

## UNDERGRADUATE RESEARCH INDEPENDENT STUDY CRASH COURSE

### Christine Tan, Craighead Research Group, 23 June 2009

*I wrote this as a set of guidelines for my first undergraduate student, when she started research with me. This set of notes is constantly being improved. Please help contribute to making it better for future generations of undergraduates by emailing me: [cpt7@cornell.edu](mailto:cpt7@cornell.edu) should you spot any errors or have any suggestions. Thank you.*

#### 1. LAB SAFETY

Laboratory safety is and should be everyone's number one priority: You leave the lab with as many digits and limbs as when you entered. Learning how to work safely on your own, as well as in a way that does not pose a hazard to your colleagues in your surroundings, are the most important lessons to learn in your undergraduate research.

Always wear the appropriate safety protection when working in the lab. Try to avoid working at odd hours in the lab when there is no graduate student or staff members around, in case you have an accident and there is no one to help you. If you really have to work after hours, always make sure that you are not doing a hazardous activity and know how to get help; if possible, *e.g.* when you are working in the machine shop alone, carry the safety alarm key-tag with you. At very least, know the telephone number of your mentor/supervisor or the lab safety manager so that you can call them if anything happens.

**When in doubt, always ask!** Do not attempt to operate equipment you have not been trained on or try to use chemicals that you are unfamiliar with. Always read the manuals, material safety datasheets (MSDS) before working with something for the first time. Even for something that seems apparently straightforward – *e.g.* pipette suction bulb – people have injured themselves with it.

#### 2. MEASUREMENTS

*“When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be.”* – Lord Kelvin (1883)

While Lord Kelvin's quote could be construed as a somewhat extreme view, it does highlight the importance of quantitative measurements in advancing science. As a researcher, your goals are firstly, to take measurements in the most accurate manner possible, with all the limitations of equipment, systematic errors *etc.*, and secondly, to figure out how reproducible your measurement is. Without reproducibility, no one would be able to repeat what you have done, so how can anyone trust your data?

The first thing to learn is the source of errors/uncertainties. There are generally two types – *systematic* and *random* errors, which you should be familiar with. You cannot do anything to correct for random errors (*e.g.* due to equipment fluctuations, the experimentalist's subjective interpretation *etc.*) and they are inherent in a set of measurements. However, systematic errors

(e.g. calibration offsets) can be detected through careful experimentation and corrected for in your measurements.

Taking multiple measurements helps reduce uncertainties. A good scientist never takes a measurement only once and then makes a conclusion out of it. Data really cannot be interpreted unless the associated uncertainties are understood! Following this line of thought, try to see how many scientific journal articles (or your friends' data) you can spot out there that have no error bars (standard deviations) in their graphs – you might be surprised!

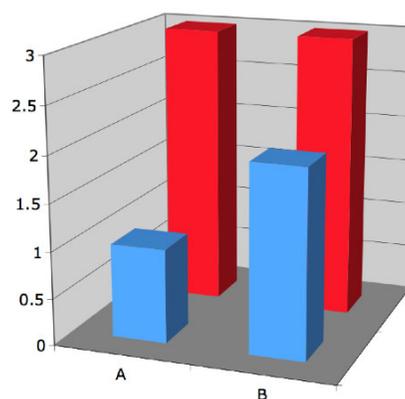
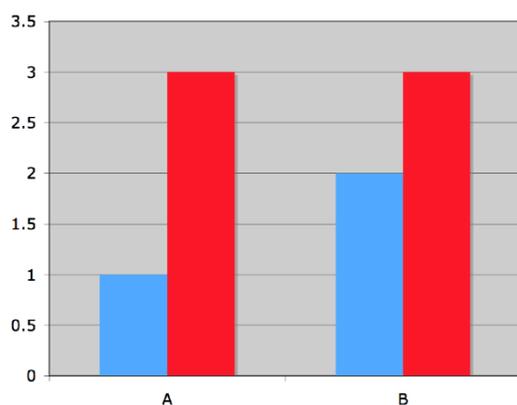
And finally, always keep a laboratory notebook! Even if you are a superhuman with impeccable memory, you always want to write down what you have done – the necessary precautions, any errors you made during your measurements, your train of thought as to why this measurement was important to record to begin with *etc.* The idea is that someone could then follow the notes in your notebook to repeat the exact experiment that you did and get similar results.

### 3. GRAPHICAL REPRESENTATION OF DATA

In an experiment, the independent variable is the parameter fixed by you, or the parameter you know the value of. The dependent variable is the observed parameter, the parameter that you do not know and are measuring in your experiment. Always plot the independent variable in the x-axis and the dependent variable in the y-axis – this is standard convention and furthermore, helps you later when you are trying to do a linear fit of  $y$  vs.  $x$ . An example is shown here:



The rule of thumb for graphical representation of data is simplicity. The eye is easily distracted by irrelevant information. Do not put in extra colours or add shadows to data points *etc.* that are unnecessary to get across what you are trying to say. Let me illustrate my point here. Consider here two graphs which is supposed to present the same data:



Even though the 3D graph on the right might seem more “aesthetically appealing”, the shadows and three-dimensional nature of the drawings add nothing to the value of the graph. In fact, these extra features take away the meaning of what we have otherwise clearly conveyed in the 2D graph on the left. Notice how you could very clearly tell what is the value of the tallest bars – which is 3 – in the 2D graph, but lose this information in the 3D graph such that you would even be tempted to say the tallest bars are less than 3 (an optical illusion)!

The visual representation of data is an art form. Always keep this question in mind: *how can I convey the meaning of my data in the most minimalist fashion?* An excellent reference and rigorous treatise on this subject is the book entitled, *The Visual representation of Quantitative Data*, by Edward R. Tufte [1]. In this picture book, you will find many examples good and bad ways of presenting data.

#### 4. STATISTICAL MEASURES

By this time in your undergraduate career, you would have been exposed to some formal training statistics. Here, I recapitulate some of the fundamental statistical measures and their definitions (mathematical equations that we sometimes take for granted simply because we can pull up the functions in Microsoft Excel).

A word of caution to the wise: Excel usually rounds values during the process of running the calculations, so sometimes you can get vastly different values of standard deviation compared to doing the calculation by hand (manually), or using another dedicated scientific graphing/data analysis software such as Igor, Origin or Kaleidagraph. Moral of the story: if you are working with data requiring great accuracy/computing precision (*e.g.* in the field of analytical chemistry/pharmaceuticals), be aware of this problem in Excel!

- **Arithmetic Average or Mean**

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} (x_1 + \dots + x_N)$$

where N is the total number of samples, and  $x_i$  is the measurement on the  $i^{\text{th}}$  attempt.

- **Median** =  $[(n+1)/2]^{\text{th}}$  value; or the middle value that separates the higher half of the data set from the lower half.
- **Mode** = the most frequent value in the data set.
- **Standard Deviation...**

... of the samples (assuming small sampling, such that  $n <$  total population size):

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

... of the population (with Bessel's correction):

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

- **Standard Error or Standard Deviation of the Mean**

$$\sigma_{\text{mean}} = \frac{\sigma}{\sqrt{N}}$$

## 5. REPORTING UNCERTAINTIES/SIGNIFICANCE FIGURES

After you have calculated the average and standard deviation in your measurements, how do you actually report it? How many significant figures would you include? Typically you round the error/uncertainty to one or two significant figures. Thereafter the reported value of your arithmetic mean would simply follow the number of decimal places of your rounded error (e.g.  $2.36 \pm 0.08$ ,  $10.0 \pm 2.0$ , or  $4 \pm 1$ ). A good reference for how scientists report uncertainties is the National Institute of Standards and Technology (NIST) website of fundamental physical constants [2]. Remember to include the correct units when reporting your results.

## 6. ERROR ANALYSIS AND PROPAGATION

My favourite reference book on statistical analysis is *Statistics and Chemometrics for Analytical Chemistry* [3]. It is full of simple and straightforward examples explaining how to apply each concept/test/analysis they have just discussed, which is useful for even non-chemists. Luckily, life in the research world is more interesting than just taking one measurement and stopping at that. It usually involves using the measurements you made to further calculate and derive something meaningful, such as the acceleration of gravity, or the proliferation index of cells growing *etc.*

For random errors (*i.e.*  $\Delta x$  and  $\Delta y$  are uncertainties independent of each other), a calculated variable,  $z$ , that is a function of two measured variables  $x$  and  $y$ :

$$z = f(x, y)$$

has an associated uncertainty,  $\Delta z$ :

$$\Delta z = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 (\Delta x)^2 + \left(\frac{\partial z}{\partial y}\right)^2 (\Delta y)^2}$$

**Equation 1**

Based on this partial derivative equation, we can derive the formulas for  $\Delta z$  in various scenarios (derive them using Equation 1 to check they are indeed true!):

<b>f(x,y)</b>	<b>Formula for <math>\Delta z</math></b>
$z = x + y$	$(\Delta z)^2 = (\Delta x)^2 + (\Delta y)^2$
$z = x - y$	$(\Delta z)^2 = (\Delta x)^2 + (\Delta y)^2$
$z = x \cdot y$	$(\Delta z/z)^2 = (\Delta x/x)^2 + (\Delta y/y)^2$
$z = x/y$	$(\Delta z/z)^2 = (\Delta x/x)^2 + (\Delta y/y)^2$
$z = x^n$	$\Delta z/z = n \cdot \Delta x/x$
$z = \ln x$	$\Delta z = \Delta x/x$
$z = e^x$	$\Delta z/z = \Delta x$

## 7. HOW TO CONDUCT A LITERATURE SEARCH (AND USE A REFERENCE MANAGEMENT SOFTWARE)

The university has a subscription to “Web of Science” (WoS), a product from ISI Knowledge Web (Thomson). This is an excellent tool for conducting literature search. Other tools include Medline, Google Scholar *etc.*

What makes WoS so great is: i) the ability to search by year, authors, title keywords, and allows you to use Boolean operators (AND, OR) and wildcards (\*); ii) the ability to directly link you to the article based on the university’s journal subscriptions; iii) the ability to order the results of your search based on “number of times cited”; and finally iv) the ability to import the search result directly into a reference management software, saving you the time and effort to type the reference into your computer yourself.

Let’s start by going to WoS. Search for “Web of Science” on the Cornell library page:

**Search:**

Click on the correct links to go to the WoS site:

- 

**Web of science**  
by Institute for Scientific Information.;  
eBook : Updating database  Continually Updated Resource  
Language: English  
Publisher: [Stamford, CT.] : Thomson Corp.  
Held by: Cornell University Library  
[View all editions and formats](#)

**Cornell University Library**

Let’s say we wish to search for articles pertaining to “nanotechnology”, published within the time period 2003-2005. We can use select the appropriate fields and type in our search parameters.

However, if we simply key in “nanotechnology” under the “Topic” field, we might miss out on articles that contain the word “nanotechnologies”. To overcome this limitation, we employ the use of the wildcard \*, and key in “nanotechnol\*”. Boolean operators can be used here.

A useful tip when first starting to research on a new topic is to search for “Review” articles. Reviews usually provide a comprehensive summary of the literature, as well as future outlook and challenges *etc.* Again, WoS allows you to search based on “Document Type”.

ISI Web of Knowledge<sup>SM</sup> Take the next step

All Databases | Select a Database | Web of Science | Additional Resources

Search | Cited Reference Search | Advanced Search | Search History | Marked List

**Web of Science®**

Search for:

nanotechnol\* in Topic  
Example: oil spill\* mediterranean

AND 2003-2005 in Year Published  
Example: 2001 or 1997-1999

AND Record Review  
Reprint  
Review  
Script in Document Type  
Select one or more from the list above.

After searching and obtaining your results, you can additionally sort them based on “Times Cited”, which will bring up the most highly cited article first. You can also refine results by searching within your results for a specific term, say “nanotubes”. Finally, click on “Get it at Cornell” to directly retrieve the document.

**Results** Topic=(nanotechnol\*) AND Year Published=(2003-2005) AND Document Type=(Review)  
Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI.

Scientific WebPlus<sup>SM</sup> View Web Results >>

Results: 261 Page 1 of 27 Go

Sort by: Times Cited

Print | E-mail | Add to Marked List | Save to EndNote® Web | Save to EndNote®, RefMan, ProCite | more options

Analyze Results | Create Citation Report

**Refine Results**

Search within results for: nanotubes

Subject Areas

- CHEMISTRY, MULTIDISCIPLINARY (57)
- MATERIALS SCIENCE, MULTIDISCIPLINARY (50)
- PHYSICS, APPLIED (37)
- NANOSCIENCE & NANOTECHNOLOGY (33)
- PHYSICS, CONDENSED MATTER (32)

more options / values...

- Title: Gold nanoparticles: Assembly, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology  
Author(s): Daniel MC, Astruc D  
Source: CHEMICAL REVIEWS Volume: 104 Issue: 1 Pages: 293-346 Published: JAN 2004  
Times Cited: 1,996  
Get it! Cornell
- Title: Self-assembled monolayers of thiolates on metals as a form of nanotechnology  
Author(s): Love JC, Estroff LA, Kriebel JK, et al.  
Source: CHEMICAL REVIEWS Volume: 105 Issue: 4 Pages: 1103-1169 Published: APR 2005  
Times Cited: 1,125  
Get it! Cornell

If you click on the title of the document directly, you will be brought to the following page. Here you can import the information of this reference into a reference management program such as Endnote, Refworks, Connotea *etc.* Refworks and Connotea are available free online, while you have to pay for Endnote. The latter offers sophisticated pre-formatting for various journals.

You can also add the selected reference to a “Marked List”, similar to a shopping cart, for a simultaneous download of all the references information later.

Another useful button is “Create Citation Alert”. This only works if you signed up for a free account with Thomson and logged in. You can set up e-mail alerts to let you know when this article is cited in literature, a good way of keeping abreast with latest research.

<< Back to results list | Record 1 of 261 | Record from Web of Science®

**Gold nanoparticles: Assembly, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology**

Get it! Cornell | Print | E-mail | Add to Marked List | Save to EndNote Web | Save to EndNote, RefMan, ProCite | more options

**Author(s):** Daniel MC, Astruc D

**Source:** CHEMICAL REVIEWS Volume: 104 Issue: 1 Pages: 293-346 Published: JAN 2004

**Times Cited:** 1,996 **References:** 820 [Citation Map](#)

**Document Type:** Review

**Language:** English

**KeyWords Plus:** LANGMUIR-BLODGETT-FILMS; MONOLAYER-PROTECTED CLUSTERS; NONLINEAR-OPTICAL RESPONSE; SHAPE-CONTROLLED SYNTHESIS; AQUEOUS CHLOROAUROATE IONS; ENHANCED RAMAN-SCATTERING; CORE-SHELL NANOPARTICLES; SINGLE-CRYSTAL SURFACES; QUANTITATIVE COLORIMETRIC DETECTION; WATER/AOT/N-HEPTANE MICROEMULSIONS

**Reprint Address:** Astruc, D (reprint author), Univ Bordeaux 1, LCOO, UMR CNRS 5802, Mol Nanosci & Catalysis Grp, F-33405 Talence, France

**Addresses:**  
1. Univ Bordeaux 1, LCOO, UMR CNRS 5802, Mol Nanosci & Catalysis Grp, F-33405 Talence, France

**Publisher:** AMER CHEMICAL SOC, 1155 16TH ST, NW, WASHINGTON, DC 20036 USA

**Subject Category:** Chemistry, Multidisciplinary

**Cited by: 1996**  
This article has been cited 1996 times (from Web of Science).  
Kryachko ES Nonconventional Hydrogen Bonding and Vertical Electron Detachment in Anionic Complexes of Gold with DNA Bases: Few Essayistic Fragments POLISH JOURNAL OF CHEMISTRY 83 5 917-931 MAY 2009  
Astruc D Electron-reservoir complexes and other redox-robust reagents: functions and applications NEW JOURNAL OF CHEMISTRY 33 6 1191-1206 2009  
Zhu MZ, Qian HF, Jin RC Thiolate-Protected Au-20 Clusters with a Large Energy Gap of 2.1 eV JOURNAL OF THE AMERICAN CHEMICAL SOCIETY 131 21 7220-+ JUN 3 2009

[ view all 1996 citing articles ]  
Create Citation Alert

Finally, there are many more cool features in WoS, so go ahead and explore!

Why bother with using a reference management software? If you start to cite more than 10 references in the report you are writing, it will save you a lot of time and heartaches. First, it helps you with formatting in Microsoft Word. You can insert the reference anywhere you like into your report and Endnote automatically numbers and adds the details of the reference at the end of the report. If you change the place you inserted your reference within the text of your report, Endnote automatically updates the numbering and order of the references for you – easy peasy japanesey.

## 8. CONCLUSIONS

I trust that this crash course is beneficial and inspires you to want to do research. I wished I had something like this when I first started working in the lab as an undergraduate. Fortunately I had many good mentors and I learned all this verbally and on the job. This is the beginning of your journey in science, so building a good foundation, cultivating good habits and logical thinking are important at this stage. It is my hope that you will find enjoyment and fulfillment in the lab, and become a responsible and proficient scientist.

## 9. REFERENCES

- [1] Edward R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, 2nd edn., 2001.
- [2] NIST Fundamental Physical Constants; <http://physics.nist.gov/cuu/Constants/>
- [3] James N. Miller and Jane C. Miller, *Statistics and Chemometrics for Analytical Chemistry*, Pearson Prentice Hall, 5th edn., 2005.