# Introduction

This document is intended as both a general-purpose lab manual and a more detailed set of guidelines and suggestions related to open science principles, ethical research practices, and conscientious data management in the context of the kinds of research and data we most often deal with in this lab. The lab manual component lays out general information about the lab, our standard practices and guidelines, and what is expected of lab members in various roles. The latter part (from [Lab Practices](#_Lab_Practices) onward) goes into more detail and takes steps toward establishing formal lab policy around methodological transparency and data storage, preservation, and sharing, as well as some less prescriptive suggestions and points regarding issues I would encourage students to think critically about in their own work. This latter component is more so aimed at, and intended to be accessible to, early graduate students, but some parts may also be useful to honours students and other undergraduates curious about or considering pursuing research.

# Lab Spaces and Resources

The lab’s graduate student office is located in Cornett A193 and has desk room and desktop PCs with 2 monitors for three students. Currently only two of these are occupied, so the third is available for use by honours students and research assistants on a drop-in basis as long as the room is open.

We have two individual participant testing rooms, each with a desktop PC and a single monitor, located in Cornett A178 and A179. Both are equipped to run experiments using E-Prime software. These rooms are also available for student use when they are not occupied for the purpose of participant testing. Reserving these rooms for either purpose is done through a shared Google Calendar, which you will be given access to as needed or upon request.

In recent years we have shifted toward testing participants in groups for some projects, which requires reserving computer lab space in the Clearihue or Human and Social Development (HSD) buildings. Schedules for these labs can be viewed [here](https://www.uvic.ca/systems/facilities/classroomdetails/index.php). Postdoctoral fellows, graduate students, and honours students working in the lab may request to reserve these labs for up to 6 hours/week once ethics approval has been obtained for a project using [this form](https://www.uvic.ca/systems/facilities/scfbookings/researchbookings/) (students will have to list Steve as a sponsor). Booking requests and changes associated with the same project should all be made by the same person (e.g., a graduate student will typically coordinate their own bookings along with those of any research assistants working on their projects), and it is important these reservations be cancelled with as much notice as possible if you do not plan to use the room.

We currently have two active licences for the experiment development and administration software E-Prime (one for version 2 and one for version 3). These licenses are in the form of USB keys that can be used on any lab computers with E-Prime installed and are usually stored in the graduate student office. University Systems has a number of E-Prime licenses that allow for experiment administration only and can install the program in a single computer lab (20-25 people), usually one of HSD A150, A160, or A170, upon request.

The CaBS department meeting room, Cornett A069, is most often used for lab meetings but is available for other purposes (e.g., holding tutorials or training sessions, practicing presentations, etc.). Reservations are made through a shared Google Calendar that Steve has ongoing access to, and CaBS graduate students can request personal access from the administrator (currently Clay Holroyd: [holroyd@uvic.ca](mailto:holroyd@uvic.ca)).

# Lab Philosophy and General Code of Conduct

This lab is subject to university-wide policies on [harassment and discrimination](https://www.uvic.ca/equity/discrimination-harassment/index.php) and [sexualized violence](https://www.uvic.ca/equity/sexualized-violence/policy-procedures/index.php). You are of course protected by, and expected to follow, these policies in all of your activities at UVic (and in many cases by law), but it is worth reiterating that they apply equally to all interactions related to your work in the lab, whether with other lab members or research participants. There will be zero tolerance for personal and sexual harassment, bullying, and discrimination on the basis of race, colour, ancestry, national origin, religion, sex, gender identity or expression, sexual orientation, age, physical or mental disability, marital or family status, and socioeconomic status. I also recognize that various forms of privilege have the effect of discouraging or preventing marginalized groups from participating in scientific research and pursuing careers in science. With this in mind, I aim to go beyond the bare minimum standard of avoiding the harms associated with harassment and discrimination and take active steps to make the lab and our work more inclusive and accessible. I am always open to critique and suggestions on how to improve current practices in this regard.

Treating research participants and their data ethically is paramount in the work we do. If you are working in the lab you are most likely already familiar with the basics of human research ethics from research methods and other courses, but if you need a refresher you may find parts of [this tutorial](http://www.pre.ethics.gc.ca/eng/education/tutorial-didacticiel/) helpful. Your particular responsibilities as far as implementing these principles will differ depending on your role in the lab, but it is important for all lab members to know the basics. Also under the data ethics umbrella is a goes-without-saying plea to *never* falsify or alter data to make it say something more “exciting”, or what you think I want it to say; it is a fact of research that our hypotheses will often be wrong, but truth is always more important than “success”.

I strongly value open and transparent research practices and this lab, like many others, is in a bit of a period of transition from the longstanding status quo – various methodological, analytic, and data management practices we now know to be suboptimal or downright wrong – toward how we know things *should* be done. Specific expectations on this theme for students conducting original research in this lab are outlined elsewhere in this document, but I hope all lab members will gain some exposure to the principles and practices of open science through working in the lab.

We also have other ethical duties as scientists: things like striving to make knowledge and the scientific process more accessible to the public; taking steps to make our discipline more inclusive, and amplifying the voices of the marginalized within it and taking their critiques seriously; considering the potential future implications of our research, for better or worse; and thinking critically about who our work – and the ways we conduct and present it – may be serving, harming, or leaving out. Most individual research projects are unlikely to dramatically change the world, but your potential positive influence is not limited to the details of the work you are involved in; just as valuable are the examples you set in *how* you conduct you work and your contributions to changing scientific culture (in this lab and beyond), which will require both small-scale and large-scale efforts. Systemic change is somewhat beyond the scope of a lab manual, but I encourage you to think critically about the smaller steps you can take in your own work toward making things better, and to feel free to discuss such issues with me if there are things you think we can do better as a group.

# Lab Meetings

All lab members will be expected to attend weekly lab meetings if you are available at the scheduled time. Accommodating graduate and honours students’ schedules will take priority in determining lab meeting times, so it may not always be possible to accommodate all research assistants’ schedules as well. RAs who have a scheduling conflict will generally be expected to meet or check in informally with your project group/student supervisor on a weekly basis or as agreed. If you usually attend lab meetings but can’t make it on a particular day, I appreciate advance notice if possible so I know who to expect and can cancel meetings if enough people are unavailable that it does not seem worthwhile to hold a meeting.

Lab meetings are usually fairly informal and are an opportunity for everyone to get up to date on the progress of various projects, collaboratively solve issues that come up in the course of data collection or other parts of the research process, and discuss results as they come in. Some lab meetings may be reserved for training opportunities, discussing important topics (e.g., related to open science), or for lab members to present their research findings or practice for an upcoming presentation. Graduate and honours students will typically give at least one presentation per year in lab meeting, but all lab members are welcome to take advantage of this opportunity to practice and get feedback on presentations in a friendly, supportive environment. If you are interested in doing so, please get in touch with me and I will try to reserve time for you.

# Roles and Expectations

## Research Assistants

Research assistants (RAs) are most often involved in scheduling and testing undergraduate research participants, but sometimes work on other tasks such as data entry, transcription, developing stimulus sets, and experiment programming. The availability of paid positions varies from year to year depending on our current funding situation, so RAs are often volunteers. It is my goal for all RAs to have a rewarding and educational experience in the lab, but this is especially important if you are not being compensated for your time and labour. If at any point you feel you are not getting what you hoped out of your experience, I encourage you to talk to me so that we can try to find work and training opportunities more in line with your goals or, if that is not possible, point you toward other labs that might be a better fit.

Hours and scheduling  
RAs are typically expected to be able to commit to around 6 hours/week of lab work on average (including weekly lab meetings, other meetings related to lab projects, and training sessions), although the availability of hours may vary substantially over the course of the semester depending on lab needs.

Scheduling for most projects is highly flexible and you do not need to commit to working the same hours every week. If there are weeks you are particularly busy with midterms or other commitments, it is perfectly fine to take those weeks off from lab work entirely.

Projects run in the lab’s individual testing rooms (Cornett A178 and A179) will require you to book space on the lab’s Google Calendar. These rooms are kept locked so if you are working on a project that requires access we will request the department assign you a key that opens both rooms as well as the external door (A177), which will require you to bring ID and a $5 deposit (that will be returned when you return the key at the end of the semester) to the Psychology main office.

If you are working on a project that requires booking computer lab space, you will typically request space through the graduate or honours student in charge of that project. If you need to cancel or modify a requested booking it is very important to let this person know as soon as possible.

### Communication

#### Personal email

Most lab-related communication will occur by email, so please make sure the email address you have on file with me/other lab members is current and an account you check regularly. You should check this account daily on weekdays while you are actively involved in data collection in case of any time-sensitive issues (e.g., room booking conflicts that might require cancellation or room changes), unless you have arranged for another mode of contact with your student supervisor/colleagues on the same project. This is also where we will contact you about lab meeting scheduling and cancellations, which, although we try to avoid it, sometimes occur fairly last minute.

For messages requiring a response or informing you personally of something important (e.g., a change to your usual testing room), please try to respond or acknowledge by the next business day. In the latter case a quick “got it” is sufficient; it’s just helpful to know the message wasn’t lost in the ether. You are not expected to check or reply to lab email on weekends, holidays, or anytime you have arranged for time away from data collection.

#### Lab email ([lindlab@uvic.ca](mailto:lindlab@uvic.ca))

If you are involved in data collection you will be given access to the lab email account, which is generally where research participants will contact us with questions, problems, or to let us know why they were or will be unable to make it to their scheduled session. You should check this account before and after your scheduled sessions in case there are messages relevant to you. Often this will not be clear from the subject line, so if you open a message and discover it’s not relevant to you, please make sure to mark it as unread.

When communicating with participants through this account (or via SONA), keep in mind you are representing the lab; there is no need to be overly formal, but please follow basic email etiquette and include an informative subject line, a greeting, and signature. It’s preferable to sign these messages with your own name rather than just Lindsay Lab so others using the account can keep track of who’s dealing with a particular issue.

### Research and Scientific Ethics

As a research assistant you will probably not be involved with ensuring the project you are working on adheres to ethical guidelines, but you will be responsible for putting these principles into practice by ensuring research participants are treated ethically and that any data you have access to or are involved in analyzing are handled properly. If you will be running participants recruited through UVic’s SONA system, you will be expected to be familiar with the [current guidelines](https://www.uvic.ca/socialsciences/psychology/assets/docs/research-participants/Researcher-Instructions.pdf) regarding use of this system and under what circumstances participants should receive credit, be marked as excused, or be penalized for not showing up for a session.

#### Research participants

If you are involved in testing research participants, your ethical duties will include ensuring (1) informed and ongoing consent, (2) the educational value of participation, by providing verbal or written debriefing at the end of the experiment, and (3) a positive overall experience with research participation, by treating participants fairly and kindly (within reason). You will receive additional training on these expectations before you start testing, but I will review some basics here.

The informed consent process will usually take the form of a printed sheet or a screen at the beginning of the experiment that participants are required to read and sign/acknowledge. Not all potential participants read this as carefully as they should, but you should do what you can to ensure this within reason (e.g., if someone turns the sheet over and signs it as soon as you hand it to them, you should politely request they take time to read it). You should be familiar with the contents of the consent form and discuss any parts you have questions about with your project supervisor so that you are prepared to address questions prospective participants might have.

All participants who complete or get partway through a study should be debriefed regarding the purpose of the research. We usually do this verbally, based on previous experience suggesting written debriefing documents tend to go to waste. In this case you will likely receive a script for reference. If it is not possible to deliver a verbal debriefing – e.g., if a session goes over the scheduled time, a participant has to leave quickly, or extenuating circumstances require you to terminate a session early – you should make sure all participants receive this information by email, either by doing so yourself or by letting your project supervisor know.

Finally, it is important to treat participants with standard professionalism and politeness (within limits), even if they do not always do the same. This means being patient with participants when they are a few minutes late, seem not to be paying much attention, or otherwise inconveniencing you; taking an approach of reasonable leniency in deciding whether or not to penalize no-shows or latecomers who could, by the rules, be penalized; and always thanking everyone for their participation at the end. This does *not* mean that the participant is always right; providing you act within PRPS rules in applying penalties you are not obligated to give in to participant demands that you reverse such penalties, nor are you expected to put up with yelling, harassment, or otherwise disruptive/threatening behaviour. You are always entitled to terminate a session under such circumstances.

#### Confidentiality and data ethics

RAs involved in data collection will at the very least have access to participant names and contact information, and, for some projects, may also have access to other sensitive information. It is important to treat this information with appropriate care. Participants consent under the assumption of anonymity and confidentiality, which most obviously means we agree not to share any data that can be traced back to them, but also means you should not discuss individual research participants or their data with anyone outside the lab (even simply to disclose that they participated). Additionally, you must not abuse your access to participant contact information to contact anyone for reasons unrelated to their research participation. Paperwork containing participant names is typically stored securely on lab premises but if you are working on a project that requires you to keep such information on your person, or accidentally take it with you, please be very careful to store it securely.

As for electronically stored data, such as that generated by experimental software, you should **never** delete or alter raw data files for any purpose. Even if something occurs that you know will require excluding the file from later analyses, such as a computer glitch partway through, or you realize you have accidentally assigned someone the wrong participant number, you should leave the file itself as is and note the issue on the anonymized record sheet provided by your direct supervisor next to the appropriate participant number. Also, in this lab we are more interested in scientific truth than success; accordingly, you should **never** try to falsify or alter data, or encourage participants to respond in a certain way, in hopes of making results turn out the way we might want them to.

## Honours Students

Honours students will take the lead on an independent research project with my guidance, and sometimes that of a graduate student co-supervisor. Most often, this takes the form of a start-to-finish study for which you collect your own data (sometimes with RA assistance), but analytically-focused projects using already existing data from our lab or elsewhere are also possible.

For projects that require data collection, it is important we settle on a topic early in the fall semester (ideally sometime in the summer beforehand) so things can get going as soon as possible. I will usually have a few project suggestions in mind but am open to discussion with the goal of finding a mutually beneficial project that both has some elements that interest you and fits well with lab standards and expertise.

General expectations regarding [communication](#_Communication) and [ethics](#_Research_and_Scientific) discussed in the previous section also apply to honours students. If your project requires a new HREB application you will be more involved in the planning stage of research ethics than RAs, but this will be a collaborative effort and you will receive guidance throughout this process.

In addition to attending lab meetings, honours students will also meet with me (and, if applicable, others working on the same project) on a weekly basis for progress updates, to work out any problems that arise, and to discuss the research literature relevant to your project. Usually I will provide some relevant readings at the beginning of your project as a starting point and may have additional suggestions as the project proceeds, but you are also encouraged to branch out and start exploring the literature on your own. You will be expected to maintain an [annotated bibliography](https://www.lib.sfu.ca/help/cite-write/citation-style-guides/annotated-bibliography) throughout the course of your honours year, and I will provide you with feedback on this document and more general advice aimed at improving your skills in efficiently reading and evaluating scientific literature.

Regarding the thesis itself, we will set informal, section-by-section writing deadlines throughout the year in line with those recommended by the honours coordinator. I will provide feedback on draft sections as you complete them.

Because you are likely to be dealing more directly with data than RAs, you will also be subject to certain expectations regarding good data management and thorough documentation. You will also be expected, time permitting, to pre-register your research plan. Some information and guidelines related to these expectations are available in the [Lab Practices](#_Lab_Practices) section of this document, but you will receive further guidance as your project proceeds.

## [Graduate Students – to be filled in by a non-graduate student]

# Authorship

All lab members may have the opportunity to contribute to research in a way that merits authorship credit on scientific journal articles, book chapters, spoken presentations, or conference posters. I have worked with student coauthors at all levels.

Recommendations and practices regarding how authorship credit and order (first author, second author, etc.) are determined vary widely across labs and disciplines, and can be a source of conflict (or – sometimes worse – quiet, simmering resentment). My view is that the best way to pre-empt such problems is to start discussing authorship issues at the very beginning of any project that may eventually lead to a publication or presentation, and continue to discuss and renegotiate if needed as a project proceeds and roles change. I know these discussions can be intimidating for new researchers and that differences in personality and background can make them more difficult for some people than others, but I encourage you to develop confidence in taking credit for and asserting your intellectual contributions, and hope to foster a lab environment where all members feel comfortable hashing out this issues with me and/or each other.

The American Psychological Association (APA) provides some [useful resources](https://www.apa.org/science/leadership/students/authorship-paper.aspx) regarding authorship determination and “tiebreaking” when contributions are almost equal, but projects vary widely enough that adhering to a [scoring system](https://www.apa.org/science/leadership/students/authorship-determination-scorecard.pdf)  does not always make sense. The following are contributions that would usually merit authorship credit (inspired in part by the APA recommendations):

* Writing or substantially editing any part of a scientific manuscript, or developing a poster or presentation
* Substantial involvement in developing the research project, including conceptual or methodological suggestions that heavily influence the direction of the project, designing the experiment(s) or instrument(s)/questionnaire(s), or conducting literature reviews
* Analyzing data or selecting/planning the analytic approach, developing and testing relevant computational models, independently interpreting analytic results
* Developing specialized stimulus sets (e.g., one that requires a lot of photography or videography, or substantially modifying images or sound files, or a particularly time-consuming search component). Note that you would not receive authorship credit on all publications and presentations using this stimulus set in perpetuity if you are not otherwise involved with the project, but would typically be offered credit on at least the initial publication (definitely if the stimulus set itself is published), and acknowledged/thanked in footnotes thereafter
* Involvement in data collection or entry for projects where this component is particularly intensive, e.g., conducting and/or transcribing lengthy interviews

Contributions that would not usually, on their own, meet the bar for authorship, but for which you may be acknowledged by name in spoken/poster presentations and/or the footnotes of scientific publications (with your permission):

* Data collection (i.e., testing undergraduate participants) for most projects
* Stimulus set development that is limited to searching, downloading, and/or straightforward aggregation
* Simple data entry or organization
* Literature searches limited to the search itself (i.e., no reading/reviewing, curation, or judgment involved)

Determining authorship order adds a layer of complexity, but the numbers on the scoring system linked above are a good rough guide to the kinds of contributions that are usually weighted more heavily. As a general rule, Masters, PhD, and honours students will be assigned lead/first authorship on any publications or presentations heavily based on their thesis or dissertation work. By the nature of the supervisory relationship I will often be involved enough in your projects to be listed as a coauthor (second or later, project-dependent), but supervision and provision of funding/other resources are not themselves authorship-level contributions, so this need not always be the case (e.g., for projects you pursue mostly independently, as part of a course, or with others in the department).

# Funding

Most general funding issues (e.g., graduate student fellowships, payment rates for research assistantships) will have been discussed and negotiated prior to your arrival in the lab.

My capacity to fund travel expenses for conferences or workshops, purchase equipment and software to make your research easier or enable new projects, and cover registration fees for societies, courses, or certificates relevant to your research or professional development will vary depending on the expense and current grant funding. I cannot guarantee funding for these things beyond what was negotiated prior to your arrival, but if there is something that would be helpful for you to have or attend I encourage you to ask. You will never be expected to fund expenses *necessary* for your research out of your own pocket. I will always cover expenses related to printing or copying necessary paperwork, paid room rentals (if required), etc.

## Conferences and workshops

With respect to travel expenses in particular, all graduate students are eligible to apply for travel grants offered by the UVic [Faculty of Graduate Studies/Graduate Students’ Society](https://www.uvic.ca/graduatestudies/finances/financialaid/travel/), and students currently working as Teaching Assistants are also eligible to apply for funding through [CUPE 4163](https://4163.cupe.ca/resources/component-1/) Component 1 (see the section titled “Conference Fund”). Many professional societies also offer general or conference-specific travel funding opportunities for member and/or presenters. Both graduate and undergraduate students who are members of [Psi Chi](https://www.psichi.org/page/awards) may be eligible for their travel awards as well.

I will cover the costs of professional poster printing for conferences at UVic or elsewhere for lab members at any level who are interested in presenting. UVic’s Psi Chi Chapter hosts an excellent annual undergraduate research conference called [Making Waves](https://uvicmakingwaves.weebly.com/) that honours students are usually required to present at, and undergraduate research assistants are strongly encouraged to consider doing so as well. It is a very friendly, low-stress introduction to conference presentations, and you can present on research or analyses you have done in the lab, as part of a class, or even a research proposal.

Another excellent yet low-key local-ish conference members of this lab often attend and/or present at is [NOWCAM](http://web.uvic.ca/~nowcam/), which is held in mid-May and aims to showcase graduate and undergraduate student research. NOWCAM 2019 will be held right here at UVic so all students are especially encouraged to apply to present (and undergraduates who have already presented at Making Waves can apply to present the same poster).

# Lab Practices

This section is intended as a starting point toward establishing more formal lab policy regarding methodological transparency, data management practices, and research ethics above and beyond the mandated baseline. Best practices in these domains are constantly evolving so I expect this document will be regularly updated. These policies (and less formal recommendations) will be mostly applicable to graduate students who will be conducting original research in the lab over a longer term, but honours students will also be expected to adopt some of these practices, and others interested in research may find some tidbits of interest.

These guidelines have been heavily influenced by the open science movement and the avalanche of best practice recommendations that have come out of it. As mentioned earlier, we are, as a lab, currently in a bit of a period of transition in attempting to bring our own practices in line with what we believe *should* be done from an open science perspective. Some of the steps toward improving practices may involve a steep learning curve such that it may take weeks, months, or even years for it to be feasible for you to fully implement them; shifting toward a script-based workflow, for example, may be especially daunting for those with little or no programming experience. Some steps can also be easy to forget if you are not used to incorporating them into your research workflow (such as hitting the “register” button on a [pre-registration](#_Pre-registration)). There are bound to be stumbles, but as they say, “perfect is the enemy of good” – you might not be able to make your work 100% open, transparent, and reproducible tomorrow, but there are still steps you can take to make it *more* open, transparent, and reproducible. My main expectation of students and others conducting original research in the lab is thus that you be willing to learn and strive to improve your research and data management practices, and of course feel free to push me to improve mine when needed.

## Summary of Expectations

The sections below are organized roughly according to the stages of the research process, from hypothesis generation and experiment planning to long-term data management. These sections go into some detail and discuss a number of considerations/points of advice that should be considered suggestions, or things I want you to think critically about, rather than formal policy. To make it clear from the outset what I do consider policy for anyone conducting independent research in the lab:

1. All research and analysis plans should be [pre-registered](#_Pre-registration), preferably on the Open Science Framework (OSF) unless you have reason to prefer a different option.
2. To the extent that is possible, applicable, and ethical, your data, analysis scripts, and experimental materials should be made available to others in a way that does not require them to contact you for access (e.g., posted on the OSF or deposited in a dedicated data repository). These should be shared as publicly as possible given the constraints of your materials in terms of confidentiality, participant consent, and copyright considerations.
3. You should get into the habit of maintaining detailed records and organizing all files – physical and digital – related to your lab work in a fashion that will be straightforward for others (and your future self) to navigate and understand. Good practices include clear documentation of your storage system/organizational approach (e.g., in the form of “readme” files that detail the folder structure) and things like the meaning of variable names within files; using descriptive and consistent names for files, folders, and variables; and extensively annotating any scripts used for data analysis or experiment programming. A formal [data management plan](https://libguides.uvic.ca/researchdata/planning) is encouraged and may be helpful to you, but is not required outright.
4. Back up all lab data regularly, and ensure they are always stored in at least 3 locations. Your office computer and the backup server available through your office computer can count as two; the third may be a personal machine or hard drive, or the lab hard drive if needed. Cloud-based services are convenient and may be acceptable for some purposes, but should not be used for identifiable or potentially sensitive data, or anything else you would prefer kept private.
5. Related to (3) and (4), you will be expected to create a well-organized and extensively documented archive comprising all critical materials each time a project is completed. This too should be stored in multiple locations, including the lab hard drive and at least one on-site computer to facilitate access by others when you are not around or move on to other things.
6. Finally, and less concrete than the others, I hope that you will be ever mindful of the barriers to scientific participation, access to knowledge, and the pursuit of scientific careers, and think critically about the ways you can help break these down in your own work.

## Open Science in Brief

For those who are unfamiliar it may be useful to start with a brief introduction to the “open science” movement, as I will use much of the associated terminology throughout the rest of this document. In brief, open science encompasses a philosophy regarding what science should be –a collaborative, transparent, and accessible endeavour, with researchers accountable to both each other and the public – and a set of ever-evolving methodological practices aimed at bringing us closer to this goal. The utopia envisioned by open science proponents (e.g., Nosek & Bar-Anan, 2012; Nosek, Spies, & Motyl, 2012; [OSF Strategic Plan](https://osf.io/x2w9h/)) is in many ways similar to what children are taught about what science is and how it works. Unfortunately, most who are familiar with its inner workings, whether in academia or industry, are rapidly disabused of this rosy view of science as guided solely by curiosity, conscientiousness, and a commitment to sharing and updating knowledge for the betterment of all. Societal and economic realities, existing incentive structures within research and academia, and standard human weaknesses and follies such as desire for fame, power, and wealth intermix to create a science that often falls well short of what we may want it to be.

The open science movement in its current form is often traced back to the late 2000s/early 2010s, when concern regarding a number of high-profile failures to replicate research results, as well as appreciation and understanding of a panoply of poor research and analytic practices that were (and perhaps still are) in widespread use, began to grow in psychology and other fields (e.g., Collaboration for Open Science, 2015; Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011). Organizations such as the multidisciplinary [Center for Open Science](https://cos.io/) and the psychology-focused [Society for the Improvement of Psychological Science](https://improvingpsych.org/) (SIPS) began to spring up in the mid-2010s with the goals of finding solutions to the various problems that had been identified and advocating for the adoption of better methodological, analytic, and communicative practices as well as larger scale structural changes.

The open science movement is not a monolith, and many who would count themselves among its proponents disagree on its definition and the relative importance of various kinds of “openness” that have ended up under the open science umbrella. These include open in the sense of transparency regarding research plans (stating hypotheses and planned analyses from the beginning, before analyses are undertaken) and methods (sharing data, analysis code, and other research materials publicly, or at least openly with other researchers, to facilitate reproducibility); open source (the adoption of software and systems that make their source code freely available for scrutiny and modification); [open access](http://www.carl-abrc.ca/advancing-research/scholarly-communication/open-access/) (making research outputs such as published manuscripts freely available); and openness in the sense of facilitating collaboration, taking steps to make science as a whole more inclusive of groups who have been historically excluded or marginalized, and making knowledge more available and accessible to the public. I will mostly focus on the first sense in this document as it is the one most amenable to lab-level policy changes, but some of the others will crop up here and there.

# Hypothesis and Experiment Planning

## Pre-registration

Although the formal act of pre-registration will probably be one of the last steps you undertake in the experiment-planning process, this step will be much easier if you have it in mind from the start, and as it is one of the most important from an open science perspective it is worth discussing first. Much of the below is adapted from Lindsay, Simons, and Lilienfeld (2016).

### What is pre-registration?

Pre-registering a study involves creating an unalterable, time-stamped copy of your research plan, ideally before you start data collection (but at the very least before you look at your data).

### Why should I pre-register?

Pre-registration is one of the major methodological pillars of the recent push for more open science, and for good reason; done properly, it addresses a number of known problems simultaneously. One of the major benefits of pre-registering hypotheses and analyses is that it limits questionable research practices like HARKing (Hypothesizing After the Results are Known; Kerr, 1998) by removing any doubt as to which of your results were predicted *a priori* and which were unexpected. The act of compiling a thorough pre-registration is also very likely to benefit you personally by forcing you to really think through the specifics of your hypotheses, design, and analysis plans before you put them into practice, and reminding you of exactly what those hypotheses and plans were when a particular study wraps up some weeks or months after you originally generated them.

### What should I pre-register?

The very short answer: as much as possible, but anything is better than nothing. If time has gotten away from you and you are supposed to start analyzing your data tomorrow (not that this has *ever* happened to me, of course), it is probably better to pre-register a few lines stating your hypotheses than nothing at all. As cultural and institutional expectations change, so too will this advice; I expect (and hope) that in ~5 years, detailed pre-registrations will be so entrenched that any forum where you might publish or disseminate your research will require them, and there will be little point in anything that falls short of those standards. But for now I think even baby steps have some benefit, if nothing else for the purpose of getting used to adding the pre-registration step into your workflow.

The more detailed answer:

1. **Research question(s).** At the very least, a pre-registration should include a clear statement of the research question(s) your experiment and/or analysis is designed to address.
   1. **Confirmatory research:** If you are conducting a study with the goal of testing a specific hypothesis or replicating a known effect, you should state your hypotheses as precisely as possible. “I expect [variable A] to be higher in [condition X] than [condition Y]” is better than “I expect [variable A] to differ between conditions”. Arguably best of all is predicting an effect within a specific range of effect sizes (Velicer et al., 2008), if you have some theoretical or empirical basis for doing so. If your prediction involves an interaction, you should describe the form you expect this interaction to take. For complex predictions of this sort you may find it easier to provide a figure than to describe the predicted pattern verbally.
   2. **Exploratory research:** If you are doing exploratory research and don’t really have a specific hypothesis, it is perfectly fine to say so, and still worth pre-registering. Confirmatory studies often also have exploratory elements, so it is common to have an assortment of explicit confirmatory hypotheses and vaguer, more exploratory questions.
   3. **Purely analytic approaches:** Most pre-registration advice focuses on original research, but there is also merit in pre-registering secondary analyses of existing data, computational modelling projects, and (perhaps especially) large-scale analytic projects such as meta-analyses. These will also often have both confirmatory and exploratory components.
2. **Data analysis plan.** In part, this goes hand in hand with the first point; especially if you have confirmatory hypotheses, you should go into the project with some idea of how exactly you will determine whether your data support them or not. Ideally, this plan should include:
   1. **Which statistical test(s) you plan to conduct with which variable(s).** It is most important to provide this information as pertains to your primary hypotheses, but if you have exploratory tests in mind it is worth including these as well. For any analyses you have planned, make sure to state your independent and dependent variables, as well as moderators, covariates, etc. as applicable. Describe any steps you will need to take to get from your raw data to the variables you will be analyzing (e.g., adding up multiple responses on a questionnaire, transforming variables to different scales, determining parameters via model fitting, etc.).
   2. **Exclusion and/or outlier criteria.** Decisions to exclude outliers or other data from further analyses represent a major source of researcher degrees of freedom, so it is good practice to get into the habit of thinking about how you will do this before your data even exist. You should state:
      1. **How you will identify outliers, and how you plan to deal with them.** How you *should* determine what constitutes an outlier on some measure and whether or not you should exclude them from further analysis is a separate question. What matters at this stage is that you are transparent regarding your plans, even if you have none, so that anyone evaluating what you *end up* doing can compare with the original plan (and decide whether they think your decision was justifiable).
      2. ***A priori* exclusion criteria.** If you foresee any other reasons you might want to exclude a participant’s data – in whole or in part – from further analysis, these should also be pre-registered. Common criteria at the participant level in the kind of research typical of our lab might include performance below a certain threshold, failing an attention check question, or reported lack of naiveté at debriefing. At the within-participant level, examples might include responses made too quickly or too slowly, certain kinds of errors, or some subset of trials (e.g., the first *n* trials in some task that requires practice). Unforeseen reasons may well arise throughout data collection or come to mind at the analysis stage; pre-registration does not prevent you from implementing these later, but again makes the whole decision-making process transparent to others.
   3. **Relevant evidentiary “thresholds”.** E.g., for NHST approaches, you should include your planned alpha level and how you will correct for multiple comparisons.
   4. **Any planned stopping rules, or other less principled data “peeking” plans.**
   5. **Plans for dealing with missing data.** This will tend to be heavily intertwined with both points (a) and (b), as common approaches to missing data handling involve exclusion from further analyses and transformation (e.g., imputing missing values from other participants). Additionally, data might be missing on some measure crucial to your exclusion criteria (e.g., if you plan to exclude left-handed individuals from analysis, you ought to consider what you will do if someone skips the handedness question).
   6. **How you plan to evaluate whether your data violate any statistical assumptions, and how you will deal with this.**
3. **Method and design.**
   1. **Participants:** type (e.g., UVic students), any restrictions on participation, intended sample size (with justification, e.g., power analysis based on designated effect size) or stopping rule,
   2. **Variables:** Describe any variables you will manipulate and/or collect data for, especially those central to your hypotheses. For independent variables, describe their levels and the nature of the manipulation (including whether it is within- or between- subjects).
   3. **Materials and procedural details:** Essentially a standard methods section of the sort you might include in a scientific paper, but preferably with a level of detail exceeding what is often seen in the literature.
4. **Background information and justification.** This is the least crucial for pre-registration purposes, but given adequate time you may wish to include it.

### How do I create a pre-registration?

The most popular platform for filing pre-registrations, and the only one our lab has used so far, is the [Open Science Framework](https://osf.io/) (OSF). You are not obligated to use the OSF if you have a preferred alternative, but if you are new to pre-registration and/or open science more generally I highly recommend it as a starting point. The OSF also facilitates other beneficial research practices, such as collaboration, version control, and sharing data, so it will be referenced in other sections of this document as well.

The basic steps you will have to follow to file a pre-registration on the OSF are (1) [create a project](http://help.osf.io/m/projects/l/481539-create-a-project), (2) [upload](http://help.osf.io/m/projectfiles/l/482002-upload-files) your pre-registration plan to the project page (or link the project with some other service where the document is located; more detail below), and (3) [create and approve a formal registration](http://help.osf.io/m/registrations/l/524205-register-your-project)before you analyze your data (ideally before you start collecting your data). This third step is the actual *act* of pre-registration that creates a permanent, time-stamped copy of your project and any accompanying documents, so although the vast majority of the work and time that go into a pre-registration will be at the writing stage, it is crucial not to skip this last step. In addition to the links above, the OSF provides a number of other [detailed, step-by-step guides](http://help.osf.io/) to using the service, and there is no one “correct” way to use it as far as project organization and storage, so I will not attempt to provide a tutorial here.

If you are pre-registering a study for a specific purpose, e.g., a competition such as the OSF’s preregistration challenge or to comply with the guidelines of a journal you intend to publish in, there may be specific requirements as far as the contents and formatting of your pre-registration document(s). For general purposes, though, this document can be written and formatted however, and using whichever program(s), you want. Try not to worry too much about writing quality; although you’ll still want to write clearly enough to make sure someone outside the project could read the pre-registration and understand your plans/easily evaluate whether or not you stuck to them after the study has been conducted, it doesn’t need to be a publication-quality document. Bullet points and sentence fragments are fine.

There are a number of pre-registration templates available that you might consider using. I find the [AsPredicted](https://aspredicted.org/) approach a particularly accessible starting point. AsPredicted is a preregistration service that asks the author to answer a series of 7 pre-set questions (with additional space for information that does not neatly fit with any of these questions), and then generates a nicely formatted pdf along with a unique URL at which the pdf will be permanently stored; this can serve as a standalone pre-registration, or it can be used in concert with the OSF. The OSF actually has the AsPredicted template and [several other such templates](https://osf.io/zab38/wiki/home/), including some more specialized ones (for the prereg challenge, for replication studies, etc.) built in at the registration stage. When you initiate a project registration on the OSF you will have the option of selecting one of these templates, so you can, if you want, skip the step of uploading a separate document altogether and just answer the necessary questions here. The [Pre-Registration in Social Psychology template](https://osf.io/k5wns/) developed by van ’t Veer and Giner-Sorolla (2016) is a particularly solid option (and although developed with social psychology in mind is just as applicable to most research in our lab).

### What if I change my mind about something I’ve already pre-registered?

You cannot change the registered copy of your plan, but you can make changes to the project itself and file a new registration. It only makes sense to do this as long as data analysis has not yet begun; for example, if you come across a new analytic method halfway through data collection and decide it makes more sense than what you were originally planning, it would be worthwhile to make that change and re-register the project with a descriptive comment so there is a clear record of when that change was made (and evidence it was not made on the basis of how the planned analyses turned out). You are still free to make changes or conduct unplanned, exploratory analyses once data collection has begun, but at this point there is nothing “pre” about such changes.

The OSF facilitates updating your original registration via:

1. **Straightforward version control.** You can edit the original file detailing your preregistration, upload it to the same project, and, providing it has the same name as the original, OSF will recognize the new file as the “current version”. Copies of all previous versions will still be retained.
2. **Allowing multiple registrations for the same project**. You can do the above regardless of whether you have already registered the project or not; the difference if the project is already pre-registered is that any changed/added files will not appear in the pre-registration copy. If you make changes following the initial registration but still haven’t analyzed your data, it is worth re-registering the project, thus creating a new time-stamped copy (but also maintaining the original). You can, as far as I know, do this as many times as you want. Although all versions will be retained such that anyone could in theory figure out what had changed each time, it is good practice to include a brief description of the changes when prompted (e.g., “changed planned sample size”, “lowered presentation time at study to get performance off ceiling”).

A personal, admittedly anecdotally-rooted recommendation for anyone getting used to pre-registration or who may be a bit unnerved by the thought of creating a permanent record of your no doubt error-ridden first draft: get into the habit of registering early and often. Register as soon as you fill out/upload the first version of your plan, even if data collection/analysis will not start anytime soon. Re-register if you change anything prior to data collection, and each time you change something after data collection has begun. There may be some downsides to this in the form of confusion on your end as to which is the “correct” pre-registration for a particular purpose, and potentially raised eyebrows if others see you have a long list of pre-registrations for the same project (but this should be mitigated by the total transparency of this process). But if you are at all intimidated by the process it may help entrench the habit and get rid of any intimidation associated with the finality of hitting that “Register” button.

Preregistration is a major step toward making your research more transparent, and composing a thorough preregistration can also encourage you to think more critically about the whats, whys, and hows of the research you’re conducting. You might catch logical errors in the arguments underlying your hypotheses and/or planned methodological approach, recognize certain elements of your research questions are vague or ill-formed and sharpen them accordingly, or realize you aren’t exactly sure of the best way to analyze your data to address your primary hypotheses. The potential benefit to science of doing things this way thus extends beyond preventing (or at least rendering transparent) more obviously questionable research practices.

The remainder of this section focuses on this theme of thinking critically about the decisions made at the stage of hypothesis and experiment planning, including decisions about research participants, experimental setup, and stimulus selection. Some of these decisions are often made with so little thought they do not even feel like decisions; they are simply the default, or how things are usually done in a particular lab or the field as a whole. I will not get into higher-level questions of how to formulate better hypotheses nor basic research methods/experimental design principles (although these are both worthwhile topics). Instead, I will focus specifically on lower-level changes that can be made at each step to make your research more open in the open science sense – i.e., more reproducible, transparent, and easily accessible – and in the sense of inclusivity.

## Human Research Ethics

The dominant ethics framework governing scientific research involving humans in Canada is the [Tri-Council Policy Statement](http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/), currently in its second edition (TCPS 2). Eligibility for tri-council agency funding ([NSERC](http://www.nserc-crsng.gc.ca), [SSHRC](http://www.sshrc-crsh.gc.ca/), or [CIHR](http://www.cihr-irsc.gc.ca/)) is contingent on compliance with this policy (and other tri-council policies pertaining to other kinds of research), and institutions involved in research in Canada have their own research ethics committees responsible for ensuring all research conducted under their jurisdiction complies with these guidelines (and relevant laws). At UVic the relevant committee, from whom approval must be obtained prior to initiating any data collection (aside from informal pilot testing with friends, other lab members, etc.), is the [Human Research Ethics Board](https://www.uvic.ca/research/conduct/home/regapproval/humanethics/) (HREB).

Although seeking HREB approval will not likely be one of the first chronological steps in your research planning process, the decisions discussed in later sections of this manual must of course be made with ethical principles in mind, so I have opted to start that discussion here. The vast majority of research conducted in this lab is in the form of experiments or surveys administered on campus or online, all of which unambiguously require formal HREB approval. The only possible exceptions are:

* **Projects based on the analysis of existing data from our lab, or other labs, for which ethics approval was obtained prior to collection:** You generally do not need to apply for a new HREB approval to re-analyze data from previous projects in our lab or to analyze data obtained from other researchers, providing the original approval can reasonably be assumed to cover your intended use.
* **Meta-analyses:** Similar to the above, projects based on synthesizing data that have already been disseminated in scientific articles or at conferences, and/or unpublished data that have been provided to you, do not require new HREB approval.
* **Projects based on the analysis of existing data that is publicly and legally available:** You do not have to apply for HREB approval to conduct analyses using public datasets (e.g., from [Statistics Canada](https://open.canada.ca/data/en/dataset?organization=statcan)), or other publicly accessible information with “no reasonable expectation of privacy” (e.g., data collected from books; newspaper articles; or online sources such as websites, forums or social media posts, providing they are entirely public [i.e., not restricted to friends or members]).

Making determinations such as under what circumstances people have a reasonable expectation of privacy is not always clear cut. As such, if you are planning to conduct a project that falls into one of these categories and are at all uncertain as to whether it meets the exemption criteria, you should always err on the side of contacting someone at the HREB.

If you are planning a new study, Steve/your other supervisor(s) will help you determine whether your proposal falls under an already approved ethics protocol, can be accommodated by amending an existing protocol, or will require an entirely new HREB application. This manual does not aim to provide a guide to that process, which the HREB website has well covered, nor to fully outline the TCPS guidelines. Instead, I hope to (1) reiterate the importance of a few fundamental ethical research principles worth keeping in mind when planning, designing, and disseminating the results of a study, and steps that can be taken to bring psychological research in closer adherence with these principles (above and beyond complying with HREB policies, which is of course non-negotiable), and (2) draw particular attention to situations where open science ideals and human research ethics may conflict, and suggest ways of balancing these concerns.

### Potential harms, benefits, and the importance of “taking the perspective of the participant”.

Most of the research we conduct is considered “minimal risk” from an ethical perspective. Standard memory experiments wherein participants are tested on memory for a list of banal words or pictures, or asked to watch a video and then determine whether someone in the video is featured in a subsequent lineup, are generally unlikely to cause harm, and tend not to be deemed any riskier to participants than the baseline of everyday life. As the TCPS emphasizes, however, researchers should not assume all participants will perceive the balance of harms and benefits the same way they do. It is important to keep in mind that research participants (UVic undergraduates or otherwise) come from a variety of backgrounds and may have histories of oppression, abuse, and/or trauma we are not privy to. In the specific context of memory research, this possibility should be kept in mind when:

* 1. **Selecting stimuli:** Most of the stimulus sets we use are relatively neutral or unlikely to cause harm (lists of nouns, images of paintings, photos of faces, videos of staged, non-violent crimes, etc.). Certain research questions (e.g., regarding the influence of emotion on memory) may require stimuli that are overtly violent, sexual, or generally unpleasant. Although the HREB will make the ultimate decision as to whether such a project is justified by the potential benefits, this is a question worth considering very seriously before you even get to this stage.
  2. **Developing questions, and deciding whether to ask for certain information:** There is a need to balance collecting information that may be relevant to your research question (or which you are just curious about) with not only the confidentiality standard, but ethical treatment more generally. Some would argue that collecting certain kinds of demographic information about your sample – even if they are not directly relevant to your research question – is a kind of methodological transparency, and that this is important information for would-be replicators of your work to have. There is some validity to this, but for most purposes in our lab I think an aggregate-level summary based on the overall characteristics of the UVic participant pool is probably sufficient. That said, I am not steadfastly opposed to collecting extraneous information related to variables you think may be interesting to analyze in an exploratory fashion providing you collect and store this information with due care. If you do opt (or need) to collect potentially sensitive and personal information related to things like ethnicity, gender identity, sexual orientation, etc., here are a few considerations and suggestions to keep in mind:
     1. **Privacy and confidentiality.** Allow participants to fill out such questions themselves rather than asking them directly, and preferably not with the experimenter or other participants looking over their shoulder. You will have to think critically about what you will do with this information, how you will store it, whether you will remove it from any data you share, etc. to ensure confidentiality is preserved (discussed more [below](#_Maintaining_and_storing)).
     2. **Transparency**. In addition to your ethical mandate to be transparent about how this information will be stored and whether you plan to share it as part of the consent process (also discussed more below), you should be open with participants about *why* you are collecting this information. If you do not need it, but are just interested in using it in exploratory analyses, say so. Participants can then decide whether they are okay with providing this information for the described purpose.
     3. **Autonomy, sensitivity, and respect.** Keep in mind that a question that may not seem at all fraught or sensitive to you personally might be perceived differently by others. Consider leaving all personal/demographic questions open-ended so participants are entirely in control of how they describe themselves, or including “rather not say” and/or “prefer to self-describe” options. If you think you need categorical response options, be sensitive in selecting and wording these options. For example, you should avoid the common well-intentioned formulation of gender options as “male, female, or transgender”, and options that conflate sexual orientation and gender identity. [This article](https://medium.com/@anna.sarai.rosenberg/respectful-collection-of-demographic-data-56de9fcb80e2) by Sarai Rosenberg is an excellent resource on respectfully requesting demographic information, including that related to gender identity and sexual orientation specifically.
  3. **Discussing the purposes and implications of your research (e.g., when debriefing participants**): In talking about research, we want to convey to participants why it is important, and an easy way to do this is to bring up concrete, evocative, real world examples. This is to be encouraged, but these examples should be chosen with care and an understanding that what may be, to us, a subject of intellectual inquiry foremost may map onto painful lived experiences for others. Perhaps one of the most widely known real-world influences of academic memory research has been in the context of debates (especially explosive in the 1990s, but still ongoing) regarding the validity of recovered memories, especially the potential of suggestive therapeutic practices to elicit false memories of childhood sexual abuse. This is undoubtedly a hugely important research subject students of psychology should be aware of, but is sometimes discussed in academic circles with a level of disconnect I imagine may be anywhere from alienating to deeply hurtful to individuals who have personally dealt with the traumas of abuse. Such examples should not be presented lightly or without warning.

### Informed Consent

I will not deal with defining informed consent or reviewing the basic elements of the process in this document (although certain elements are discussed [below](#_Free,_ongoing,_and)), but a crucial part of the HREB application will be conveying how participant consent will be assured, and any relevant forms must be provided. The lab has a number of templates you can adapt if needed. With respect to open science and data management, it is important to incorporate clear information regarding which data will be shared, where you intend to share them, and who will have access into the consent process. Here is an example from a consent form recently submitted for HREB approval:

To protect your privacy, all of the information that you provide to us after this consent form will be coded by a number untraceable to you. We will retain the resulting data files indefinitely. It is anticipated that the results of this study will be shared with others in published academic papers and presentations at scholarly conferences and may also be described in theses written by students. We also plan to share our data files publicly on the Open Science Framework (see osf.io) to facilitate collaboration with other researchers, increase the transparency and accessibility of our research, and ensure our analyses can be fully checked and reproduced by anyone who wishes to do so. These data files will include any responses you enter into the computer program used for the experiment, and in some cases the number of seconds taken to respond (i.e., response/reaction times). You will not be asked to enter any sensitive or potentially identifying information into this program, so your data file will be fully anonymized, labelled only with the untraceable number mentioned above. Date and timestamps will also be removed from any files shared publicly to ensure that even someone who happens to know exactly when you participated would be unable to identify your data.

In most cases (especially if you are participating in a group session) we will be unable to identify your file after you finish the experiment, so if you are not comfortable with your data being shared in this way you should decline to participate. That said, near the end of the session you will be given an opportunity to change your mind about your anonymous data being shared.

For experiments where it may be difficult to explain exactly what data participants are consenting to share before they actually take part, you should include a secondary data consent step at the end of the study, giving participants an opportunity to withdraw consent to share their data after they have entered all the data they will be asked to enter. This may lead to some amount of data loss from very cautious, privacy-conscious folks, but I think it is good practice. If you find you are losing >5% of data to opt-outs at this step, you may want to reconsider your wording, or whether there may be particular questions you are asking that are producing this attrition (in which case you could either remove these questions or, if they are important, plan to exclude them from the data you share).

## Participants

This section discusses questions of determining the nature of your study sample and sample size in more detail than is probably necessary for most projects, but the main points can be summarized as follows:

* **Generalizability concerns:** Although we rely on the UVic undergraduate participant pool for most of our research, this is a skewed sample of Canadians (let alone humanity as a whole, the population psychological researchers are ostensibly studying) on many dimensions (education, socioeconomic status, etc.). For certain questions and goals it may be necessary or advisable to seek out alternative populations by collaborating with other researchers, advertising in the community, or recruiting participants online.
* **Online data collection**: Collecting data online has the potential to mitigate problems related to the unrepresentative nature of undergraduate samples; however, this approach does come with its own concerns such as lack of control over the experimental setting, the prevalence of professional participants (i.e., people who have participated in a lot of academic research) on platforms designed for this purpose, and the ethics of compensation. If you are planning to collect data online using a crowdsourcing platform, you should be cognizant of the business model of the platform you intend to use and the reality of who participates on that platform. As a general rule, participants should be paid at least the US minimum wage, and work should only be rejected (i.e., not paid for) in extreme cases.
* **Fairness and equity in participation:** Groups of people should only be prevented from participating in research on the basis of personal/demographic characteristics if the research question necessitates this, and excluding data after the fact (instead of restricting participation outright) is not a viable option.
* **Principled sample size planning:** Sample sizes should not be pulled from thin air or based on what other people have done, but should be based on principled consideration of factors including the effect size you expect (or the smallest one you are interested in detecting) and how important it is to limit the probability of type I and II errors. There are programs available to calculate the necessary sample size given these and other parameters. Alternatively, you may wish to look into approaches that allow for validly stopping data collection once evidence in either direction reaches a certain threshold.

### Who will they be?

The “right” answer to this will of course depend on your research question, but many questions asked in this lab are indifferent on this point such that choosing participants tends to come down to ease of access. At a broad level, by far the most commonanswer in psychological research is “undergraduate students”, and this lab is no exception (indeed, the population is even more restricted, namely “undergraduate students registered in a Psychology course who choose to participate in research for bonus points”).There is a growing appreciation of the problems this may pose for generalizability in scientific disciplines that ostensibly aim to understand human behaviour as a whole. The bulk of social science research is conducted with WEIRD participants (Henrich, Heine, & Norenzayan, 2010) – that is, individuals hailing from Western, Educated, Industrialized, Rich, and Democratic regions, meaning many effects and conclusions in the literature are based on a highly skewed demographic sample, and psychology undergraduate pools like UVic’s represent a further restricted sample that tends to be mostly young and female-identifying.

In cognitive psychology, the underlying goal of many of the questions we ask is to better understand the fundamental mechanisms of memory, attention, perception, etc. At some level, of course, these mechanisms and the machinery that allows for them are shared by all humans; an optimistic view, then, is that it probably doesn’t matter who participates in a given experiment, as the basic principles we eventually uncover will generalize to everyone. Of course, because we can only ever peer at this level *indirectly*, this is quite a logical leap. The truth is that we don’t know how generalizable much of our research is, and it is entirely possible there are oft-cited effects floating around that we think tell us something about how the mind/brain works, but are in fact specific to WEIRD participants for some reason we don’t yet understand. This is a field-level concern that will require large scale, systemic changes to make serious progress on. At the level of an individual experiment, one obvious solution is to collect data from a broader-than-usual sample, perhaps by recruiting from the community rather than only the undergraduate pool or, better yet – in terms of both ease of recruitment/administration and potential demographic span – the internet. The following two sections will discuss some important considerations associated with collecting data from undergraduate and online samples, by far the two most common approaches for past research in this lab.

#### UVic Undergraduates

For much of the work we conduct, it makes sense to at least start by running studies with participants from the UVic undergraduate research pool, so it is very likely you will be working with this kind of sample at some point during your time in the lab. Much of the discussion surrounding human research ethics focuses on ethical treatment of participants, and this will be discussed in the [Experiment](#_Experiment) section, but there are also ethical considerations involved in deciding who does and does not get to participate.

“Researchers shall not exclude individuals from the opportunity to participate in research on the basis of attributes such as culture, language, religion, race, disability, sexual orientation, ethnicity, linguistic proficiency, gender or age, unless there is a valid reason for the exclusion.” (TCPS 2, Chapter 4, 2014)

If you are planning to run your study with UVic undergraduate participants, you will have the option to impose eligibility criteria and restrict participation on certain bases. It is an ethical imperative that participation be open to as many people as reasonably possible given the constraints of your research question. The goal of absolute inclusivity will sometimes conflict with other pragmatic and/or ethical considerations such that “reasonably possible” can be a matter of personal judgment. This section is not intended to be prescriptive, but attempts to set out some general guidelines and suggest possible alternatives to restricting participation under particular circumstances.

The last clause in the above quotation leaves a lot of wiggle room, and inclusion/exclusion criteria are a clear example of harms that can become amplified at the collective level. It is not a huge problem if one experiment requires “normal or corrected-to-normal” vision, for example, and this exclusion may even be necessary, but if visually impaired participants are excluded from most research at UVic this is clearly unfair, and if people with visual impairments are underrepresented in psychological research in general this is a failure of science to serve society.

To be clear, there *are* valid reasons for exclusion, and for particular research questions exclusions may be unavoidable. Exclusion criteria can even, somewhat paradoxically, serve to *increase* fairness and equity overall, e.g., if you are studying some effect that has predominantly been investigated with WEIRD participants and want to expand your inquiry to underrepresented groups. But such criteria should only be imposed after careful consideration of ways you might adjust the design of your experiment and/or data analysis to allow for as few restrictions on participation as possible.

For the purposes of recent and ongoing research in our lab, I can think of no instances where it has been or would be remotely justifiable to exclude people from participating on the basis of race, gender or gender identity, age, sexual orientation, national origin, culture, or religion. This does not mean none of these variables are ever relevant to our research questions – e.g., there is extensive research into memory for faces, voices, etc. of individuals who are similar vs. dissimilar from participants on the basis of various demographic characteristics – but any such questions are typically addressed at the analysis stage rather than restricting participation. This discussion will therefore focus on exclusion criteria most likely to come up in our work: English fluency, visual impairments (including restricted colour vision), and hearing impairments.

Because “successful” participation depends to some extent on comprehending verbal instructions (written and/or described aloud) and our research questions sometimes directly pertain to semantic meaning, prior experience with certain words/concepts, etc., it may sometimes seem appropriate to impose exclusions on the basis of English fluency or “first language” requirements. Similarly, the validity of a participant’s data for the purposes of a given research question may depend on their ability to see and comprehend an image, discriminate colours, etc. More rarely, we may use auditory stimuli, which may pose a challenge for participants with various hearing impairments and/or the interpretation of such data. In all of these cases, **whenever possible, it is better to make exclusions at the data analysis stage rather than restrict participation outright.** There are several ways to address this:

1. **Exclusion decisions can be left up to experimenter discretion.** This is particularly easy in one-on-one sessions – if you can tell a participant is struggling with the experiment or does not understand the instructions, you can simply make a note next to their participant number after the session, and train RAs to do the same.
2. **Participants can be excluded from analyses on the basis of their responses to a direct question about the criterion of interest.** We often include a few multiple choice and/or open-ended questions at the end of experiments, so you might ask about the criterion in question at this stage and decide to exclude participants who select a certain response.
3. **Participants can be excluded on some indirect basis that will tend to capture most cases relevant to the criterion in question**. If some part of the experiment cannot be performed successfully without understanding the instructions, discriminating between certain colours, etc., setting a simple performance cut-off should suffice. Alternatively, I have seen examples of tasks or questions designed to evaluate things like English proficiency without participants necessarily knowing, particularly in the context of experiments administered online. This should be balanced with the ethical imperative of informed participation and thorough debriefing.

#### Online samples

The internet seems to provide a clear if imperfect solution to the problem of overreliance on WEIRD samples in psychology. Of course, conducting research online is not a panacea; regular internet access remains an inequitably distributed privilege (Poushter, 2016) such that even a sample representative of the entire internet user base would be far from representative of humanity, and in reality demographics tend to be further skewed by factors like how participants are recruited, publishing surveys in English only, and providing little or no financial compensation. Nonetheless, the internet can at least in theory provide access to a more diverse sample than the UVic participant pool and even the local community, so an online sample may be worth considering if you are interested in trying to improve the overall generalizability of your research, establish the generalizability of a particular effect, or specifically reach out to underrepresented populations.

Providing you follow ethics guidelines and receive explicit HREB approval for a given recruitment/advertisement method, you can recruit participants online in a number of ways, e.g., by linking to your study on social media, posting on any of the numerous dedicated platforms for research recruitment, or even advertising on classified websites. The best approach in a particular case will depend on your goals in collecting data online, and this is something you will discuss with Steve and others involved in your project, but some guidance on crowdsourcing platforms in particular is provided below.

##### Crowdsourcing platforms

An increasingly popular approach to online data collection among academics is the use of crowdsourcing platforms, including [Amazon’s Mechanical Turk](https://www.mturk.com/) (MTurk), Crowdflower[[1]](#footnote-1), and [Prolific](https://prolific.ac/). We have used the first two of these for research in the past and do not currently have a formal lab policy regarding which platform should be used, or forbidding the use of any particular platform. With that said, for reasons discussed in more detail in the appendix for those interested, I would strongly discourage the use of MTurk unless you have access to sufficient funding to ensure participants are paid **at least** the US federal minimum wage of $7.25/hr[[2]](#footnote-2) or the minimum wage in the region you are targeting. I would also strongly encourage you to take a look at these [Guidelines for Academic Requesters](http://wiki.wearedynamo.org/index.php/Guidelines_for_Academic_Requesters) developed by MTurk workers. I am less familiar with the demographics and inner workings of other, smaller platforms, but would encourage generous compensation as a general rule. Prolific takes part of the decision-making out of your hands by imposing a minimum payment of £5.00 (British Pounds) / $6.50 (US Dollars) per hour. Concerns have also been raised regarding data quality on sites like MTurk, such as the potential for bot infiltration and the implications for generalizability and validity of collecting psychological data from people who are essentially “professional participants”. These considerations are all worth weighing seriously.

### Sample size planning

A major thread of the replication “crisis” and responses to it has been increasing attention to the high prevalence of underpowered studies in psychology and neuroscience (Button et al., 2013; Vankov, Bowers, & Munafò, 2014). Low statistical power – to which small sample sizes are one contributor – compromises researchers’ ability to detect true effects, especially if they are small, and means effects that *are* detected are likely to be overestimated (because when power to detect smaller effects is low, only samples in which the effect happens to be large by chance will produce significant results). At the collective level, the tendency for studies to be underpowered compromises the validity of meta-analyses and can lead to skewed perceptions of just how robust some effects are. As with many issues in psychology, this is not a new discovery (Cohen, 1962), but the potential scope of the problem and the need to change practices in response have only begun to be taken seriously on a large scale within the last decade.

Part of the response to this has been a movement toward more principled sample size planning. There are differing opinions on the best way to approach this and what exactly those principles should be, and you may develop your own philosophy over time, but you should at least get into the habit of running some sample size calculations[[3]](#footnote-3) prior to initiating a new experiment. You can do this using free programs such as [G\*Power](http://www.gpower.hhu.de/) , which offers a point-and-click interface, or R packages such as [pwr](https://CRAN.R-project.org/package=pwr). The key things you will have to consider are:

**The size of the effect you are looking to detect**. If you are working with a very well-known, widely studied effect, there may be reference points in the literature for how large you should expect this effect to be. Sample size planning in attempts to replicate previous work are often based around the effect size reported in the original sample, which seems reasonable enough on its face. Although basing sample size estimates on effect sizes from the literature is common and perhaps better than nothing, you should be very cautious with this approach for the reason outlined above – that is, *widespread low statistical power means effect sizes in the literature are very likely to be overestimates.* Estimates from meta-analyses may be a better way to go, as such approaches often incorporate corrections for publication bias, but even this is not failsafe. For all but the most robust, widely replicated effects, effect sizes from the literature should only guide your decision, not be strictly relied upon (unless you are explicitly only interested in finding an effect of that size).

This leads to my own personal preference as far as the effect size upon which sample size planning should be based, namely the smallest effect size of interest (or SESOI; Lakens, 2014). If you are only interested in detecting a sizable effect, use a correspondingly large effect size in these calculations; if you think the effect is important enough that detecting a difference of 0.2 or 0.3 standard deviations between groups would still be worth the additional sample size burden, that is a perfectly principled decision. You may find in the course of these calculations that for some designs the required sample size to detect such small effects stretches the limits of what is practical (or even ethical, if you are relying on a limited pool such as the UVic one that other researchers also draw from), and will have to balance these considerations. Ultimately, there is no correct answer to the question of what your sample size should be, but you should get into the habit of thinking critically about it and justifying your choices in preregistration documents and manuscripts.

**Your alpha level** (probability of a type I/false positive error)**.** 0.05 (5%) is the go-to benchmark, and although arbitrary it is perfectly fine unless you have a principled *a priori* reason (e.g., planned multiple comparisons) to adjust it.

**Your desired power level (1 – beta, beta being the probability of a type II/false negative error).** Although alpha and beta, and the notion of a need to balance them in considering your analytic approach, are often introduced in tandem in introductory statistics courses, historically the norm seems to have been (and may still be, despite some positive change) to largely neglect consideration of beta after that. But determining how much statistical power you want to have – that is, how confident you want to be of avoiding false negative errors – is central to sample size planning. Much like setting the alpha level, there is some inherent arbitrariness in this choice, and like all other parameters that go into sample size determination it ultimately comes down to trade-offs between desirability and pragmatism. A common benchmark is 80% (so a 20% chance of failing to detect a true effect under your various other parameters), but I would advise aiming higher when practical.

**Your design.** If it is valid and possible to address your research question using within-subjects manipulations, the same sample size will yield much higher statistical power in such a design than in a between-subjects one. Of course, this is not always desirable or even possible. There may also be additional parameters you have to set depending on your particular design – for example, in a mixed design with both within- and between-subjects manipulations, you will need to set things such as the anticipated within-subjects correlation and the number of factors of each type.

## Experiment Design

This section will briefly discuss the open science-y considerations associated with the actual design and programming of an experiment. From this perspective, the ideal experiment is one that can be programmed using open source, freely available means, such as [OpenSesame](https://osdoc.cogsci.nl/), [PsychoPy](http://www.psychopy.org/) (both of which use [Python](https://www.python.org/), but offer free graphical user interfaces [GUIs] that allow you to do at least some things via point-and-click), or [jsPsych](https://www.jspsych.org/) (de Leeuw, 2015), which uses JavaScript. This offers advantages for reproducibility, knowledge-sharing, and research efficiency, as anyone can in theory examine and test the resulting experiment script, clearly see what settings were used, and adapt it for their own purposes rather than starting from scratch. These tools also allow you to avoid the costs associated with purchasing proprietary software such as E-Prime, SuperLab, or Matlab, which can be steep or insurmountable for researchers with little access to funding or working at underprivileged institutions. More widespread adoption of such tools can thus indirectly benefit such researchers.

That said, there are various reasons open source tools may not currently work or make sense for you – perhaps you have inherited a project from someone else that relies on proprietary software, or need access to some feature that is not currently available or difficult to implement using these tools (e.g., I am not sure if any of them yet offer as fine-grained control over timing/screen refresh rates as E-Prime; this is not a concern for most behavioural experiments in our lab, but may be if you are interested in EEG). If this is the case, there are still steps you can take during and after the course of designing your experiment to improve the transparency, accessibility, and reproducibility of your research. You can share your experiment design files (e.g., the .es/.es2/.es3 files generated by E-Prime) on the OSF or elsewhere so at least those who have access to the program can reproduce your work. You can (and should) also carefully document, report, and share the details of your method, including any settings you adjust in a point-and-click fashion. You should do this in sufficient detail that someone would be able to adapt your exact experiment in whichever program they might have access to.

Finally, if you are using any custom code/script in your experiment – whether Python, Javascript, the Visual Basic-based language used by E-Prime, etc. – you should get into the habit of annotating extensively, as discussed in more detail in the [Data Analysis](#_Data_Analysis) section.

### Stimulus Selection

Another decision point at which your research can be made more open and reproducible is in choosing and/or developing stimuli, which may be words, sentences, images, videos, etc. The ideal here is to use stimuli that are freely available, or to design your stimuli with the intention of making them freely available (what this means will depend on the stimuli; e.g., if you are taking photos of people, you will have to design your consent process and ethics application to ensure you have approval to share them in perpetuity). This will make it possible for others to reproduce your work, conduct replications of your studies that are as exact as possible, or use the stimuli in original research. However, it may also be important for your research question that certain kinds of norm data – e.g., word frequency, visual complexity, emotion ratings – be available for your stimuli, and the goals of “freely available” and “extensively normed” will sometimes be in conflict. There may also be valid justifications for keeping certain stimulus sets, or other experimental material, difficult to access; it may be important for data quality or research validity that materials not become widely known, or that the experiment setting be participants’ first ever encounter with a particular set of materials. Questionnaires used in clinical or other assessment settings, for example, would lose diagnostic value if the details, scoring criteria, etc., were widely available.

If you are using a stimulus set designed by others (freely available or otherwise), make sure to keep track of the source so you can provide proper credit in any presentations or publications based on studies using these stimuli. Relatedly, you should pay attention to any details provided regarding usage rights – although the fair dealing provision of the Canadian [Copyright Act](https://laws-lois.justice.gc.ca/eng/acts/c-42/) allows for even copyrighted material to be used freely for research purposes under most circumstances (UVic Libraries is a [great resource](https://www.uvic.ca/library/featured/copyright/) for details on copyright considerations, including [fair dealing](https://www.uvic.ca/library/featured/copyright/fairdealing/)), certain kinds of licenses may prohibit you from modifying any stimuli or sharing the set outside the research context. You should also be careful to track any changes you may make to the set (e.g., removing images) and your reasons for doing so.

# Experiment

## Ethical treatment of participants

If you are working, studying, or volunteering in this lab, you most likely have at least some familiarity with the fundamental ethical pillars of psychological research: free, ongoing, and informed consent; the importance of minimizing possible risks and harms, and ensuring these are outweighed by benefits; maintaining participant confidentiality, and so on. Although the HREB is responsible for ensuring a given project adheres to these principles, anyone involved with testing participants or handling data in the lab has a responsibility to ensure they are implemented in practice.Below are some guidelines for ensuring ethical conduct and going beyond what is required to make participants’ experience as pleasant as is reasonably possible. If you are new to testing participants, this will be part of your training, so don’t worry if you are overwhelmed by the information below; you will be given examples directly relevant to the project you’ll be working on and are encouraged to ask questions during this training and on an ongoing basis.

### Free, ongoing, and informed consent

Consent must be fully voluntary and informed, and can be withdrawn at any time. This is, in theory, accomplished by providing participants with a detailed consent form outlining what they can expect, what will be done with their data/how we will ensure confidentiality, and emphasizing their right to decline participation outright or leave at any point. You will find most participants do not take this step very seriously and it starts to seem like an administrative box-ticking exercise. Still, you can do your best to ensure the integrity of this process by:

* Trying to make sure participants read and understand the consent form before signing – you can only do so much here, but if someone very obviously just flips the sheet over to sign, please ask them politely (and discreetly in a group setting) to read the contents before starting the experiment.
* Letting participants know they are welcome to ask questions before signing if they have any concerns.
* Reiterating at the beginning of the experiment that participants are free to leave at any time without penalty, and to request their data not be used.
* Being prepared to deal with the (very rare) event in which a participant does choose to withdraw consent during or after the experiment. Participants must be allowed to leave without question at any time (although you can of course ask if they are okay, need you to contact anyone, etc.), and any request their data not be used must be honoured – these are the only circumstances under which data should be wholly deleted.

### Minimizing unnecessary risks and potential harms

Most of this is taken care of in designing the experiment**,** and for most experiments we conduct participants will not generally be at risk of anything more serious than boredom. But it is not possible to eliminate the possibility of harm in the experimental setting entirely – unexpected harms may arise, or participants may experience crises that are not directly related to the experiment itself. In the event a participant is visibly in crisis or approaches you with concerns related to the experiment or otherwise:

* Try to respond with basic compassion and discretion, even to concerns that may seem silly or invalid to you. In a group setting you may want to offer privacy if this seems appropriate (e.g., leaving the room to talk in the hallway).
* Keep in mind that you are not a mental health (or legal, security, etc.) professional nor are you expected to behave as such, and offer to refer them to services as needed. You will be provided a list of relevant contacts as part of your testing materials, including supports for personal crises and reporting ethical concerns.
* Safety, security, and mental health – both participants’ and yours – are more important than research, and you should ALWAYS feel free to end an experimental session early for these reasons. In a group setting, something like “sorry, something unexpected has come up and I have to end the session early; you’ll still receive full credit, and someone will contact you with debriefing information” is entirely sufficient, but these details can also be sorted out afterwards so don’t worry if you don’t remember in the moment.
  + Related to this, you do NOT have to tolerate abuse, harassment, or mistreatment by a participant, even if they are in crisis or you feel somehow responsible for the situation. If an interaction escalates to this point, feel free to take measures such as asking a participant to leave and contacting emergency services or Campus Security if your safety is at risk.

### Debriefing

It is important that participants are made aware of the purpose of the research they have contributed to. There is a particular ethical imperative, clearly stated in policy documents such as the TCPS, to do so if the experimental design involved any deception. In such cases, the deception must be disclosed, and participants must be provided with justification as to why it was necessary and explicitly given an opportunity to withdraw their data. But even for run-of-the-mill, deception-free, minimal risk studies – i.e., most of what is conducted in our lab – it is important people be given the opportunity to glean some educational value from the participation experience. For studies conducted on campus, researchers in our lab usually deliver a verbal debriefing at the end of each study session. It is possible to administer a written debriefing instead, and there are certain circumstances under which this may be desirable or necessary, such as when a session runs late (in which case the debriefing can be emailed)or for experiments conducted in groups that are likely to have large variations in completion times such that it would be unduly inconvenient to make all participants stay until the others are finished. However, personal experience suggests printed debriefing sheets are rarely read, so discussing the experiment verbally is usually preferable in terms of making it *slightly* more likely participants learn something. For online studies, debriefing is usually delivered in the form of a text summary at the end of the experiment.

There are some circumstances under which participants should be debriefed even if they do not finish the experiment. A good rule of thumb in the on campus context is that *almost* everyone who receives credit for a study should be debriefed regarding the purpose of that study, the only exception being individuals who could validly participate in the same study in the future – e.g., if a computer glitch in the very beginning forces you to end a session, or the experimenter does not show up at all, individuals in such sessions can participate another time for full credit and should not be debriefed unless they indicate they are not interested in participating. Potential participants with disabilities that preclude them from standard participation should be fully debriefed, as should individuals who opt to withdraw consent during the experiment (although they of course should not be forced to stay to hear the debriefing information if they are dealing with an urgent situation or clearly uncomfortable).

You are encouraged to consider going above and beyond in terms of providing opportunities for participants to learn from their participation. An easy example is making yourself available after experimental sessions to answer questions and further discuss the study with participants who might want to do so, or inviting them to contact you via email after the fact. Researchers also sometimes offer to send out the results of a study to interested participants once it is complete, but given the realities of academic life it is easy for this to slip down the priority list and be forgotten when the time actually comes to deal with the results. We are moving toward implementing a system that is less susceptible to the vagaries of researchers’ prospective memory, namely identifying projects on the OSF with particular distinctive tags or keywords that are easy to provide participants with either verbally or in printed form.

## Maintaining and storing records

This section will mostly focus on the paper records that should or can be maintained while running on-campus experiments. Considerations pertaining to secure storage of digital data files will be discussed in more detail in the [Data Management](#_Data_Management:_Short) section.

Depending on the experiment and the setting in which it is conducted, the paper records for which you will be responsible may include:

* **Signed consent forms**: Because these contain participant names, they should be securely stored on lab premises whenever possible and treated with particular care when in transit/accidentally taken off premises. They should be retained for 5 years (in case of the very rare event an ethical complaint is filed with the HREB) and then securely shredded.
* **PRPS Record of Participation form** (available [here](https://www.uvic.ca/socialsciences/psychology/research/participants/index.php)): This form should be filled in every session with the names of participants who receive credit or individuals who are marked as excused or unexcused no-shows, along with the date and indication of credit amount/status (e.g., for a half-hour study, you would mark 1 for participants who show up, -1 for unexcused no shows, and 0 or N/A for excused no shows) . Participants who do show up to the experiment should also be asked to fill in their student numbers. These records should also be securely stored and shredded, but can be destroyed as soon as 3 months after the end of term because they are primarily kept for the purposes of settling any credit discrepancies.
* **Anonymous record sheet:** The exact form this takes will vary depending on the experiment, but most studies will require some kind of record of participant numbers, their associated experimental condition/group, testing times/dates, and a place for comments/notes. This sheet provides an easy reference point for the experimenter (e.g., to avoid reusing participant numbers) and a means of implementing quasi-random assignment (e.g., by pre-populating a spreadsheet with participant numbers and randomly shuffled group IDs, and testing individuals in this order). Additionally, the experimenter should use this sheet to note any errors on their part or unexpected events that may have implications for data quality (e.g., if someone does not seem to understand the instructions or is visibly not paying attention). Such events should be described in enough detail that Future You (or the experimenter in charge of the study) can understand it and make an informed decision regarding whether the corresponding participant(s) should be excluded from analysis. Because this record is fully anonymous (providing you take care not to put any sensitive participant information on it, nor note any participant numbers on the Record of Participation sheet discussed above), it does not need to be securely destroyed and can be stored indefinitely. A better option is to create a digital copy of any information from this sheet that may be needed in the future, such as the participant numbers of individuals who were excluded from analyses and the corresponding justification(s), and store this alongside the rest of the data associated with the study.

When running computerized experiments, it is always a good idea to check the file for the first participant (in each condition, if applicable) to make sure everything is recording as expected. Data files should be backed up regularly (options for this discussed in more detail [below](#_Data_Management:_Short)), ideally after every experiment session, and changes should *never* be made to the original files.

# Data Analysis

This is arguably the stage of the scientific process where the greatest number of “researcher degrees of freedom” (see Simmons et al., 2011) come in. Many transparency-related and methodological concerns related to data analysis can be dealt with long before you actually get to this point by preregistering your analysis plans and potentially seeking analytic advice/comments from others. This section will deal with issues related to implementation and ensuring reproducibility.

## Reproducibility

Ideally, the entire data analysis process should be reproducible by anyone given your raw data files, from the initial data cleaning, checking, and inspection all the way up to the final analyses (and even beyond, to the manuscript or research summary). The most straightforward way to implement this is by establishing a well-annotated, script-based workflow, ideally in languages that can be run using programs that are entirely free and open source (e.g., Python, R/RStudio). Depending on your particular background and experience, this may be an overwhelming prospect, and learning how to adapt each step of your workflow may require a substantial time commitment. Lab policy is therefore flexible on this point: you are expected to make as much of your data analysis workflow reproducible as is possible given your current skillset, and to prioritize filling in the gaps as you proceed through your career in the lab.

My own bias is to encourage you to use R or Python. Both are open source, highly flexible, have vast communities that are constantly developing new packages and ways of doing things such that it is almost certain someone has already solved the problems you may need to solve (and can help you if not), and may be helpful to you in various career pursuits. However, the learning curve is not trivial for either, so it may not be feasible for you to jump right in. Below I will outline some general principles related to enhancing reproducibility that may help guide your choice of program or suggest ways you might increase the reproducibility of work done in “imperfect” programs.

**Scripting is better than point-and-click.** Documenting your entire data merging, cleaning, organizing, and analysis process in code makes it possible for others to reproduce and check your work (and for you to remember exactly what you did). Once you have a script you are confident in, adapting it for future projects reduces the possibility of human error that is introduced by steps that are often necessary in point-and-click interfaces such as copying and pasting and having to manually select options anew every time.

Some programs that are mostly used for working with data in a point-and-click fashion also have custom scripting options, such as SPSS’s [syntax editor](https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/base/syntax_editor.html) or Microsoft’s [Visual Basic](https://docs.microsoft.com/en-us/dotnet/visual-basic/), which can be used with Excel. [JASP](https://jasp-stats.org/) does not, as far as I know, have customizable scripting options at this stage, but ensures reproducibility by storing analysis settings in the output file; it is probably the best option reproducibility-wise if you prefer or are more used to point-and-click interfaces.

**Open source is better than proprietary.** In addition to the general benefits of open source programs and languages (better security because more people can access, identify problems with, and modify the source code; less vulnerability to the vicissitudes of the market; greater likelihood of still being available far into the future, as others can take over development and updating if the original developers move on), using them in your analysis workflow will increase the number of people who will be able to reproduce your analyses. This is good for transparency and error-checking, as well as collegial knowledge-sharing (e.g., others may adapt your code for their own purposes) and making research more economically accessible.

JASP, mentioned above, is open source. [PSPP](https://www.gnu.org/software/pspp/) is an open source SPSS-like program that can work with SPSS files and offers syntax options compatible with SPSS’s. There are also open source equivalents to [Excel.](https://www.openoffice.org/product/calc.html)

**Annotate, annotate, annotate!** Regardless of the program you end up using, one of the best things you can do to ensure the reproducibility of your workflow and to benefit your future self is to keep track of everything you do with your data and why. If using R, Python, or other script-based systems, this means commenting or annotating your code to clarify what each section or component is doing and why you have set custom options a certain way. There may be other things it is useful to note, such as instructions on how to adapt particular sections to do something different or use different settings, what to do if a given section results in an error, etc.

I advise always erring on the side of over-commenting; what seems intuitively obvious to you now may not be so in a few weeks, let alone to others who are less familiar with the language or packages you are using.

## Version Control

Version control systems (VCSs) are a means of managing files likely to undergo multiple revisions and maintaining copies of each version in an integrated way. If you have ever ended up with a mess of files named along the lines of “term paper.doc”, “term paper final.doc”, term paper final FINAL.doc”, etc., you are familiar with the gist of version control, but for obvious reasons it is preferable to formalize this process and automate some of the organizational elements.

Formal version control can be useful for things like papers, but its advantages are perhaps especially apparent when it comes to code such as R scripts you may develop data analysis and other purposes, particularly those you are collaborating on with others. In the case of code you will be regularly adding new lines or sections and tweaking things as you discover they don’t work as intended. Inevitably, sometimes you will make changes to your code that unexpectedly make things worse or break the entire process, and you may want to revert to an earlier version; used properly, a VCS makes this kind of thing relatively painless. They also usually allow for some kind of straightforward comparison across versions so you can see exactly what has been added, deleted, or changed, and – in collaborative situations – who has done so.

The OSF has its own built-in [version control system](http://help.osf.io/m/projectfiles/l/524182-file-revisions-and-version-control), so this is one option for incorporating version control into your workflow. However, you will have to remember to manually upload each new version of the file in question to take advantage of this function; it is possible to sync your files with the OSF via a third-party storage provider like Dropbox, but in this case you would be limited to the number of versions maintained by that provider.

An increasingly popular VCS is [Git](https://git-scm.com/), which has an accompanying hosting service called [GitHub](https://github.com/) that can serve as an online repository for your project files and even generate and host [websites](https://pages.github.com/). Although Git primarily has a programming/code-based focus, it can also work with other file types. Git integrates smoothly with RStudio, and the OSF allows you to link projects with GitHub pages, so as you develop your own research workflow you may find Git a useful addition. We are in the early stages of learning it in this lab, but there are a number of online resources and tutorials available (e.g., [this one](http://happygitwithr.com/) by Jenny Bryan with reference to R/RStudio integration specifically) if you are interested.

## Statistical Considerations

We are very fortunate to have several quantitatively skilled people who are also excellent teachers within the department, so I am certain you will develop and sharpen your statistical chops throughout your time at UVic. I will not delve very far into the subject here other than to touch on a few common interpretive pitfalls of null hypothesis significance testing (NHST), address the question of how to determine support for null effects, and strongly recommend [this online course](https://www.coursera.org/learn/statistical-inferences) by Daniel Lakens on improving your statistical inferences, which explains both of the previous points much more eloquently than I am able to.

### Common NHST pitfalls

**Misinterpretation of *p*-values:** These are so common there’s an entire (pretty good) [Wikipedia article](https://en.wikipedia.org/wiki/Misunderstandings_of_p-values) about them. It is easier to start with what they are not: the *p*-value does not reflect the probability the null hypothesis is false, or the probability of the observed pattern of results occurring by chance alone if the null hypothesis is false. *p*-values *are* probabilities, but they are long-run error probabilities: if the true state of the world is that the null hypothesis is true (e.g., there is no difference in variable *x* between populations *A* and *B*), and you re-ran the same study and analyses using samples of the same size some very high number of times (say 1000), you would only obtain an effect equal to or greater than the one you have observed *p*% of the time (so if you get a *p* = 0.04, 40/1000 samples from a population in which the null hypothesis is true would produce effects equal to or greater than the one you have observed).

**Misinterpretation of confidence intervals.** I learned in my first year of graduate school that my understanding of confidence intervals (CIs) was 100% wrong, and have been engaged in a gradual process of becoming slightly less wrong every since. Again, it is useful to start with what they are not, and I will borrow heavily from a paper by Hoekstra, Morey, Rouder, and Wagenmakers (2014). Consider a 95% CI with lower and upper bounds of 0.3 and 0.7, respectively. This CI does *not* indicate there is a 95% chance the true population mean is in the 0.3-0.7 range, nor that you would find means in this range 95% of the time if you repeated your experiment and analyses a great number of times. Really, a 95% CI tells you nothing about the particular range it captures in any one case; instead, if you repeated the study and analyses which generated the CI a great number of times, 95% of CIs of that size would contain the true population mean.

The correct interpretations of p-values and CIs are neither catchy nor intuitive, and some people argue they are fundamentally uninformative or uninteresting and advocate for a full shift away from frequentist approaches/NHST and toward likelihood-based or Bayesian approaches. I subscribe to the more measured perspective of Lakens and others who emphasize that these approaches are all just tools designed for slightly different aspects of the overarching problem of uncertainty in sampling and measurement; none is inherently superior to the others, but one may be more appropriate than another for a particular purpose. Assuming your statistical background is, as with most students of psychology, primarily rooted in NHST, I do encourage you to branch out to other approaches, but will not insist you use one or the other.

### Null effects

You are probably familiar with the idea that standard NHST only allows for two conclusions: you can reject the null hypothesis (if you get a significant result), or fail to reject it (if you get a non-significant result), but you can never *accept* it. But null results can be theoretically and practically important – it can be perfectly reasonable for your primary hypothesis to be that you will find a null effect (e.g., no difference between groups), and even if not it may be useful for you to have more information on exactly how null-ish your results are before making any decisions about methodological changes or abandoning your original hypothesis altogether. There are both frequentist and Bayesian approaches to the problem of quantifying support for the null hypothesis/null effects, and I am sure I’m not aware of all of them, but will introduce you to two here.

#### Equivalence testing

The essence of equivalence testing is that although the NHST/frequentist framework does not allow you to test the hypothesis that a given effect size is *exactly* zero, you can define a range of effect sizes around zero that you consider too small to be worthwhile. For example, you might decide that in the context of your effect of interest, Cohen’s *d* values falling between -0.1 and 0.1 are so small as to be meaningless to you. Equivalence testing then enables you to statistically reject the hypothesis that the effect in question is large enough to be outside these bounds. In other words, the approach does not exactly provide support *for* a null effect, but does provide compelling evidence *against* the possibility the effect is large enough to be of interest (however you have defined it).

There are different kinds of equivalence testing, but one that is commonly used is the two one-sided test (TOST) procedure, which Daniël Lakens has laudably attempted to make accessible to psychologists via tutorials (Lakens, 2017) and an accompanying R package, [TOSTER](https://CRAN.R-project.org/package=TOSTER).

#### Bayes factors

The intuitive way Bayesian approaches allow you to quantify relative support for different models/hypotheses, including those in which there is no effect, is a major advantage. This takes the form of the Bayes factor, which is a likelihood ratio conveying the relative likelihood of the observed data under one hypothesis versus another. E.g., you might have a Bayes factor of 5 indicating the data are 5x more likely under the alternative hypothesis than the null hypothesis (and the inverse will of course give you the ratio corresponding to the opposite). These analyses can be conducted in JASP or using various R packages (e.g., [BayesFactor](https://richarddmorey.github.io/BayesFactor/)).

# Disseminating and Communicating Results

Once you have finished a study and analyzed your data, the next step will be to do something with your results. This might be presenting at a conference, publishing in a journal, sharing your results in some informal way, or some combination of the above.

One recommendation I think is good practice (although I regularly forget to do it myself) is writing a concise, 1-3 page summary of your study purpose, methods, and results while all are fresh in your mind. You may sometimes find yourself moving onto other projects very quickly and – memory being fallible as we know it is – forget many of the details when the time comes to publish or revisit your work for other purposes. Although you should always be able to track down these details if you follow the other advice in this manual, having a brief summarized version may save you time, and also provides an easy way to share your results with others if they are not yet or never published formally.

## Reproducible Documents

Although reproducibility initiatives tend to focus on methods and analyses, efforts to go beyond this and generate fully reproducible theses/dissertations, journal publications, and even conference posters are also proliferating. [RMarkdown](https://rmarkdown.rstudio.com/), for example, is a tool for R/RStudio that allows you to write and format reports in a language called markdown, and embed all the code for your statistical analyses, figures/tables, etc. within the same document. You end up with some (fairly ugly-looking) code that, when “knit”, will run all of your analyses, create your figures and tables, and so on, and embed the results within the text of a nicely formatted pdf, Word, or html file. This saves you having to copy and paste or manually rewrite your statistical results, which is a potential point of failure, and can make editing much less of a hassle – if you decide to, say, remove an outlier’s data from all of your analyses, you will likely just have to change a single line of code and re-generate the document instead of tracking down all the references to those data in text, figures, and tables, and manually changing them. Although the initial learning curve can be steep, it will save you time in the long run

A related system (and one that can be integrated with RMarkdown) I am less familiar with but may also be worth looking into is [LaTeX](https://www.latex-project.org/), which seems to have particular advantages for work that requires embedding mathematical formulas.

## Data Sharing

Regardless of what comes of your results, you should plan to share your data as publicly as possible given the applicable ethical constraints. Although it may make sense to share your data in multiple forms/at multiple levels of analysis, an ideal to work toward is being able to share your data exactly as they were collected (see e.g., Jeff Rouder's [2016] notion of "born-open data") – this may not always be possible or advisable given the nature of the data you are working with, but is a means of maximizing transparency and working toward a fully reproducible research workflow. The [Tri-council policy on data management](http://www.science.gc.ca/eic/site/063.nsf/eng/h_547652FB.html), including guidelines and expectations for data sharing, is still in the draft stage, but it seems likely the councils will gradually move toward requiring (1) all data and code associated with journal publications and other formal “research outputs” arising from grant-funded work be stored in an institutional or recognized third party digital repository, and possibly (2) that formal data management plans (DMPs) be submitted as part of grant applications. Using digital repositories and developing DMPs are practices worth adopting even if they are not mandated; I will discuss the former here, and the latter in the final section.

There are a wide range of digital repositories available – some are generalist in the sense of being multidisciplinary and/or not limited to data specifically, such as the OSF, [Figshare](https://figshare.com/), and [Zenodo](https://zenodo.org/), while others are specific to the social sciences or Psychology or dedicated exclusively to research data, not other parts of the research workflow. The service that makes the most sense for you will depend on the sensitivity and format of your data, as well as your purpose in sharing them (e.g., some journals or grant agencies may only recognize certain services or require you submit to their own). These services differ with respect to features such as where data are hosted, privacy and protected access options (e.g., only allowing verified researchers to access your data), documentation and metadata requirements, whether submissions are open or curated, and the formats for which they are specialized. Eventually it may make sense to standardize which services we use as a lab, but for the time being a few examples include:

* The UVic instance of [Dataverse](https://dataverse.scholarsportal.info/dataverse/uvic) (more information about the international project [here](https://dataverse.org/)), which does not currently have third party protected access options but does allow for various levels of restricted access;
* [ICPSR](https://www.icpsr.umich.edu/icpsrweb/), a large international repository for social science data that offers multiple levels of [restricted access,](https://www.icpsr.umich.edu/icpsrweb/content/ICPSR/access/restricted/) including an option to release data only to researchers whose applications are verified by ICPSR staff. ICPSR is a paid service, but UVic currently has an institutional subscription.
* [Databrary](https://nyu.databrary.org/), which is specialized for video data and also allows for multiple levels of restricted access.

An exciting recent development I will continue to keep an eye on is the [Federated Research Data Repository](https://www.frdr.ca), which has not yet fully launched but is dedicated to hosting research data from Canadian institutions.

### Privacy and confidentiality concerns

The best case scenario in terms of methodological transparency is to publicly share enough information that anyone could reproduce your entire analysis, from start to finish, including merging, cleaning, exclusions, etc. Ideally, this means sharing the original raw, unaltered data files. However, this ideal must be balanced against privacy and confidentiality concerns. The small decrease in full start-to-finish reproducibility caused by sharing data files with some information removed and/or changed is far preferable to the risk anyone who participated in your study with the assurance their confidentiality would be maintained being identified.

The TCPS distinguishes between 5 categories of data:

1. **Directly identifying information**, which includes identifiers like names, social insurance numbers, phone numbers, etc.;
2. **Indirectly identifying information**, which can be reasonably expected to identify an individual in combination (e.g., date of birth, city, distinctive or rare personal characteristics);
3. **Coded information**, which does not include any direct identifiers but can still be linked to individuals given access to some code;
4. **Anonymized information**, which contains no direct identifiers (or has been stripped of any such information), cannot be linked to individuals via code, and does not contain sufficient indirect information to identify a particular individual; and
5. **Anonymous information**, which is and has always been free of any identifying information.

These categories are not always clear cut; technological advances have already substantially blurred the lines between what information can and cannot be considered potentially identifying, and we have no idea what might be possible within a few years let alone decades. As we progress toward increasing reproducibility, transparency, and collaboration by sharing data and analysis code more widely, our ethical duty to do everything we can to protect participant confidentiality must be kept in mind. Some of the implications here are obvious; data files containing participant names should of course not be shared publicly nor with anyone not directly associated with the research project, and things like highly personal and/or sensitive stories or responses to questions, audio and video recordings of experiment sessions, etc. should not be shared unless this was approved by the HREB and participants consented to this use of their data. Similar considerations apply to things such as stimulus sets comprising photos or videos of people; while making stimulus sets publicly available is desirable, this should only be done with the express consent of the depicted individuals, and the exact nature of the original indication of consent must be kept in mind when using older stimulus sets or those obtained from other labs. Someone consenting to their photo being shared in 2005 likely did so with a very different understanding of who was likely to access it and the potential uses it could be put to than someone providing such consent today.

Things get particularly sticky when it comes to the possibility of identifying an individual on the basis of some combination of indirectly identifying information and/or metadata. Essentially, it can be assumed that the more data we collect, the more this possibility increases; this is especially obvious in the context of data such as birthdate, location, and various demographic information, but as machine learning grows more complex and powerful these possibilities may extend to information we don’t currently see as risky.

For our purposes, the most straightforward way to protect participant privacy and confidentiality in most of the research we do is by fully anonymizing the data that are stored. Because the vast majority of our data collection is computerized, this is usually easily implemented by the use of participant numbers or IDs; these serve to individuate participants, but providing (1) there is no other information in the file that could reasonably be used to identify the participant (e.g., extensive demographic information, responses to personal questions, sufficiently long responses to reveal a distinctive writing style, etc.) and (2) these numbers or IDs are not stored in connection with the participant’s name anywhere else, the risk the participant will ever be identified on the basis of, let alone harmed by, public release of their data file drops near zero the second they leave the experimental session.

There are a number of things we can do to keep this risk as low as possible. One, of course, is to limit the collection of personal information to only what is directly relevant to your research question. If you don’t care about participant gender or have any reason to believe it is important to what you are studying, maybe you don’t need to ask participants for this information. That said, there are valid arguments to be made for collecting more information than you need. Exploratory analyses of variables secondary to your research aims can reveal patterns worthy of further investigation, your *a priori* sense of which variables you *think* are (or are not) important to what your are studying need not be correct, and you – or other researchers studying or trying to meta-analyze similar questions – may come to lament the lack of certain data later. You can also take particular care to make sure participant names and anonymized IDs are not stored together somewhere outside the original data file, and train RAs to do the same.

**Multi-session experiments**: There is some research for which it is not feasible to store data in a fully anonymized form, as the researcher needs some way of tracking the same individuals over time or across sessions. Probably the most common case of this in our lab will be experiments involving a memory test after some delay; if the study material is not the same for everyone, we will need some way to figure out who’s who in the test data, or at the very least which condition they were in at study.

Probably the best approach in terms of balancing security and ease is to still use participant numbers or IDs in the data files themselves, but keep a record elsewhere matching these IDs with names. This record should be securely stored, accessible only to those who need it, and securely destroyed as soon as it is no longer needed. Maintaining this record on paper is arguably more secure than keeping a digital record as it is easier to control access to and to permanently destroy a physical copy, but practical circumstances may outweigh this advantage - e.g., if a project involves multiple experimenters and/or testing locations, a paper record may be not only inconvenient but also at a higher risk of being lost. Providing reasonable measures are taken to ensure security, either format is fine.

**Anonymizing after the fact:** Although I have not worked with them myself, there are a number of means by which you can anonymize data using R, including the methods described [here](https://jangorecki.github.io/blog/2014-11-07/Data-Anonymization-in-R.html) and the packages [fakeR](https://cran.r-project.org/web/packages/fakeR/vignettes/my-vignette.html) and [anonymizer](https://CRAN.R-project.org/package=anonymizer)**.**

# Data Management: Short and Long Term

Last but not least, it is important research data – including “data” in the usual sense of participant data and results, but also other critical research materials including stimuli, experiment files, and important references – be managed with an eye toward making it easy not only for you to find the files you need tomorrow or next week, but also for you or others unfamiliar with the project to find what might be needed or helpful 5 or 10 years down the line. It’s a bit morbid, but I try to regularly ask myself “if I died tomorrow, would someone else given access to my files be able to take over this project?” The usual answer is “not without some amount of pain and/or cursing my spirit”, but I am taking steps to improve the situation, and the earlier in your research career you start adopting good data management practices the easier it will be in the long run. I have already touched on some topics relevant to data management above, including [data sharing](#_Data_Sharing), [reproducibility](#_Reproducibility), and [version control](#_Version_Control).

## Data Management Plans (DMPs)

More and more funding agencies are moving toward requiring data management plans as part of applications, and it is a good habit to get into regardless. DMPs generally include high-level information such as why and how you are collecting the data in question; information about the data themselves (what kind, which variables, what file format, how much, how sensitive/confidential, etc.); how you will structure, organize, and store your files; plans for sharing and longer term preservation, and the kinds of documentation/metadata that may be required for your data to feasibly be used by others in the near or far future. The UK-based Digital Curation Centre offers a useful checklist [here](http://www.dcc.ac.uk/resources/data-management-plans) that may be helpful in introducing the kinds of questions you should start thinking about early in your research career and before data collection begins on a given project. Starting with a standardized tool that guides you through the process, such as the Portage Network’s [DMP Assistant](https://assistant.portagenetwork.ca/), will also probably make your life easier.

## File Management and Storage

As someone who still occasionally finds myself with ~10+ files on my desktop named various iterations of “asdglkj”, I can assure you you will thank yourself later for establishing a good file management system now. Ideally, your system should be intuitive and well-documented enough that others who may eventually need access to some of your files will also thank you.

### Backups

All digital lab data – experiment files, stimuli, raw data, analysis scripts, etc. – should be stored in at least 3 separate locations: for example, your office computer, your home computer, and an external hard drive. If using one of the office or testing room computers as one of your storage locations, you should use Tivoli Storage Manager to regularly back up your machine, which will back your data up on a secure server (and can serve as one of your three sources if used regularly enough).

By far the most convenient way to ensure data are stored in 3 locations is by using a cloud storage service capable of syncing data across multiple devices (e.g., Dropbox or Google Drive). Because many of these services store the data locally on each synced machine as well as “in the cloud” (i.e., on the company’s servers, wherever they are located), this is a near-effortless way to work with your files in various locations and know you will always have recent copies elsewhere in case one link in the chain breaks down without having to worry about manually transferring files from place to place on a regular basis. Ultimately, unless you are tech-savvy enough to securely encrypt files on your end before uploading them to such services, you should assume the contents of any files stored in this way are fully accessible by higher-level employees at these companies and in many cases any governments with jurisdiction over the server location. Unlikely as it may be that anyone would ever actually do so, if you are dealing with sensitive data collected at UVic or elsewhere – or any data, sensitive or otherwise, participants have been assured would be securely/locally stored – [you should not store it in the cloud](https://www.uvic.ca/systems/support/informationsecurity/datasecurity/datasec-cloudstorage.php). This also applies to any other university or student data you may be working with, e.g., as part of your duties as a Teaching Assistant. For files you are not concerned about keeping private (e.g., your own papers), and participant data you have consent to share publicly (e.g., via the Open Science Framework), it is fine to use such services.

Lab hard drive

After an experiment has wrapped up, you will be expected to transfer all relevant data onto a shared external hard drive following the file naming system outlined below. At a minimum, this should include:

1. Any files (including stimuli) or scripts required to run the experiment, and for online studies for which this is not possible a detailed enough description that someone could reproduce the full procedure;
2. All completely raw, unaltered data files, including those of participants who were ultimately excluded from data analysis (if data are not already in .txt, .csv, .xml, or .html format, please convert to one of these formats and include these copies as well – this is primarily a “future-proofing” mechanism, as files in non-proprietary, plain text formats are more likely to be immune to obsolescence);
3. Any files (e.g., merged data, or data at other stages of processing that could not be automated) or scripts required to reproduce your analyses, as applicable;
4. Sufficiently robust documentation regarding variable and file names – e.g., in the form of a “readme” file – that someone completely new to the project could navigate your files and understand what’s going on within them.

Additional information you might include are things like the final summary recommended [above](#_Disseminating_and_Communicating), more extensive results such as informative figures and tables you have generated, background information, and key references justifying or otherwise relevant to the project.

The archive you create for this purpose should also adhere to the backup rules above; in this case, the hard drive would of course count as one of the three locations, but should not be the only place these archives are stored.

### Guidelines for naming & organizing files

I will not prescribe a particular approach to file naming and organization for your personal use; as long as you develop a system that works for you, that’s fine. That said, I will want the final archived version of each project formatted according to these guidelines, and you may find it useful to adapt some of them in your own system.

#### General advice

Much of the following has been adapted from a [slide presentation](https://speakerdeck.com/jennybc/how-to-name-files) given by Jenny Bryan at a Reproducible Science workshop in 2015.

* In naming both files and variables, use hyphens and/or underscores rather than spaces to separate words. This preserves human readability while vastly improving machine readability.
* File and variable names should follow the general principle of being “descriptive but brief”.
* Some recommend naming files themselves fairly generically and relying on folder structure to identify what the file is. I prefer some redundancy here so that files will be identifiable even if they somehow end up stored out of context (maybe unlikely, but better safe than sorry). You don’t have to reproduce the whole folder structure in the name, but at least include the project name/number if applicable, year, and if it’s something you have authored or generated (e.g., script, or highly processed data), and your initials.
* You may wish to include a version number for files you are likely to change substantially but may also need to revisit old versions of, but using a formal version control system can obviate the need for this.
* You can’t always rely on “date created/modified” metadata to determine the history of a given file as they may get moved around, and this information sometimes gets overwritten in the process. So for files for which you anticipate this being important information, it may be worth including a date in the filename. If you do so, make sure to stick with a uniform format across files, preferably the near-universal YYYY-MM-DD.

##### Lab hard drive format

The following is a sample folder structure of the sort I would want you to store on the lab hard drive (albeit ultimately zipped or otherwise compressed), with each additional bullet level representing a new folder level. This specific example pertains to a project with multiple paths/sub-projects conducted over multiple years.

metamemory-judgments\_08-06-391b [*parent project name with HREB identifier*]

* materials-based-bias-effect [*child project name*]
  + 2018-2019 [*school year, as this is the usual data collection schedule*]
    - svdw\_naming [*individual study/experiment name*]
      * readme.txt [*file containing necessary documentation*]
      * data
        + processed
        + raw
      * experiment
        + stimuli
      * preregistration
      * references
      * results

## Reference Management

Although it used to be the norm for students and researchers to end up with huge piles of printed or photocopied research articles at the end of a project, the more common modern equivalent is the folder full of pdfs with unintelligible names. Better-organized folks than myself may be able to keep on top of this manually in a way that works for them, but in my experience it can spiral out of control very quickly, especially when you are trying to juggle literature associated with multiple projects and/or courses at the same time. Even if you think your current system is working, I highly recommend trying out a reference manager or two early in your career. Most of these tools have some capacity to extract important metadata such as article titles, journal and author names, and publication dates from pdfs or other file formats; organize and rename your files in a meaningful way (e.g., into folders named by first author, and files named “[Article title (Year)].pdf”; some combination of in-program folder- and or tag-based system for sorting articles by project, class, or topic(s); plugins that allow for easy in-text citations and bibliography generation in programs like Microsoft Word, and cloud-based syncing systems that allow you to access articles across different devices.

UVic Libraries provides a [guide](http://libguides.uvic.ca/citations) to some of the most commonly used reference managers, including a [comparison](http://libguides.uvic.ca/c.php?g=256836&p=1714605) of various features that might be of importance to you. I can vouch for the ease of use and high-quality metadata extraction of both Zotero and Mendeley, but in terms of open science and data preservation Zotero, which is open source, is a better bet. Mendeley is currently owned by the academic publishing giant Elsevier, and thus comes with some [privacy](https://www.elsevier.com/legal/privacy-policy) and [ethical](http://thecostofknowledge.com/) concerns, as well as some risk of losing control over your library/annotations or becoming locked in to the platform given its proprietary nature. York University’s library provides a [detailed comparison](https://www.library.yorku.ca/web/research-learn/citations/zotero-vs-mendeley-comparison/) of various features and pros and cons of the two programs.

Other points of caution, regardless of which program you use, are

* + - * 1. Be aware of your privacy settings (e.g., whether the program you are using adds citation records into a searchable public library by default, and how to turn this function off or exclude certain documents, such as personal communications or draft manuscripts you have been sent privately)
        2. If you have any references you want to keep private and are using a program with sync functionality, look into where its servers are located (neither Zotero nor Mendley currently offer the option to store data within Canada).
        3. Metadata extraction is not perfect, so you will still want to glance through your document list anytime you add a large number of papers to make sure everything looks okay. Similarly, you should double-check any bibliographies you generate for the same purpose and to ensure APA compliance.

# References

*I have consulted and been inspired by so many formal and informal sources in the course of developing this document that I am sure to miss crediting some who deserve credit, but would be remiss not to note lab manuals created by* [*Maryam Ali*](https://github.com/alylab/labmanual)*,* [*John Paul Minda*](https://osf.io/8b6kj/)*, and* [*Maureen Ritchey*](https://github.com/memobc/memolab-manual)*, from which I gained specific inspiration.*

Button, K. S., Ioannidis, J. P. a, Mokrysz, C., Nosek, B. a, Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews: Neuroscience*, *14*(5), 365–76.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2014). *Tri-Council Policy Statement: Ethical conduct for research involving humans*. http://doi.org/1

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *The Journal of Abnormal and Social Psychology*, *65*(3), 145–153. http://doi.org/10.1037/h0045186

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716-aac4716. http://doi.org/10.1126/science.aac4716

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. http://doi.org/10.3758/s13428-014-0458-y

Difallah, D., Filatova, E., & Ipeirotis, P. (2018). Demographics and Dynamics of Mechanical Turk Workers. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining - WSDM ’18*, (August), 135–143. http://doi.org/10.1145/3159652.3159661

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., & Bigham, J. P. (2018). A Data-Driven Analysis of Workers’ Earnings on Amazon Mechanical Turk. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18*, 1–14. http://doi.org/10.1145/3173574.3174023

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. http://doi.org/10.1017/S0140525X0999152X

Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E. J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin and Review*, *21*(5), 1157–1164. http://doi.org/10.3758/s13423-013-0572-3

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*.

Ipeirotis, P. G. (2010). Analyzing the Amazon Mechanical Turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students*, *17*(2), 16. http://doi.org/10.1145/1869086.1869094

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*(3), 196–217. http://doi.org/10.1207/s15327957pspr0203\_4

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*, *44*(7), 701–710. http://doi.org/10.1002/ejsp.2023

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. http://doi.org/10.1177/1948550617697177

Lindsay, D. S., Simons, D. J., & Lilienfeld, S. O. (2016). Research Preregistration 101. *Observer*.

Nosek, B. A., & Bar-Anan, Y. (2012). Scientific Utopia: I. Opening Scientific Communication. *Psychological Inquiry*, *23*(3), 217–243. http://doi.org/10.1080/1047840X.2012.692215

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, *7*(6), 615–631. http://doi.org/10.1177/1745691612459058

Pittman, M., & Sheehan, K. (2016). Amazon???s mechanical turk a digital sweatshop? Transparency and accountability in crowdsourced online research. *Journal of Media Ethics: Exploring Questions of Media Morality*, *31*(4), 260–262. http://doi.org/10.1080/23736992.2016.1228811

Poushter, J. (2016). *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies*. *Pew Research Center*. http://doi.org/10.1017/CBO9781107415324.004

Ross, J., Zaldivar, A., Irani, L., & Tomlinson, B. (2010). Who are the Turkers? Worker Demographics in Amazon Mechanical Turk. *Chi Ea 2010*, (July 2016), 2863–2872. http://doi.org/10.1145/1753846.1753873

Rouder, J. N. (2014). Optional stopping: no problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301–308. http://doi.org/10.3758/s13423-014-0595-4

Rouder, J. N. (2016). The what, why, and how of born-open data. *Behavior Research Methods*, *48*(3), 1062–1069. http://doi.org/10.3758/s13428-015-0630-z

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. http://doi.org/10.1177/0956797611417632

Smith, A. (2016). *Gig Work, Online Selling and Home Sharing*.

Society for the Improvement of Psychological Science. (n.d.). Retrieved June 19, 2018, from https://improvingpsych.org/

van ’t Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*. http://doi.org/10.1016/j.jesp.2016.03.004

Vankov, I., Bowers, J., & Munafò, M. R. (2014). On the persistence of low power in psychological science. *Quarterly Journal of Experimental Psychology*, *67*(5), 1037–1040. http://doi.org/10.1080/17470218.2014.885986

Velicer, W. F., Cumming, G., Fava, J. L., Rossi, J. S., Prochaska, J. O., & Johnson, J. (2008). Theory Testing Using Quantitative Predictions of Effect Size. *Applied Psychology*, *57*(4), 589–608. http://doi.org/10.1111/j.1464-0597.2008.00348.x

# Appendix

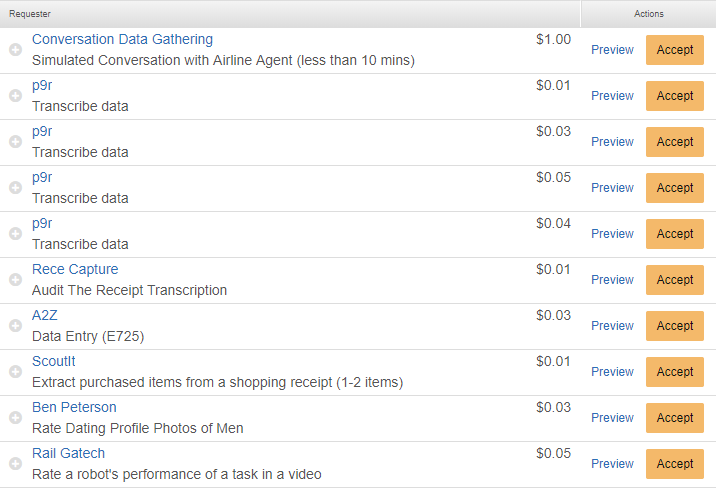
**Some Extra Notes on Ethical Concerns Regarding Mechanical Turk**

As a researcher, it is surprisingly easy to conduct studies on platforms like MTurk without understanding much about how the platform works or who uses it. Anecdotally, the first time I posted an experiment on MTurk, the vast majority of my time and thought went into the details of setting up the task itself and figuring out logistics such as how to ensure the same person could not participate more than once. Once I had posted it, the process felt almost like magic; for probably well under $100 total, I reached my intended sample size within an afternoon. It was a fairly short and simple task and participants’ comments were mostly positive, so I assured myself in the same way I suspect a lot of academics using the platform do – many people are eager to contribute to scientific research with or without pay, the task was probably more interesting and paid better than most work on MTurk, etc.

Most research labs are not wealthy and those working within them tend to be very busy, so the ability to collect data from hundreds of participants in a few hours for a minimal cost has straightforward appeal. And it is true that many people are eager to contribute to science on a strictly volunteer basis – the recent success of so-called citizen science initiatives (see e.g., the [Government of Canada’s Citizen Science Portal](http://science.gc.ca/eic/site/063.nsf/eng/h_97169.html), [Zooniverse](https://www.zooniverse.org)), which allow people to participate in research in various ways depending on their skills and interests (from transcribing century-old ship logs to identifying penguins in Antarctic webcam stills) is just one demonstration of this. But having learned a bit more about the realities of MTurk’s inner workings and the experience of people who use it (Workers), I am much less comfortable with my original justifications for using it.

I will first outline a bit about how MTurk works on the researcher (Requester, in MTurk jargon) side. You create an experiment or survey; if it requires anything complex, such as conditional jumps (e.g., if participant answers “Yes” to a question, skip the next block of questions) or careful timing, you will probably need to use an external website (e.g. Qualtrics). Cognizant of potential quality control issues with online data, you might take precautions such as including attention checks to screen out bots or people who are not paying attention. Once the experiment is ready, you create a project on MTurk. At this stage you will decide how much you want to pay each participant (Worker), how many unique Workers you want to complete the experiment, how long they will have to complete the experiment after starting, and how long you want the project to stay up. Additionally, as a further quality control measure, you might include participant requirements; academic requesters often restrict participation to Workers with very high HIT approval rates (99% is not uncommon) and/or who have been assigned the “Masters” qualification. Regardless of your decisions, you are very likely to reach your intended sample size within a few days at most. You are only required to pay for satisfactory work, so you might choose to reject HITs from participants who failed the attention check.

Things look much different from the workers’ perspective. You start out with a 10 day probationary period, during which you are limited to a maximum of 100 tasks (Human Intelligence Tasks, or HITs) per day. Because most higher-paying tasks will have restrictions that exclude you (e.g. requiring at least 1000 completed HITs), your options are likely to look something like this:



And because so many workers are competing for HITs, even the ones paying $0.01 will be gone within seconds; to make more than a few dollars per day during this period you will have to put in additional time installing scripts and extensions that check for and accept new tasks automatically. Although requesters are supposed to provide reasonable estimates of how long a task will take so you can judge whether the payment is worth your time, Amazon has little incentive to enforce this, so identifying bad actors to avoid finding yourself 10 minutes into a task that only pays 2 cents requires further time investment – e.g., checking external review sites that allow workers to rate requesters, or installing additional extensions that will bring up this kind of information on the MTurk page. You are also very vulnerable to having your work rejected at this point – not only will you not be paid for rejected work, but even one rejection can have a dramatic impact on your approval %, which will need to be close to 100% to be eligible for much high-paying work later on. Requesters are supposed to be fair and reasonable in making these decisions, but standards are often opaque and enforcement, again, is essentially non-existent, so if you feel your work has been rejected unfairly it’s on you to contact the requester.

Eventually, if you successfully complete at least 1000 HITs and maintain a 99%+ acceptance rate, you might receive the elusive “Masters” qualification that academic researchers often use as a quality control restriction. According to Amazon:

**A Master Worker is a top Worker of the MTurk marketplace that has been granted the Mechanical Turk Masters Qualification. These Workers have consistently demonstrated a high degree of success in performing a wide range of HITs across a large number of Requesters. We leverage statistical models that analyze all Workers based on several Requester-provided and marketplace data points to make that determination. Some of the key categories of data that are considered to be granted and maintain the Masters Qualification include the Worker's ability to consistently submit high-quality results (as indicated by Requester approval rates and other related factors), marketplace tenure, and variety of work performed. Master Workers have access to work that requires a Masters Qualification.**

No one knows the explicit criteria for bestowing this qualification, how many people have it, or how often it is assigned.

Then, of course, there is the issue of wages; although academic researchers are not usually profiting directly from the data they collect via MTurk, the majority of requesters using the platform are benefiting greatly from the abundance of demand for poorly paid work. Mean/median earnings on MTurk have been estimated at $2-$3/hr (Hara et al., 2018; Pittman & Sheehan, 2016), well below minimum wage (and even farther below estimated living wage) in the US, which dominates the MTurk workforce. This is allowed because MTurk workers are considered “independent contractors” instead of “employees”; Amazon itself is merely considered the “payment processor”, and thus has little responsibility toward those who use the platform. Many workers use platforms like MTurk for fun, or just to earn some extra spending money in their spare time; this is probably the type of participant most researchers using MTurk have in mind. But a nontrivial proportion rely on MTurk and/or other similar platforms as a primary or even sole source of income[[4]](#footnote-4). Although many individual tasks on MTurk pay well over minimum wage, it is difficult if not impossible to make anything close to this working full time on the platform. The majority of tasks are low paid, menial, and alienating – e.g., $0.01 to transcribe a single line on a receipt, $0.10 to transcribe a full receipt. Many higher paid tasks, such as academic surveys and experiments, include restrictions like “must have completed 1000 HITs”. Thus, even if you are personally paying participants well, it is essentially guaranteed they have already spent an enormous amount of time on poorly paid, tedious tasks and dealt with substantial unfair treatment – such as arbitrary rejection of work – to get to the point of even being eligible to participate in your study. This reality is why, although I am sympathetic to concerns that extremely generous compensation for research participation can constitute coercion, I am unconvinced by this argument against paying at least the minimum wage on MTurk specifically – no, research participation is not employment, and scientific researchers are not (usually!) greedy bosses, but MTurk is explicitly focused around the language of work and premised on a fundamentally exploitative model.

1. At the time of writing (mid-2018), Crowdflower has recently rebranded as Figure Eight and now seems to be more specifically targeted toward tasks designed to train artificial intelligence (i.e., things algorithms still struggle with such as image transcription and sentiment analysis). While it is still possible to conduct more typical psychological experiments using the platform, even the company itself [urges caution](https://success.figure-eight.com/hc/en-us/articles/201855969-Guide-To-Running-Surveys) given issues with quality control. [↑](#footnote-ref-1)
2. An admittedly arbitrary reference point, but most MTurk workers are American (Difallah, Filatova, & Ipeirotis, 2018; Ross, Zaldivar, Irani, & Tomlinson, 2010; see also [the MTurk Tracker](http://demographics.mturk-tracker.com/%23/countries/all), Ipeirotis, 2010). Some have raised concerns that such guidelines stand in opposition to the ethical mandate to avoid “undue compensation” and thus may reach the threshold of being coercive. The HREB will be the ultimate arbiter on whether this is the case given the amount you have proposed and the population you are trying to reach, but I do not find it a convincing argument in the vast majority of cases. [↑](#footnote-ref-2)
3. Or considering alternative approaches to strict *a priori* sample size planning, such as optional stopping if you plan to use Bayesian methods (Rouder, 2014) or analogous sequential analytic approaches in the NHST context (Lakens, 2014). I am less familiar with these approaches but they seem to offer great advantages in terms of data collection efficiency (i.e., not collecting data from more participants than is really necessary), so do hope to move toward implementing them. [↑](#footnote-ref-3)
4. It is difficult to get an exact handle on these numbers, but for example: a 2010 survey of 573 workers found roughly 1/3 were unemployed, and 18% reported “sometimes or always relying on MTurk to make basic ends meet” (Ross, Zaldivar, Irani, & Tomlinson, 2010, p. 2), and a 2016 Pew Research Survey found that 56% of US adults who indicated having used online job platforms (not limited to MTurk) within the last year reported that income was “essential or important” (rather than “nice to have”; Smith, 2016). [↑](#footnote-ref-4)