

FROM THOUGHT TO PLOT: REVEALING UNDERGRADUATE BIOLOGY STUDENT GRAPHING PRACTICES

SimBio Webinar

March 3rd , 2022



Stephanie M. Gardner, Ph.D.
Associate Professor of Biological Sciences



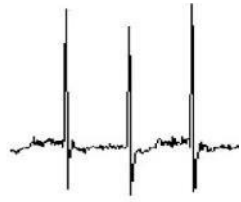
Artistic credit: Autumn Siebold

Rest



What would you do?

Exercise



ajpheart.physiology.org

Question: Will the average number of hours exercised in a week affect the heart rate recovery rate after exercise?

Prediction: Subjects who exercise more have faster recovery rates than those who do not exercise as often per week.

Excerpted data

Subject #	Exercise (hours/week)	Decrease in heart rate at 2 minutes (bpm)
1	5	10
2	4	15
3	3	21
4	5	34
5	4	30

THINK: What type of graph would you use to represent these data?
Why?

Genesis of this work: Undergraduate student graphing

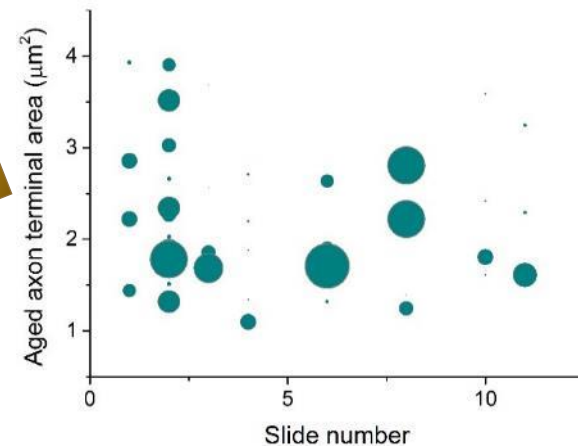
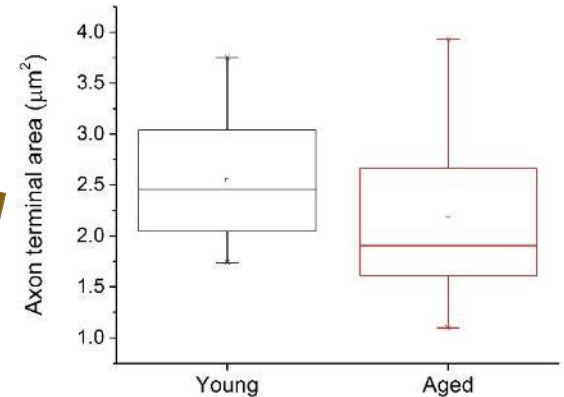
In my Physiology class students design and execute experiments and analyze and present their data.

Function of graphs (Tufte, 1983)

Graphical displays should:

- Show and reveal the data
- Focus the viewer's attention on the display
- Avoid distortion of the data
- Present large numerical data sets in a small space
- Make large data sets coherent
- Encourage visual comparisons
- Reveal the large and fine trends in the data
- Serve a clear purpose
- Be integrated with statistical and verbal descriptions of the data

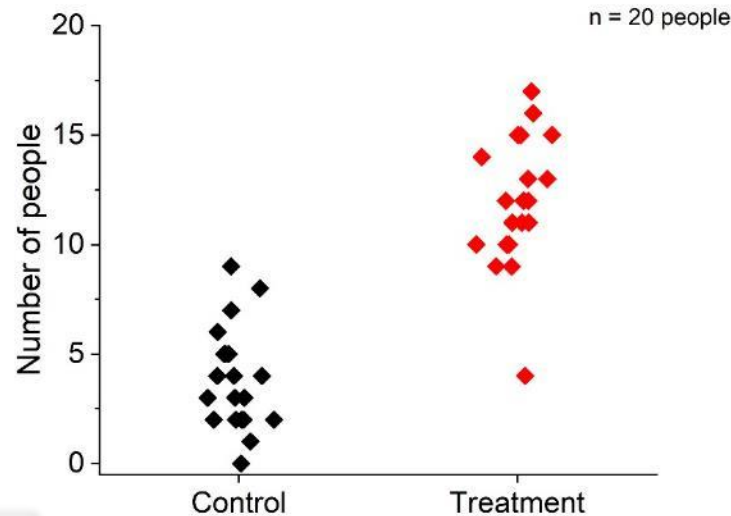
3.75174	3.929705
3.73122	3.905297
3.73122	3.685624
3.72438	3.587991
3.67992	3.514767
3.33108	3.246278
3.23874	3.026605
3.20796	2.855748
3.13614	2.806932
3.13614	2.709299
3.04038	2.660483
2.85228	2.636075
2.84202	2.562851
2.82492	2.416402
2.8044	2.343178
2.69154	2.294362
2.565	2.269954
2.565	2.221137
2.52396	2.221137
2.49318	2.196729
2.47266	2.025873
2.45898	1.903832
2.38374	
	1.879424
2.26404	1.879424
2.26404	1.879424
2.25378	1.879424
2.21958	1.855016
2.19906	1.8062
2.17854	1.781792
2.14434	1.732975
2.09304	1.708567
2.08962	1.684159
2.04516	1.610935
2.04516	1.610935
1.9836	1.610935
1.98018	1.513302
1.881	1.440078
1.83312	1.391262
1.79208	1.342446
1.76814	1.318038
1.75446	1.318038
1.75104	1.244813
1.74078	1.098365



Basic graph construction knowledge and skills

Graph types

Math and statistics

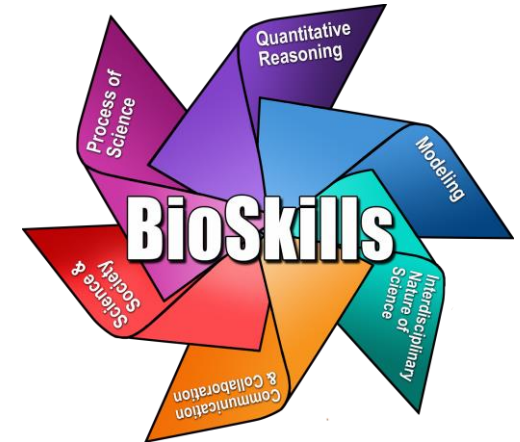
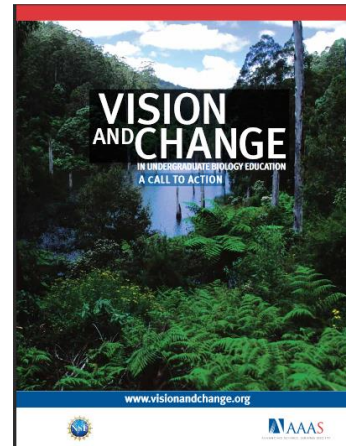
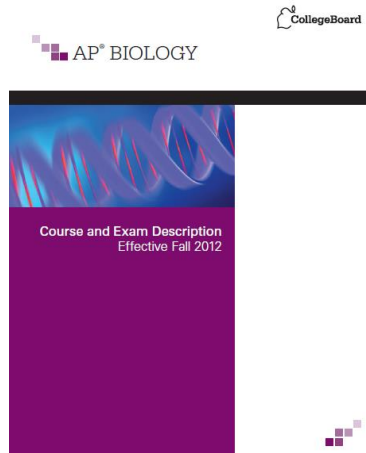


Concepts and practices of the discipline

Visuospatial and aesthetics

Undergraduate student graphing

- Reforms to STEM education at both the high school and undergraduate levels stress science process and practices



- Students will be making decisions about data analysis and representations
- However, more emphasis has been placed on interpreting graphs and not creating them (e.g. AP Biology up to 2012)
- Students are competent plotting points and identifying coordinates
- Students struggle with important concepts related to experiments, data analysis, and graphing

Padilla et al. ,1986; Brasell & Rowe, 1993; Berg and Smith, 1994; Ainley, 1995; Mevarech & Kramarsky, 1997; Mathewson, 1999; Grunwald and Garfield,2003; Bakker, 2004; Leonard & Patterson, 2004; Clase et al., 2010; Hartman, 2010; Tairab & Al Naqbi, 2010; Meletiou & Lee, 2010; Bray-Speth et al., 2010; Hattikudur et al., 2012; Bray-Speth et al. 2010; McFarland 2010; Roth & Bowen 2001; Harsh and Schmitt-Harsh, 2016; Schultheis and Kjolvik, 2020

A call for better representations by practitioners (2014–present)

OPEN ACCESS
Editorial

Ten S

Nicolas P
1 INRIA Bordeaux
4 Space Telescop
States of Ameri

Scientific
defined as
displaying a
process is E
There are
represent th
linear plots,
name just a
data, using
perceived vs
who is look
accurate def
tion would
between per
article, we
everything a
[1,2] for in
aim to impr
some of the

Rule 1: K

Given the
arise when
differs sign
the convey
tant to in
the design
message th
graphical d
informed b
making a f
direct colla
skip a num
process, b
what the f
you intend
scientific jo
figure, rele
relevant inf
care since
explain a c
have to ad
sure the c
Finally, the
most diffic
need to be
proximated
most salie
(Figure 1).
difficult ex

PLOS Comput

PLoS

Saxon BMC Biology
DOI:10.1186/s129

COMMENTARY

Beyond

Emma Saxon

Abstract

Probiotic trea
of commens
human gut, c
an important
on the level
guts of health
increased wi
control. The
thus help to

Commentary

Bar charts are
represent or
data from this
appears to sup
biotic significa
mensal species
t-test $p < 0.01$)
on publication
research group
conducted the
Fig. 1b, to de
replicable. Ind
significant tre
 $p < 0.05$), alth
control treatm
However, wh
showing the re
treatment in th
picture emerge
in gut commes
that make up
value was part
occurred in th
increase seen

Correspondence: B
BMC Biology, BioMe

Competing Interests
that no competing

PLOS Biology |

Editorial

The Effect

David J. Slutsky

¹ The Hand and W
Professor, Depart
Los Angeles, Calif

J Wrist Surg 2014

Graphs are a com

ships in the data,
are too numero
in the text and i
small amounts o
sentence. Likewi
defeats the purp
pronounced tre
graph should be
trend in the evid
Although thes
generate a grap
principles. A bas
and readable. Th
symbols but y
provide a clear
may have sever
figure number, (3
data field, (5) ax
a credit or sour
that the vertical
dent variable a
represents the i
on the X axis.²
caption, axes an
symbols need to
trast between t
ground. Open an
are more effect
open squares.³ L
should concisely
about what the
provide a summ
mental details. A
"temperature vs
graph type base
diagrams or scat
numeric, use bar
pie charts. These

Introduction

Although ba
and are ubiq
they do not al
dataset. This
graphical met
box plot, also
basic elements
how to correct
available to p
software that
publication.

Weaknesses of

A bar chart
zations and is
package. The
zing counts or
not always
comparing nur
suppose that
Berg Balance S
and after a th
Using a bar ch
tentia

Address for corres
David J. Slutsky, MD
Wrist Institute, 238
Street, Torrance, C
(e-mail: d-slutsky@



AMERICAN ACADEMY
OF MEDICAL RESEARCH
AND EDUCATION

The Effect

Bernd Klaus

Commentary

In the first pa
discussed gener
the results o
(Klaus, 2015). The
representation of t
of data presentati
mated as figures at
data analysis and
an important anal
to reveal pattem
The appropriate
strengthen or ev
statistical proced
tests. In the contex
figures should gui
article and provide
sentation of the ex

Introduction

In this article,
key principles and
ing small-to-med
comparing results
tal groups. As a g
show as much of
instead of summa
or variances. Eve
displayed efficien
plot; bars and box
statistics can set
guides. To adapt
this article, reader
mentary "notebo
code to generate t
(R Core Team, 201
tool (<http://emboj>)
the bee swarm
discussed later in
the topics discuss
are many more
attention. The "C
Course" material
unit of the Babra

1. The value of

which is reg
top line of t
conveys no
waste of sp
2. The bar lea
is an import
vertical axis
often misle
necessarily
the sample
magnitude
bar is purel

1934-1482/5 - see
<http://dx.doi.org>

Published online: July 1

Commentary

Focus: Study

Statistics

part I

Bernd Klaus

Introduction

In the first pa
discussed gener
the results o
(Klaus, 2015). The
representation of t
of data presentati
mated as figures at
data analysis and
an important anal
to reveal pattem
The appropriate
strengthen or ev
statistical proced
tests. In the contex
figures should gui
article and provide
sentation of the ex

Weaknesses of

A bar chart
zations and is
package. The
zing counts or
not always
comparing nur
suppose that
Berg Balance S
and after a th
Using a bar ch
tentia

1. The value of

which is reg
top line of t
conveys no
waste of sp
2. The bar lea
is an import
vertical axis
often misle
necessarily
the sample
magnitude
bar is purel

1934-1482/5 - see
<http://dx.doi.org>

PLoS

Received: 24 May 2018

DOI: 10.1002/pst.1912

MAIN PAPER

How can
the grap
scientist

Marc Vandenberg

Alison Margolis

¹ Biostatistical Scienc

Pharmacometrics, N

Basel, Switzerland

² CD&A Strategy & I

Pharmaceuticals Co

Hanover, New Jerse

³ Biostatistical Scienc

Pharmacometrics, N

Biomedical Research

Massachusetts, USA

Correspondence:

Marc Vandenberg

Pharma AG, Asklep

Campus 4056 Basel,

E-mail: marc.vanden

novartis.com

1 | INTRODUCTION

Chambers et al¹
then any statist
Tuftes² and Clev
reason may be
Nolan and Perre
demic teaching
obstacle happen
graphics that fa
may not even re
Increasing o
erally, quantitat
are at the core o
making. Hence,

Pharmaceutical Statist

Circulation

PRIMER

Reveal, Don't Conceal

Transforming Data Visualization to Improve Transparency

ABSTRACT: Reports highlighting the problems with the standard practice of using bar graphs to show continuous data have prompted many journals to adopt new visualization policies. These policies encourage authors to avoid bar graphs and use graphics that show the data distribution; however, they provide little guidance on how to effectively display data. We conducted a systematic review of studies published in top peripheral vascular disease journals to determine what types of figures are used, and to assess the prevalence of suboptimal data visualization practices. Among papers with data figures, 47.7% of papers used bar graphs to present continuous data. This primer provides a detailed overview of strategies for addressing this issue by (1) outlining strategies for selecting the correct type of figure depending on the study design, sample size, and the type of variable; (2) examining techniques for making effective dot plots, box plots, and violin plots; and (3) illustrating how to avoid sending mixed messages by aligning the figure structure with the study design and statistical analysis. We also present solutions to other common problems identified in the systematic review. Resources include a list of free tools and templates that authors can use to create more informative figures and an online simulator that illustrates why summary statistics are meaningful only when there are enough data to summarize. Last, we consider steps that investigators can take to improve figures in the scientific literature.

Tracey L. Weissgerber, PhD
Stacey J. Winham, PhD
Ethan P. Heinzen
Jelena S. Milin-Lazovic, MD
Oscar Garcia-Valencia, MD
Zoran Bukumiric, MD, PhD
Marko D. Savic, MPH
Vesna D. Garovic, MD, PhD
Natasa M. Milic, MD, PhD

Key Words: bar graphs ■ basic science
■ continuous data ■ data visualization

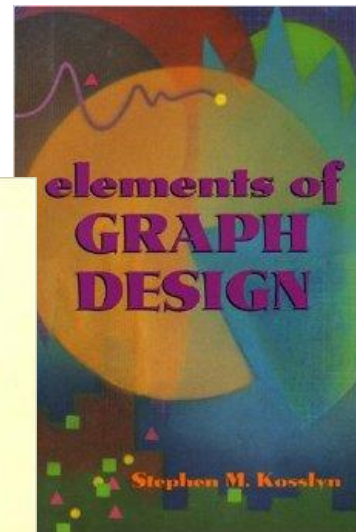
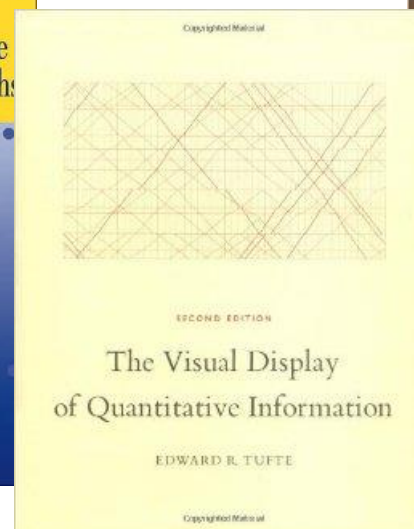
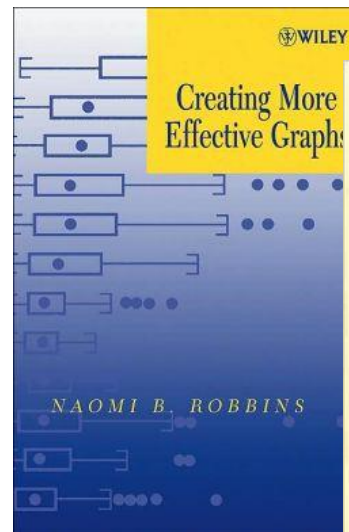
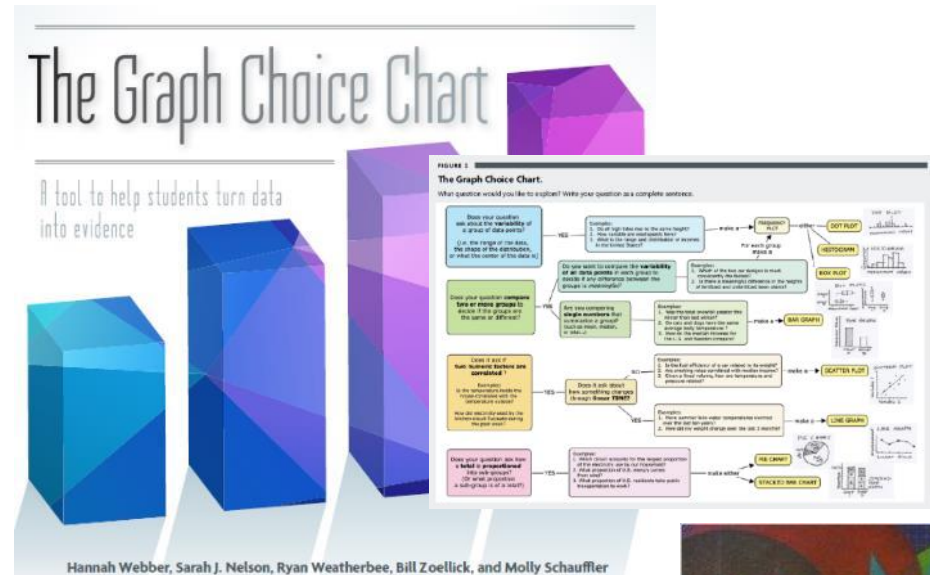
© 2019 The Authors. Circulation is published on behalf of the American Heart Association, Inc., by Wolters Kluwer Health, Inc. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited and is not used for commercial purposes.

<https://www.ahajournals.org/journal/circ>

Downloaded from <http://ahajournals.org> by on November 28, 2019

Existing graph choice and construction resources

- Generic guidelines on proper graph choice and construction
- Materials not situated in the context of experiments for science/biology
- No empirical investigation on reasoning with graph choice and construction



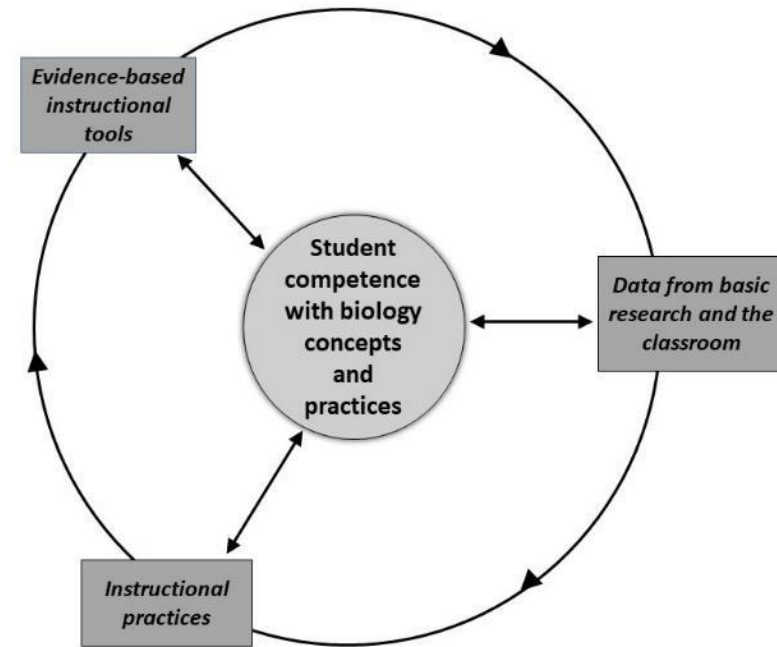
Motivating question and objectives

Overarching research question:

What are the reasons biology students struggle with graphing and how can we help them?

Objectives:

- Evaluate the graph construction practices in biology
- Reveal the reasoning that people use when creating graphs in biology
- Use those data to improve classroom instruction and student proficiency with graphing



Model inspired by: The Research and Redesign Wheel from Redish, E. F. (2003).



Aakanksha Angra

Early insights from clinical interviews

- Semi-structured think-aloud interviews
 - Pen-and-paper graph choice and construction task
- Participants from Department of Biological sciences:
 - Biology Professors (N=7)
 - Graduate students (N=12)
 - Undergraduate students (N=33; n=12 upper level and n=21 lower level)

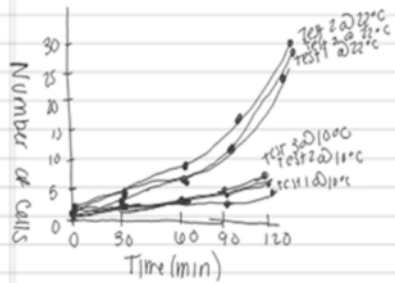


www.dorkadore.com

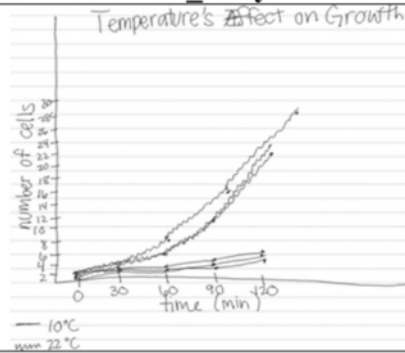
Graph attributes vary by participant group

Bacterial growth scenario

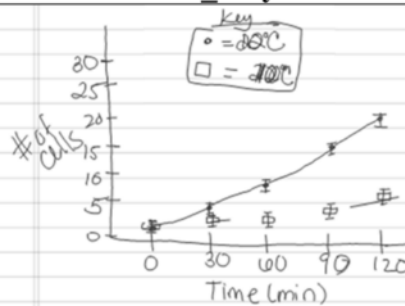
UGNR1_2nd year



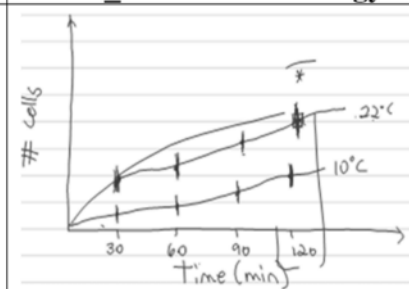
UGR2_4th year



GS1_3rd year

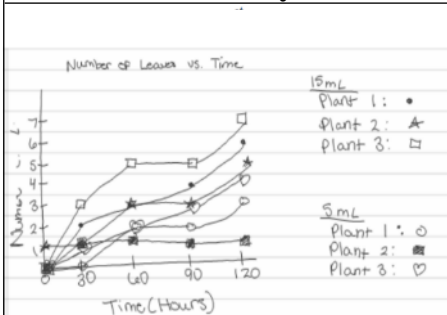


P2_Behavioral ecology

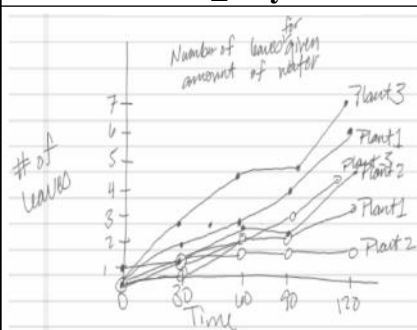


Plant growth scenario

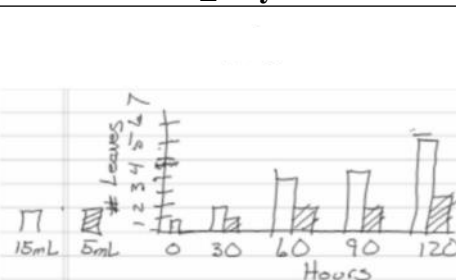
UGNR5_1st year



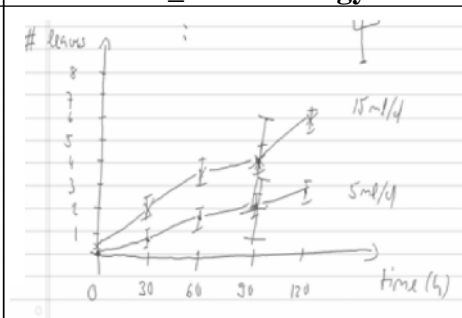
UGR1_4th year



GS2_4th year



P5_Neurobiology

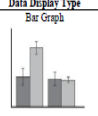





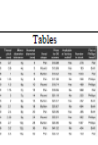




- *Minimal to no planning prior to construction*
- *Meticulous graph construction*
- *Raw data often used in graphs*
- *Data considered independent of experimental context*
- *Variability in experimental data not considered*
- *Graph choice not based on question and/or hypothesis*
- *Intuitive reasoning about data variables and experimental replicates*
- *Reflection focused on superficial graph features and aesthetics*

Guides to Improve Graph Choice and Construction

I use these with students in all my classes!

Data Display Type	Usage	Advantages	Disadvantages
 <p>Bar Graph</p>	To compare categorical data from multiple groups. ^{12, 44} Each bar represents a category; shape can be changed by moving the categories around. ⁴⁷ The independent variable is categorical and the dependent variable is continuous. ⁵	Easy to interpret. Stacked bars or shading of bars can be used to distinguish the different levels within the data. ^{12, 44}	Does not convey more information than a table unless multiple groups are compared. ^{12, 44} Obscures the distribution of data ¹² , number of data points, and their values. ^{16, 19}
 <p>Box and Whisker Plot</p>	To show distribution of data from one ¹⁰ or multiple groups. ^{14, 44}	Shows and compares distributions of large datasets. ^{14, 44}	Should not be used for small datasets. ¹² Does not show individual data (except for outliers). ^{14, 44}
 <p>Histogram</p>	To show a distribution of data with multiple categories. ^{12, 44}	Shows the shape of the distribution.	Must choose the bin size wisely to avoid influencing the shape being too compressed or too dispersed. ^{12, 27}
 <p>Line Graph</p>	To show how multiple variables change over time. ¹²	Shows trends between data. ²²	Not appropriate for representing averages of a group. ²¹
 <p>Dot Plot</p>	To show distribution of small data sets from multiple groups. ^{16, 41} The independent variable is categorical and the dependent variable is continuous.	Shows all data from multiple categories and the distribution within each category. ^{16, 41}	Not appropriate for representing a large data set because the plot will become cluttered and it will be difficult to see the individual points. ⁸
 <p>Scatterplot</p>	To show individual data points from bivariate data. ⁴⁴ May or may not include a regression line. ⁴²	Shows the relationship between variables. ^{22, 44} Shows trends in the data and any noticeable outliers.	It may be difficult to extract individual data points if they fall on the same or nearby coordinates. ^{22, 44}
 <p>Tables</p>	To summarize findings, compare various aspects in a study, display demographic information, list raw data obtained from experiments. ²²	Summarizes main points in one figure. ¹² Illustrates precise, quantitative values. ^{22, 39} A well-organized table should easily communicate patterns in the data. ⁴⁰	Should not be used for showing small amounts of data that can be summarized in the main text. ¹² Tables may make it difficult to interpret the take home message if not organized properly. ⁴⁰

Guide to data displays

The purpose of this table is to guide you through the data communication phases.

Phases 1-3	Elements	Notes for your Experiment
1. Planning- In this phase, you must organize your data and decide on the message you want to communicate in your graph. It helps to first conceptualize the whole task before executing it.	<p>Step 1- Revisit your research question and hypothesis and ask yourself, what is it that you want the graph to show?</p> <p>Step 2- Identify your independent and dependent variables.</p> <p>Step 3- Classify your variables as either categorical or</p>	
2. Execution- You execute your plan by creating the graph.	<p>Step 4- Choose the appropriate graph type for your data.</p> <p>Step 5- Adjust the scale of axes into appropriate increments for the data.</p> <p>Step 6- Include a key, if appropriate.</p> <p>Step 7- If you are displaying the graph in a report, include a figure legend.</p> <p>Step 8- Adjust the scale of axes into appropriate increments for the data.</p> <p>Step 9- Include a key, if appropriate.</p> <p>Step 10- If you are displaying the graph in a report, include a figure legend.</p> <p>Step 11- Include a descriptive title.</p>	
3. Reflection and Explanation Phase- In this phase, you will critically reflect on your graph choice, interpret your graph, and explain your answers to questions posed in steps 13-16.	<p>Step 12- Check the alignment of your representation with your research question and hypothesis.</p> <p>Step 13- What are the advantages of the representation?</p> <p>Step 14- What are the disadvantages of the representation?</p> <p>Step 15- What is the take-home message of the representation?</p> <p>Step 16- What are some other ways that you could have represented your data?</p>	

Step-by-step guide to graph construction

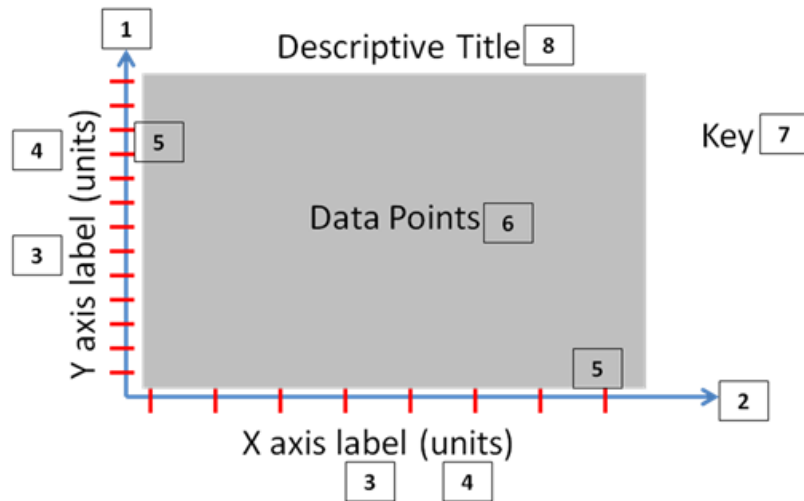
	Present / Appropriate (PA)= 0.5 pts for each category Present but Needs Improvement (NI)=0.25 pts for each category Absent / Inappropriate (NI)= 0 pts	GRAPH 1		
		PA	NI	NI
Graph Mechanics	Description/Title			
	<ul style="list-style-type: none"> PA- Should be a) in the form of a statement, b) mention the subject, c) appropriate variables, and d) include relevant details about the experiment that help understand the take-home message. NI- If the title is missing any one of the four points mentioned above 			
	Label for the X axis (e.g. time)			
	<ul style="list-style-type: none"> PA- Should be appropriate and descriptive for the experiment. For graphs with categorical independent variables, there needs to be a label under each set of data and a larger label under all data plotted. NI- If the label is missing any one of the points mentioned above 			
Graph Choice	Label for the Y axis (e.g. heart rate)			
	<ul style="list-style-type: none"> PA- Should be appropriate and descriptive for the experiment. If the data is manipulated (average, change, percentage, etc.), then it should be indicated on the y-axis. NI- If the label is missing any one of the points mentioned above 			
	Units for the X axis (e.g. seconds)			
	<ul style="list-style-type: none"> PA- Should be appropriate and descriptive for the data displayed. NI- If the units are not appropriate or descriptive 			
Communication	Units for the Y axis (e.g. average beats per minute)			
	<ul style="list-style-type: none"> PA- Should be appropriate and descriptive for the data displayed. NI- If the units are not appropriate or descriptive 			
	Scale (e.g. 0 to 100)			
	<ul style="list-style-type: none"> PA- Should be appropriate and descriptive for the data displayed. NI- If the units are not appropriate or descriptive 			
Graph Choice	Key (e.g. 0 to 100)			
	<ul style="list-style-type: none"> PA- Should be appropriate and descriptive for the data displayed. NI- If the units are not appropriate or descriptive 			
	Excellent (E) = 2 pts for each category Needs Improvement (NI) = 1 pt Unsatisfactory (U) = 0 pts			
Communication	Ease of Understanding-Aesthetics			
	<ul style="list-style-type: none"> E- If the graph is aesthetically pleasing, meaning that: a) the data plotted takes up sufficient room in the Cartesian plane, b) makes use of legible size font, c) uses and y axis lines are clear and legible, d) the graph displays data in an appropriate number of bars and lines, and e) is devoid of chartjunk elements such as distracting background colors, patterns, and grid lines. NI- If the graph has one of the following flaws: a) the graph displays too much white space, b) the font size is too small, c) the x and y axis lines are not clear and legible, d) the graph shows too many bars or lines, e) elements of chart junk are clouding interpretation of data. U- If the graph is ineffective at communicating data trends and take home message, such that it causes confusion. 			
	Ease of Understanding- Take home message			
	<ul style="list-style-type: none"> E- If the graph has sound construction and mechanics that allow for clear sorting of trends and take home message. NI- If data trends are difficult to observe or it is difficult to formulate a proper take home message. U- If the graph is ineffective at communicating data trends and take home message, such that it causes confusion. 			
Graph Choice	Graph Type (Bar, Line, Scatter, Dot, Box and Whisker)			
	<ul style="list-style-type: none"> E- If data displayed in a graph is appropriate for both independent and dependent experimental variables (i.e. categorical and continuous) and data. (Referring to the data form) NI- If data displayed in a graph is a) not suitable for either the dependent or independent experimental variables OR b) there is a better way to present data. U- If the graph type is not suitable for both experimental variables. 			
	Data Display (Bar, Line, Scatter, Dot, Box and Whisker)			
	<ul style="list-style-type: none"> E- If the graph indicates the type of data (e.g. Bar, averages, etc.) that are plotted. There should be a clear distinction between raw data and manipulated data based on the information presented in the key (e.g. sample size and number of trials) and axis label. If the graph is showing averages, then it should also be accompanied with STDDEV or error bars. NI- If the graph is missing one or more of the points mentioned above. U- If data type is inappropriate for the graph type. 			
Graph Choice	*Alignment* (at least one of the graph's elements should align with the research question and hypothesis. Other graphs can be appropriate.)			
	<ul style="list-style-type: none"> E- If the graph is completely aligned with the research question and/or hypothesis. In other words, the independent, dependent variables, and information about the experiment are explicit. NI- If the graph is partially aligned with the research question and/or hypothesis. In other words, the graph is missing information about either the independent, dependent, or details about the experiment. U- If the graph is not aligned with the research question and/or hypothesis. 			
	Total for Graph Choice and Communication			
	Overall Presentation Grade			
Comments:				

Graph rubric

Angra A and Gardner SM (2016). Development of a Framework for Graph Choice and Construction. *Advances in Physiology Education* 40: 123-128. doi:10.1152/advan.00152.2015

Angra A and Gardner SM (2018) The Graph Rubric: Development of a Teaching, Learning, and Research Tool. *CBE-Life Sciences Education* 17: ar65 doi: 1187/cbe.18-01-0007

The Guide to Data Displays



Box plots
Box plots are a statistical plot that allows you to provide very detailed descriptive statistics about a data set in a concise visual. The median of the data set (line in the middle of the box) is determined and then the data set is divided into four, equal parts (quartiles) along the range of the data values. The number of data points that fall into each of those quartiles is represented by the size of the box (for the two middle quartiles around the median) or whisker for the outer quartiles. The mean is often denoted with an asterisk or other symbol.

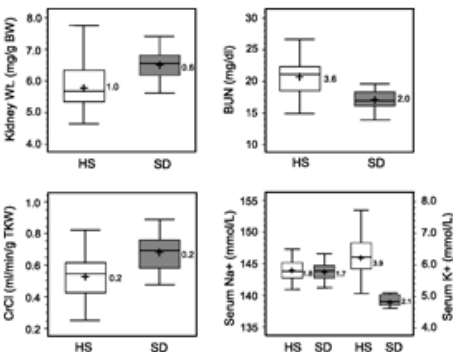


Figure 5. Example of a box plot (From: <http://ajprenal.physiology.org/cgi/content-nw/full/298/6/F1484/F2>)

Table 1- Summaries the common types of graphs, their usage, advantages, and disadvantages.

Graph Type	Usage	Advantages	Disadvantages
Bar Graph	To compare categorical data from one or multiple groups. ^{1,2} Each bar represents a category; shape can be changed by moving the categories around. ³	Easy to interpret. Stacked bars or shading of bars can be used to distinguish the different levels within the data. ^{1,2}	Does not convey more information than a table unless multiple groups are compared. ^{1,2}
Box and Whisker Plot	To show distribution of data from one ¹ or multiple groups. ²	Shows and compares distributions of large datasets. ^{1,2}	Cannot be used for small datasets. Does not show individual data (except for outliers). ^{1,2}
Histogram	To portray a sampling distribution with a continuous independent variable. ^{1,2} Uses numerical data instead of categorical data. ³	Shows the shape of the distribution of data with a continuous variable. ^{1,2}	Must choose the bin size wisely to avoid influencing the shape being too compressed or too dispersed. ^{1,2}
Line Graph	To show how a single variable or multiple variables change over time. ¹	Shows data values and slope between them. ¹	Not appropriate for representing averages of a group. ¹
Pie Graph	Effective at presenting relative frequencies or percentages. ^{1,2,4} Each segment represents a proportion of the whole pie. ^{4,5}	Good for presentations. ¹ Multiple pie charts may portray emerging patterns that may not be easily portrayed by a table. ²	Solitary pie graphs are redundant with tables. ^{1,2} Cannot communicate distribution, uncertainty, and cannot be used for displaying nested data. ² Multiple pies are ineffective at comparing proportions that vary greatly. ⁴
Scatterplot	To show individual data points from bivariate data. May or may not include a regression line. ^{1,2}	Preserves dimensions of data and individual points, and shows the relationship between the variables. ^{1,2}	Points that fall on the same or close coordinates may or may not be distinguished easily. ^{1,2}
Tables	Shows comparison between multiple groups, displays demographic information, lists raw data obtained from experiments. ¹	Illustrates precise values from the data. ^{1,4}	Should not be used for small numbers that can be summarized in 1 or 2 sentences. ¹ Tables may also make it difficult to interpret the take home message.

The Step-by-step guide

Phases 1-3	Elements	Notes for your Experiment
1. Planning- In this phase, you must organize your data and decide on the message you want to communicate in your graph. It helps to first conceptualize the whole task before executing it.	Step 1- Revisit your research question and hypothesis and ask yourself, what is it that you want the graph to show?	
	Step 2- Identify your independent and dependent variables.	
	Step 3- Classify your variables as either categorical or continuous.	
	Step 4- Decide whether or not you need to manipulate your data.	
	Step 5- Decide on a graph type that will best represent your data.	
2. Execution- In this phase, you will actively construct a graph.	Step 6- Label the axes with your variables.	
	Step 7- Add units to the axes, if necessary.	
	Step 8- Adjust the scale of axes into appropriate increments for the data.	
	Step 9- Include a key, if appropriate.	
	Step 10- If you are displaying the graph in a report, include a figure legend.	
	Step 11- Include a descriptive title.	
3. Reflection and Explanation Phase- In this phase, you will critically reflect on your graph choice, interpret your graph, and explain your answers to questions posed in steps 13-16.	Step 12- Check the alignment of your representation with your research question and hypothesis.	
	Step 13- What are the advantages of the representation?	
	Step 14- What are the disadvantages of the representation?	
	Step 15- What is the take-home message of the representation?	
	Step 16- What are some other ways that you could have represented your data?	

- Think-aloud interviews with expert professors and novice students
- 3 Phases:
 - Planning
 - Execution
 - Reflection
- Scaffolds a systematic and reflective approach to graphing



Expert-Novice Interviews

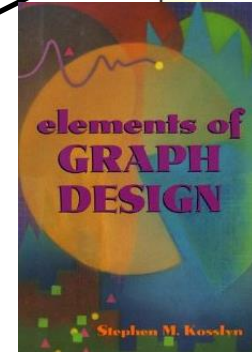


Student Reflections

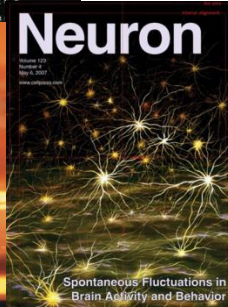
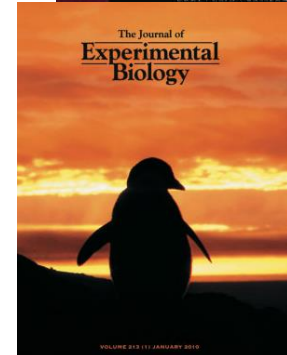
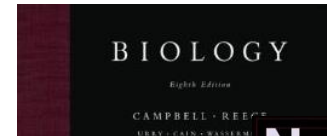
		GRAPH 1			GRAPH 2		
		P/A	NI	X/I	P	NI	X/I
Graph Mechanics	Present/Appropriate (P/A)= 0.5 pts for each category Present but Needs Improvement (NI) =0.25 pts for each category Absent/Inappropriate (X/I)= 0 pts						
	Descriptive title						
	Label for the X axis (e.g. time)						
	Label for the Y axis (e.g. heart rate)						
	Units for the X axis (e.g. seconds)						
	Units for the Y axis (e.g. average beats per minute)						
	Scale (appropriate intervals and range for data)						
	Key (defines different data sets that are plotted)						
Total Points for Mechanics:		/3.5pts					
Communication	<p>Element (1 pt)</p> <p>Its Implication (1 pt)</p> <p>Overall (2 pts)</p>						
	<p>Graph Type (Bar, line, scatter, dot, box and whisker)</p> <p>Aesthetics</p> <ul style="list-style-type: none"> E, if the graph is visually pleasing and does not contain chart junk NI, if the graph is cluttered or contains chart junk U, if chart junk is interfering with ease of understanding and interpretation of data. 						
	<p>Ease of Understanding: Take home message</p> <ul style="list-style-type: none"> E, if data trends are easy to observe or no trends are apparent. NI, if data trends are difficult to observe. U, if the graph is ineffective at communicating data trends and if it causes confusion. 						
	<p>Graph Type (Bar, line, scatter, dot, box and whisker)</p> <ul style="list-style-type: none"> E, if the graph indicates the type of data (ex. Raw, averages, etc.) that are plotted. If the graph is showing averages, then it should also be accompanied with STDEV or Error bars. NI, for displaying data in a graph that is not suitable for either the dependent or independent experimental variable. U, if the graph type is not suitable for the experimental variables. 						
	<p>Data Displayed (Raw, Averages, Changes, Percentage)</p> <ul style="list-style-type: none"> E, if the graph indicates the type of data (ex. Raw, averages, etc.) that are plotted. If the graph is showing averages, then it should also be accompanied with STDEV or Error bars. NI, if the graph is missing one of points mentioned above. U, if data type is inappropriate for the graph type. 						
Graph Choice	<p>*Alignment* (at least one of the graphs presented should align with the research question and hypothesis. Other graphs can be exploratory.)</p> <ul style="list-style-type: none"> E, if the graph is aligned with the research question and hypothesis. NI, if the graph is partially aligned with the research question and/or hypothesis. U, if the graph is not aligned with the research question and/or hypothesis. 						
	Total Points for Graph Choice and Communication:	/10pts					
	Overall Graph Presentation Grade:	/13.5 pts					
	Comments:						

THE RUBRIC

Graphing Literature



Graphs in Textbooks And Primary Literature



Student Graphs

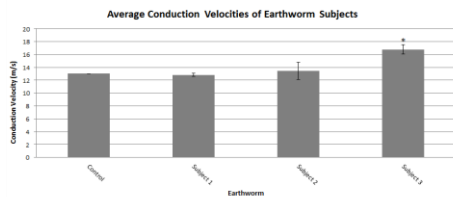
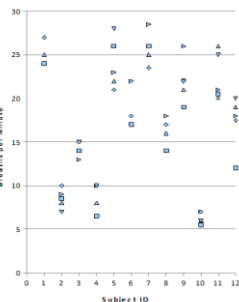


Fig. 1. The average conduction velocities of the median giant axon for the following earthworm subjects: Control at 20°C, and Subjects 1-3 at 25°C. Subject 3 has a significantly larger ($P < 0.05$) average conduction velocity when compared (Student T-Test) to the average conduction velocity of both Subject 1 and Subject 2.

Angra A and Gardner SM (2018) The Graph Rubric: Development of a Teaching, Learning, and Research Tool. CBE-Life Sciences Education 17: ar65 doi: 1187/cbe.18-01-0007

	Present /Appropriate (P/A)= 0.5 pts for each category Present but Needs Improvement (NI)=0.25 pts for each category Absent/Inappropriate (X/I)= 0 pts	GRAPH 1			GRAPH 2		
		P/A	NI	X/I	P	NI	X/I
Graph Mechanics	Descriptive title						
	Label for the X axis (e.g. time)						
	Label for the Y axis (e.g. heart rate)						
	Units for the X axis (e.g. seconds)						
	Units for the Y axis (e.g. average beats per minute)						
	Scale (appropriate intervals and range for data)						
	Key (defines different data sets that are plotted)						
Total Points for Mechanics:		/3.5pts					

- **Graph Mechanics**
- Descriptive Title
 - Axes Labels
 - Units
 - Scale
 - Key

Excellent (E) = 2 pts for each category Needs Improvement (NI) = 1 pts Unsatisfactory (U) = 0 pts		GRAPH 1			GRAPH 2		
		E	NI	U	E	NI	U
Communication	Ease of Understanding-Aesthetics <ul style="list-style-type: none">E, if the graph is aesthetically pleasing and devoid of chart junk.NI, if elements of chart junk are clouding interpretation of data.U, if chart junk is interfering with ease of understanding and interpretation of data.						
	Ease of Understanding-Take home message <ul style="list-style-type: none">E, if data trends are easy to observe or no trends are apparent.NI, if data trends are difficult to observe.U, if the graph is ineffective at communicating data trends and if it causes confusion.						
Graph Choice	Graph Type (Bar, line, scatter, dot, box and whisker) <ul style="list-style-type: none">E, for displaying data in a graph that is appropriate for both independent and dependent experimental variables (i.e. categorical and continuous).NI, for displaying data in a graph that is not suitable for either the dependent or independent experimental variable.U, if the graph type is not suitable for the experimental variables.						
	Data Displayed (Raw, Averages, Changes, Percentage) <ul style="list-style-type: none">E, if the graph indicates the type of data (ex. Raw, averages, etc.) that are plotted. If the graph is showing averages, then it should also be accompanied with STDEV or Error bars.NI, if the graph is missing one of points mentioned above.U, if data type is inappropriate for the graph type.						
	Alignment (<i>at least one of the graphs presented should align with the research question and hypothesis. Other graphs can be exploratory.</i>) <ul style="list-style-type: none">E, if the graph is aligned with the research question and hypothesis.NI, if the graph is partially aligned with the research question and/or hypothesis.U, if the graph is not aligned with the research question and/or hypothesis.						
	Total Points for Graph Choice and Communication:	/10pts					
Overall Graph Presentation Grade:		/13.5 pts					
Angra A and Gardner SM (2							

- **Communication**
- Aesthetics
 - Take home message

- **Graph Choice**
- Graph Type
 - Data
 - Alignment with RQ and HYP (and Prediction)

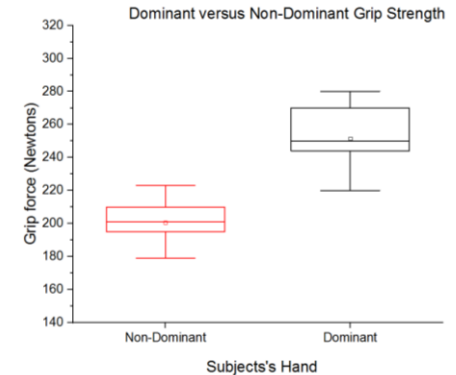
Example activity: Graph evaluation using the graph rubric

Goals of activity:

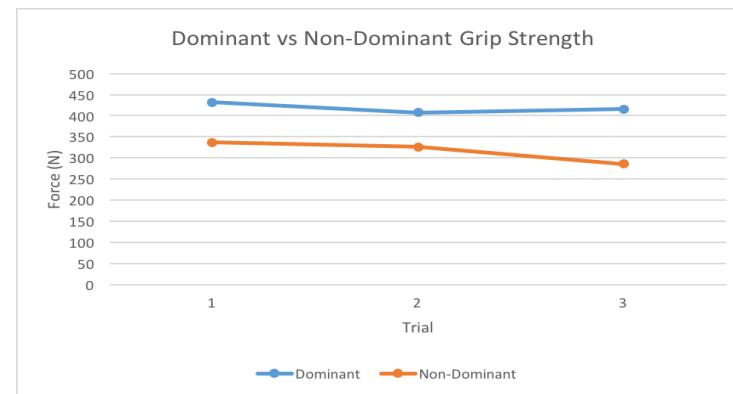
- Allow students to become familiar with the graph rubric by using it
- Reveal areas of competence and challenges
- Guided discussion allows for calibration of expectations, clarification of questions, and encouraging a critical and reflective approach

The quality of student graphs improve over the course of the semester especially in graph type and alignment criteria (Gardner et al., 2022 (in press))

Hypothesis/Prediction: You use your dominant hand more in day-to-day life so the muscles in your dominant hand are likely stronger which would lead to a greater grip force compared to your non-dominant hand.



Prediction: If a subject uses their dominant hand rather than their non-dominant hand, then it is predicted that the dominant hand will produce more grip force.



Important limitations of our previous work

- Participants all from the same institution
- Graphing behaviors with pen-and-paper may be different than in a digital format
- Labor intensive to evaluate
- Participants were given a dataset with which to work
- Bacterial and plant growth scenarios:
 - Contrived scenario
 - Research question and hypothesis are not explicit
 - Simple data set with two, relevant variables
 - Time as a variable led to the predominance of line graphs

	Number of Cells					
Time (min)	22 °c			10°C		
	Tube 1	Tube 2	Tube 3	Tube 1	Tube 2	Tube 3
0	2	2	1	2	1	2
30	4	4	3	2	2	3
60	6	8	6	2	2	3
90	12	16	12	2	3	4
120	24	30	22	4	5	6

Grappling with Graphs – Collaborative project



#1726180

Stephanie Gardner, PI



Joel Abraham, Co - PI



Eli Meir, Co - PI



Overarching research goal

- Develop **digital assessment modules** that can be used **to reveal student knowledge and skill** and **be used without the need for manual grading of individual work** in classroom settings of all sizes.



Elizabeth Suazo-Flores



Anupriya Karippadath

GraphSmarts biological scenario

How do MPAs Affect the Food Chain?

Scientists in Australia have been tracking lobster, urchin, and kelp abundance, as well as lobster fishing patterns, in the MPA and non-MPA areas of coastal Tasmania. As part of this larger project, scientists reasoned that stopping lobster fishing would increase lobster predation on urchins and therefore reduce the number of urchins in the kelp forest. Their reasoning is illustrated here:



Students are presented an overarching hypothesis and specific predictions to evaluate through graphing

Study Plot ID	Month Sampled	MPA Status	Lobster Density (#/m ²)	Average Lobster Size (g)	Urchin Density (#/m ²)	Kelp Abundance Score
1	Aug.	YES	1.10	410	9.5	HIGH
2	Sept.	YES	1.55	445	8.5	MED
3	Aug.	NO	1.15	350	12.0	MED
4	Oct.	YES	2.00	435	7.0	MED
5	Aug.	NO	0.75	385	9.5	MED
6	Oct.	NO	1.05	355	11.0	LOW

- Two conditions (MPA yes/no)
- Nine replicates (plots) for each MPA condition (18 total plots)
- Irrelevant/less relevant variables to test the prediction

Student model for graph construction in biology - Handout

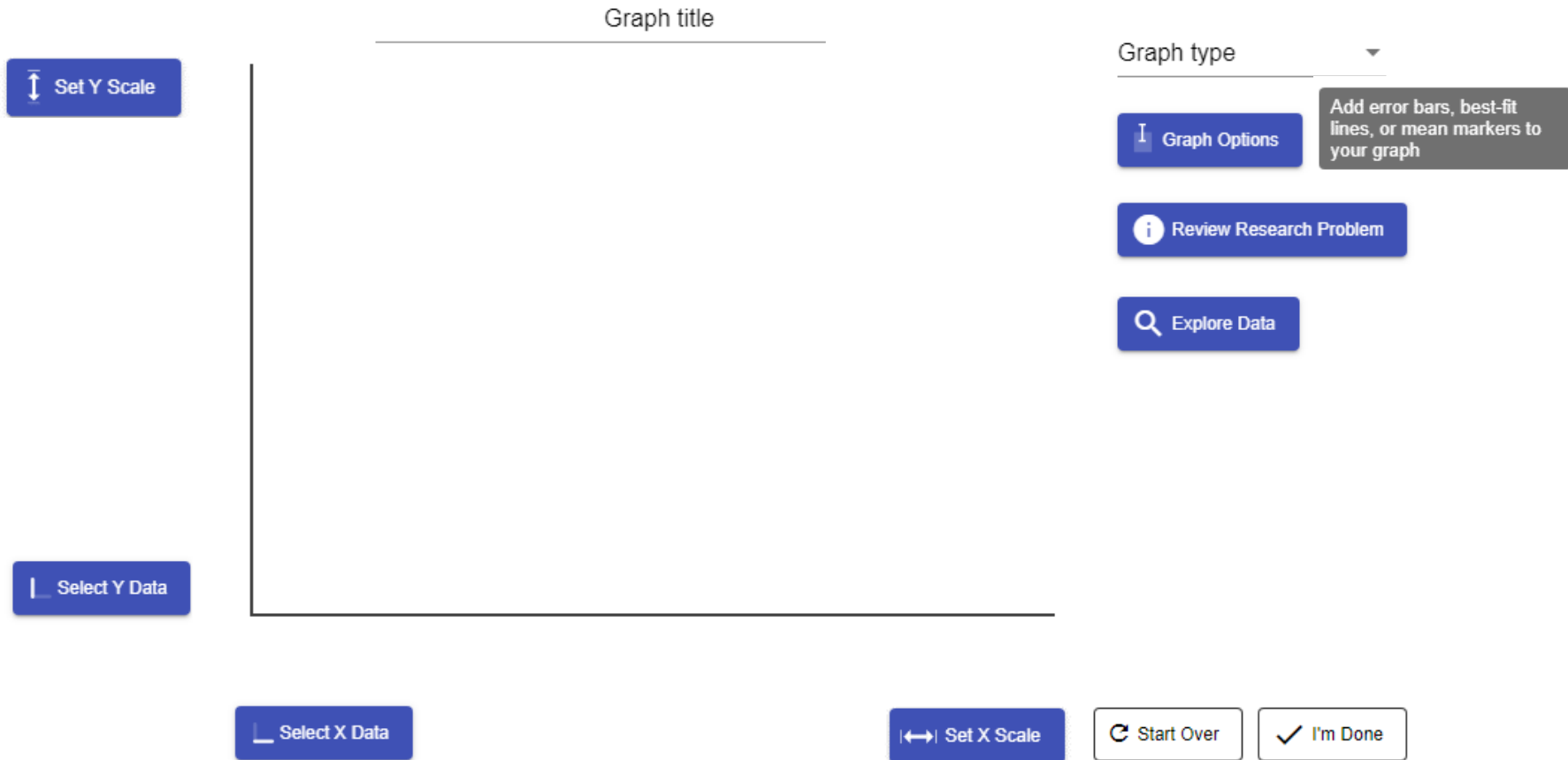
Graph Construction Conceptual Model (GCCM)

Category
<i>Data Selection</i>
<i>Data Exploration</i>
<i>Graph Assembly</i>
<i>Graph Reflection</i>

- Based on:
 - Literature review
 - Our own experiences teaching graphing
 - Our research on graphing
 - Focus groups with biology ($n = 5$) and statistics education ($n = 3$) researchers and educators

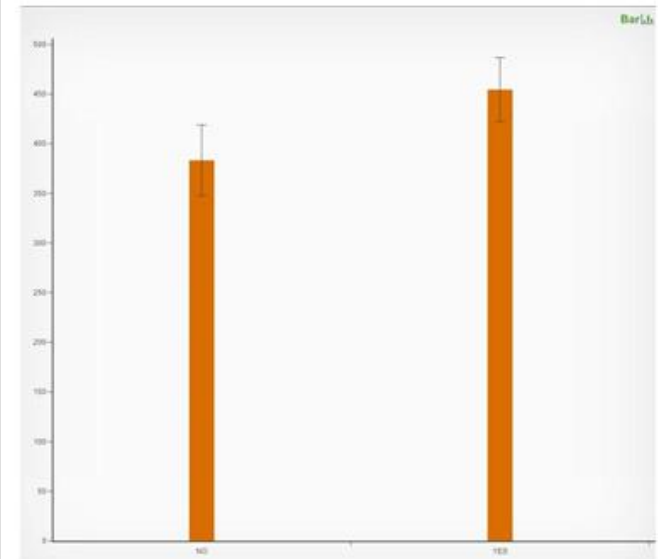
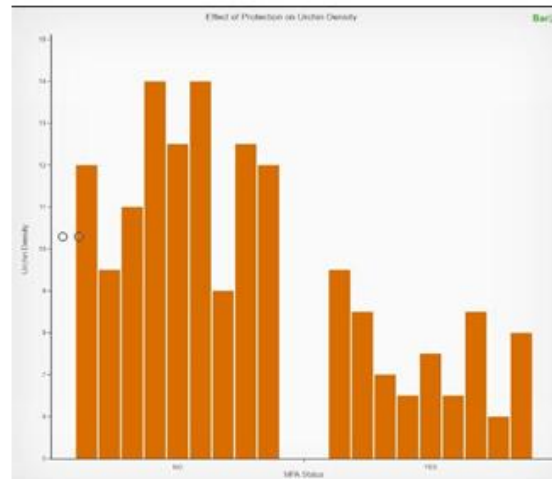
Graphing interface

Graphing interface is simple and somewhat constrained

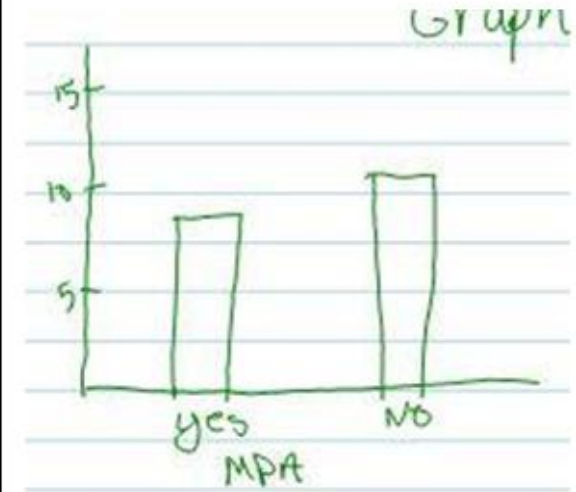
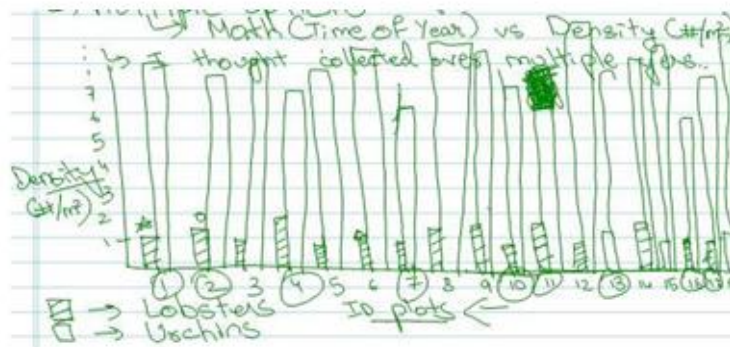


Types of graphs generated from GraphSmarts scenario

GraphSmarts



Pen-and-paper



Important insights from GraphSmarts vs. Pen-and-paper

- **Data from GraphSmarts assessment corroborates many of our previous findings with pen-and-paper graphing**
 - Similar reasoning for graph choice and construction
 - Students rarely aggregate data (i.e. Mean) and when they do they don't include a measure of variability (e.g. S.D.)
- **GraphSmarts...examples of interesting findings**
 - Students are forced to focus on plotting a small subset of variables
 - More students stated the hypothesis/prediction as reasons for variable selection and graph choice
 - Selection of relevant variables for testing a prediction is a challenge for many students
 - Students can quickly iterate on their graphs and try things out (is this necessarily always a good thing?)

Recommendations based on our data and experience

- **Provide targeted instruction around specific graphing practices:**
 - Identifying relevant variables to answer questions
 - Graph choice
 - Data form (i.e. raw or summarized data) and representing variability
- **Engage students in graphing activities and assessments that allow them to develop and demonstrate their competence**
 - Our evidence-based frameworks can help reveal student competence and scaffold learning
- **Consider the pros and cons of having students graph by hand vs. with digital graphing tools**
- **Encourage students to reflect throughout the process of graphing**

Thank you!



Photo credit: Autumn Siebold

Current group members

- Anupriya Karippadath
- Sharleen Flowers
- Nouran Amin

Former group members

- Dr. Elizabeth Suazo-Flores
- Dr. Aakanksha Angra – Georgia State University
- Jackie Mercader, Alexandrou Ivan, Mary Welker, Kenya Lee, Riley Stehr (undergraduate researchers)
- Mozhu Li, M.S. – IU School of Medicine

- Biology Education Area in the Dept. of Biol. Sciences at Purdue
- Purdue International Biology Education Research Group (PIBERG)

Graphing collaborators

- Dr. Joel Abraham – CSU Fullerton
- Dr. Ryan Baker- UPenn
- **SimBiotic Software**
 - Dr. Eli Meir
 - Susan Maruca
 - Dr. Kerry Kim
 - Stephen Allison-Bunnell
 - Joshua Quick
- Dr. Joe Harsh – James Madison University

ACE-Bio

- Dr. Nancy Pelaez
- Dr. Trevor Anderson
- Dr. Yue Yin

THANK YOU

Stephanie M. Gardner
Purdue University
sgardne@purdue.edu