentials are [mvdmlab@gmail.com](mailto:mvdmlab@gmail.com)/ozomat11.





# Computing

COMPUTING PLAYS AN INCREASINGLY IMPORTANT ROLE in (neuro)science and is central to many aspects of the lab's workflow. The safety of your hard-won data, the integrity and reproducibility of your analyses and results, the ongoing development and sharing of the best protocols and procedures, and the speed with which you can accomplish your goals all depend critically on correct use of the lab's computing resources.

Because many of these are shared (experimental machines) and/or inherently collaborative (lab database, codebase) it is especially important to be aware of the issues below. Experience with some of the more advanced concepts and tools is a highly valued skill in many labs and workplaces; mastery of these will set you apart from many of your peers.

## *Lab computers*

The lab's computers are a mixture of custom builds (assembled from parts by MvdM or someone else brave enough!), designed for a particular role, and general-purpose workstations purchased from Dartmouth Computing Services. The different roles determine the best way to use each machine. The roles are:

- Single-user machines in offices, one for each grad student/tech/postdoc.
- Shared machines. There is one in the FAR (B28), for use with the NanoZ and high-power microscope; one in B101, for use with the 3D printer; a laptop in surgery; and one computer associated with each data acquisition system (Neuralynx, Open Ephys).

### *Single user machines*

One of these will be exclusively yours to use during your tenure in the lab. You are free to install software and change settings to suit your needs as long as any changes (1) do not interfere with the machine's ability to support research, and (2) fall within the Dartmouth Computing guidelines<sup>48</sup>. However, remember that this is still a lab (Dartmouth)-owned machine, and should not be extensively used for private activity. You can feel free to choose your own username and password credentials.

<sup>48</sup> link

Single user machines have a boot drive (C:) for the operating system and frequently used software. This is a fast SSD drive with limited space, so be mindful of what you install here. Absolutely do not use this drive for any documents or data; use the data drive (D:) for this.

Many single-user machines use a RAID (automatically “mirrored”) storage array for the D: drive. This means that if one hard drive fails, the data is still available on the other. However, mirroring will not help you if you accidentally delete or overwrite files. For this reason, *data* should always be available on the lab server, *code* should be regularly pushed to GitHub and/or stored on Dropbox<sup>49</sup>, and *documents* (e.g. your grant application, paper, or thesis) should be on Dropbox or equivalent.

<sup>49</sup> ...or a similar service such as Google Drive; note that Dropbox has a free academic account upgrade.

If you see Windows updates to install on any machine, go ahead and install them. If a computer needs rebooting because of updates and you can do so safely, go ahead.

### *Shared machines*

Unlike the single user machines, when using the shared machine you will need to consider how your actions affect other users. A few guidelines:

- If need to save your work, create a folder with your name on the Data (D:) drive. Do not leave work or data files on the desktop. If you find anything that's not supposed to be there on the desktop or anywhere else, put it in the lost-and-found folder on the desktop.
- Do not install any software, or change any software and system settings, without discussing with MvdM first. Do not delete any files, but place them in the lost-and-found folder on the Desktop instead.

- If you log in to web services, consider logging out when you are done.

### *Experimental machines*

The above considerations for shared machines apply. In addition:

- If you acquire data for your project, use correct renaming and backup procedure when you finish your acquisition session. This is absolutely essential: data is expensive, and you individually and the lab as a whole cannot afford to lose any of it, EVER. Do not leave data in the acquisition folder.
- If you encounter a potential data file or folder that is not yours and seems lost, make an effort to find out whose data it might be.
- NEVER delete anything that looks like it might be data, unless you are absolutely certain that it has been backed up correctly.

### *Printing*

The departmental printer (in XX) is the primary way to print stuff. Please be considerate when using the printer. Keeping an electronic library of PDFs using an iPad, Dropbox and iAnnotate is a worthwhile investment that will make it much easier to find papers and your comments in the future, and will save a LOT of paper!

We also have a lab printer in the FAR. Only use this printer for printing experiment-related documents such as animal weight sheets and protocols.

### *Skills*

To be written.

Basic skills: learning about your filesystem, what makes sense to store where. Learn to use the command line.

Learn to code. MATLAB and Python are the most important. In our subfield of neuroscience, MATLAB is still the most used, but Python is gaining ground.

Good coding practices; style guides.

How to ask good questions. Stack Overflow on how to create a [minimal, complete, and verifiable](#) example.

## Software

In general, you should feel free use whatever software is the best tool for the job<sup>50</sup>, and you should feel free to do so. However, there are startup costs associated with learning how to use new software, and the lab has built up expertise and resources in a number of software packages that work well. Learning to use these software packages will help you be productive in the lab, and will be applicable in other settings as well.

- **Git** is a tool for “collaborative version control”: a way to keep track of changes made to files such as computer code, experimental protocols or manuscript drafts, and to coordinate those changes across multiple contributors.
- **MATLAB** is a scientific computing environment that includes a programming language, libraries for performing specific tasks, a development environment with a fully featured editor, debugger, and interpreter, and many other features. It is widely used for data analysis in neuroscience. Although it is not free, Dartmouth has a site license that enables you to use it ([download page](#)). Python is becoming increasingly popular, and is far more widely used than MATLAB outside neuroscience. However, various toolboxes used in the lab such as Chronux and FieldTrip are still MATLAB-only, and MvdM is still a beginner in Python.
- **Adobe Creative Suite** is expensive, but produces superior documents. We use Illustrator to touch up and group together figure panels for manuscripts based on raw output from MATLAB or Graphviz, and to make posters. Photoshop is sometimes used to process image files such as those from histology. Free alternatives such as Inkscape and GIMP exist, can do the job for simple documents, but I don’t (yet) recommend them for publication-quality output.
- **Graphviz** produces schematics such as the one in Figure 7. Unlike ‘what-you-see-is-what-you-get’ design software such as Photoshop and Illustrator, the Graphviz tools create graphics from a plain text source file. The source file contains a recipe for what the graphic output should look like.

<sup>50</sup> ...where “best” is defined as the intersection between the capabilities of the software itself, suitability for the task at hand, and the user’s ability to use it correctly.

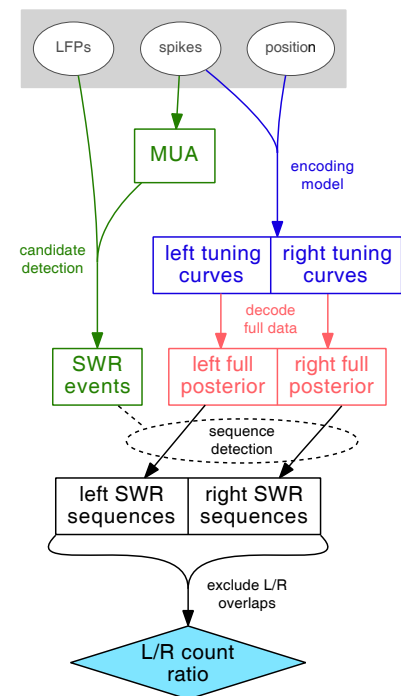


Figure 7: Example data analysis workflow, made with [Graphviz](#)

- $\text{\LaTeX}$ . Unlike “what-you-see-is-what-you-get” content creation programs such as Word,  $\text{\LaTeX}$  creates documents from a plain text source file. The source file contains instructions for what the document should look like. Compared to Word,  $\text{\LaTeX}$  does much better with equations, produces more professional-looking documents<sup>51</sup>, and makes it much easier to maintain large documents.

Note that several of these require you to use the Command Line (Terminal).

<sup>51</sup> Especially where typesetting features such as spacing, kerning, and ligatures are concerned.

### *Operating Systems*

Lab computers use Windows and/or Ubuntu Linux. There are a few reasons for this: the data acquisition machines require Windows (Neuralynx) or Windows/Linux (Open Ephys), and custom machine builds are generally cheaper for PCs than for Apple Macs. Nevertheless, many lab members have Mac personal laptops that run OS X, and most of the above software works fine on them. However, some Neuralynx loaders only work on Windows, some codebase functions have only been compiled to work on some OSes, and the lab only has licenses for some software (Adobe) to work on Windows.



# *Data management*

DATA IS VALUABLE. After taking into account the number of hours that went into training an animal, building a hyperdrive, preparing for and performing surgery, painstakingly turning electrodes; animal housing costs, costs of parts and consumables, and so on, the cost of a single subject worth of data can run into tens of thousands of dollars – and this is without the moral calculation of considering an animal's life. Data therefore needs to be treated with the greatest respect: specifically, it needs to be managed so that it can be maximally useful. The *first requirement* of data management is that data needs to be available – i.e. backed up so that it can never be lost, and accessible by those who need it. Our lab is committed to Open Science, of which public data sharing is one component. Increasingly, journals and funders demand that data is made publicly available. The *second requirement* of data management is that the data needs to be organized such that it can be easily understood and used.

Before discussing how these two requirements – data storage and data organization – are implemented in the lab, it is useful to distinguish between different categories of data.

- **Raw data** is what gets saved on a data acquisition machine, such as a running room computer connected to a Neuralynx recording system, or a machine connected to a microscope. Depending on the data you collect, a single session may yield many different files, such as behavioral tracking data and neural recording data, or a single file (an image). Raw data is only rarely suitable for analysis beyond a few quick checks. At a minimum, freshly acquired data sets typically must be annotated, and/or the files systematically renamed – for instance, with the ID of the experimental subject and some information about recording locations – so that the analyst can select which files to analyze, and combine results across sessions and subjects. More complex pre-processing steps include spike sorting (the process of assigning spike waveforms

to putative single neurons to obtain their spike times), artefact removal, and many others.

- **Promoted data** is data that is ready for analysis. Data files need to be organized in a specific folder structure, named according to the naming scheme, and supplied with annotations describing the data. If applicable, various preprocessing steps may have been carried out, for instance spike-sorting of raw voltage traces into spike trains of putative single neurons, filtering of position data to remove artifacts, and definition of trials on a behavioral tasks. PDFs of handwritten notes and procedure logs, and images of histology may also be included. A promoted data set typically does not include the raw data. Promoted data should be sufficiently well described and organized such that a competent reviewer or collaborator can use it.
- **InProcess data** is data that is being worked on to move it from raw to Promoted.

### *Data storage*

IMMEDIATELY AFTER DATA IS FIRST ACQUIRED, do the following:

- Create and/or rename the new data folder according to the Lab Data Naming Scheme.
- If applicable, compress any files that are large and compressible: typical examples include Neuralynx .nvt (video tracking) files<sup>52</sup>.
- Upload the data to the *Incoming* folder on the lab server. See section XXX for instructions.
- Move the data out of the location where it is saved, into a folder specific to your subject or experiment.
- Following completion of data collection from that subject or experiment, verify that indeed all data is correct and present on the server, and then delete it from the data acquisition machine.<sup>53</sup>

The lab server uses a redundant data storage system that can tolerate failure of a single hard drive without causing loss of data. The contents of the datavault folder are periodically backed up to “cold” offsite storage. However, it is good practice to also make your own personal backup of your raw and promoted data. One good way of doing this is to store that data on your desktop workstation, or move it to an external HDD, just in case.

<sup>52</sup> See section XXX for a quick primer on data compression.

<sup>53</sup> This is an important step, because if there is no space available on a data acquisition machine, we cannot acquire more data!



## *Data promotion*

What exactly is included in a promoted data set depends on the specific experiment, but typical features include the following:

- Data should be organized and (re)named according to the lab Data File Naming Scheme (link).
- Make sure raw and intermediate data files that are no longer needed are removed.
- ExpKeys file.
- If applicable, task-specific metadata file(s).
- A description of the task/experiment.
- PDFs of relevant task notes and histology.

As you collect data, you should start drafting an example ExpKeys file. Prior to starting data collection, you should have created a protocol that describes the procedures used. You should use this protocol as the basis for a description of the task/experiment.

Examples of nice promoted data with description include Alyssa's MotivationalT data (link) and Jimmie's CueCoding data<sup>54</sup>.

<sup>54</sup> <https://github.com/jgmaz/vStrCueCoding>

## *Data use cases*

Some examples ("use cases") that motivate careful data management include:

- When analyzing your own data, you want to be able to easily specify which subject(s) and session(s) you are working on.
- If you have a question that can be answered with someone else's data, you want to be able to "plug in" that data into your existing workflow. When I (MvdM) joined the Redish lab, I ran a comparison of dorsolateral striatum, ventral striatum, and hippocampus data<sup>55</sup>. This was made possible through the use of a consistent data formatting and annotation scheme.
- When publishing your work, it is the policy of the lab, and an increasing number of journals and funders, that your data is made publicly available. You want it to be easy to use and understand, so that your colleagues are not annoyed with you and you don't have to answer their emails telling you things aren't working.

<sup>55</sup> van der Meer et al. Neuron 2010

- Before you get to publishing, you will need to convince MvdM of your results and conclusions. This likely involves him running your code on your data. That will only work if you've organized your data correctly.
- MvdM is always working on applications to obtain research funding. Some of these applications, and the pilot data used to support them, are planned in advance; others are more spur-of-the-moment, or even if planned, new insights can happen. Thus, there is often a need to produce an analysis or figure in short order. This is only possible, or at least made a lot easier, if the data is ready for analysis..

## *Bibliography*

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: Is it spatial memory or a memory space?