FROM THOUGHT TO PLOT: REVEALING UNDERGRADUATE BIOLOGY STUDENT GRAPHING PRACTICES

SimBio Webinar

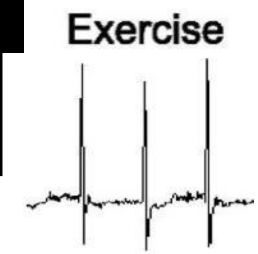
March 3rd , 2022

Artistic credit: Autumn Siebold

PURDUE UNIVERSITY

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

What would you do?



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

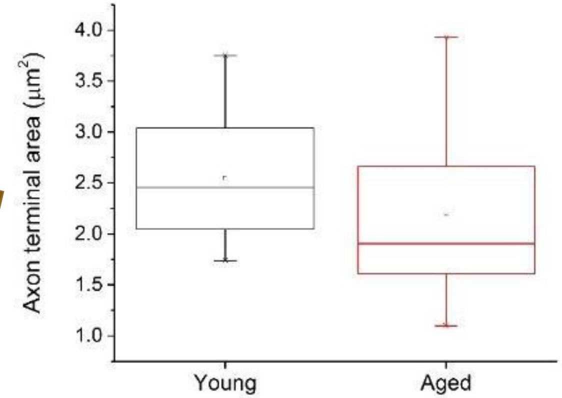
In my Physiology class students design and execute experiments and analyze and present their data.

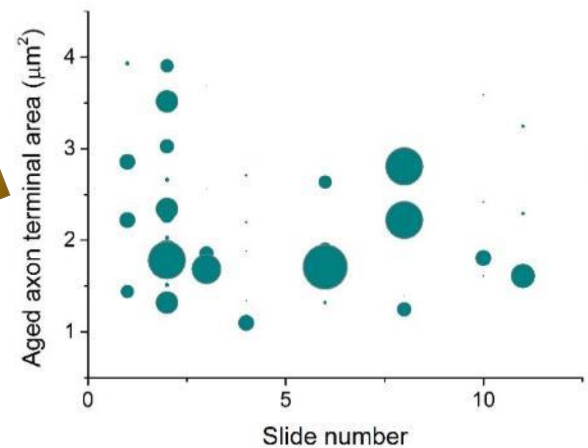
Function of graphs (Tufte, 1983)

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

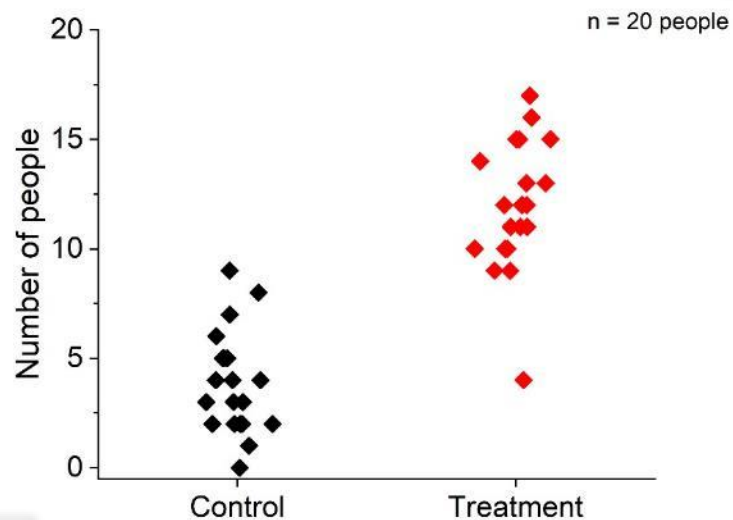




Basic graph construction knowledge and skills

Graph types

Math and statistics



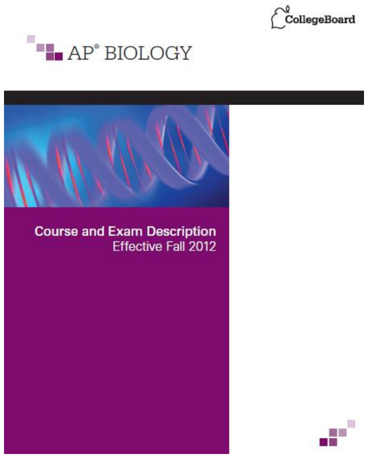
Concepts and practices of the discipline

Visuospatial and aesthetics

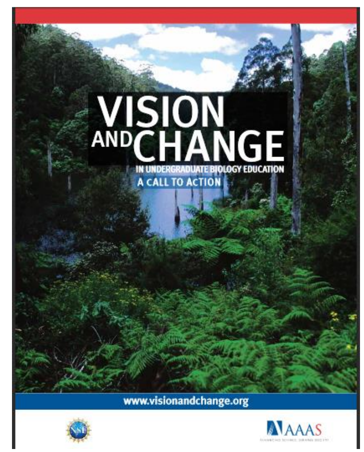
Novick, 2004; Grawemeyer and Cox, 2004; diSessa, et al., 1991; diSessa and Sheron, 2000; diSessa, 2004; Berg and Smith, 1994; Friel and Bright, 1996; Konold et al., 2015; Kellman, 2010; reviewed by Hegarty, 2011; Franconeri et al., 2012

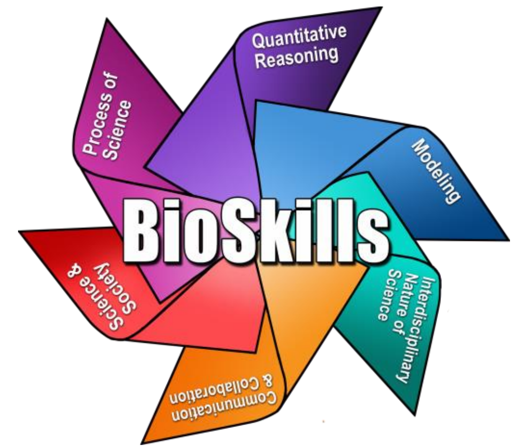
Undergraduate student graphing

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



NEXT GENERATION SCIENCE STANDARDS
For States, By States

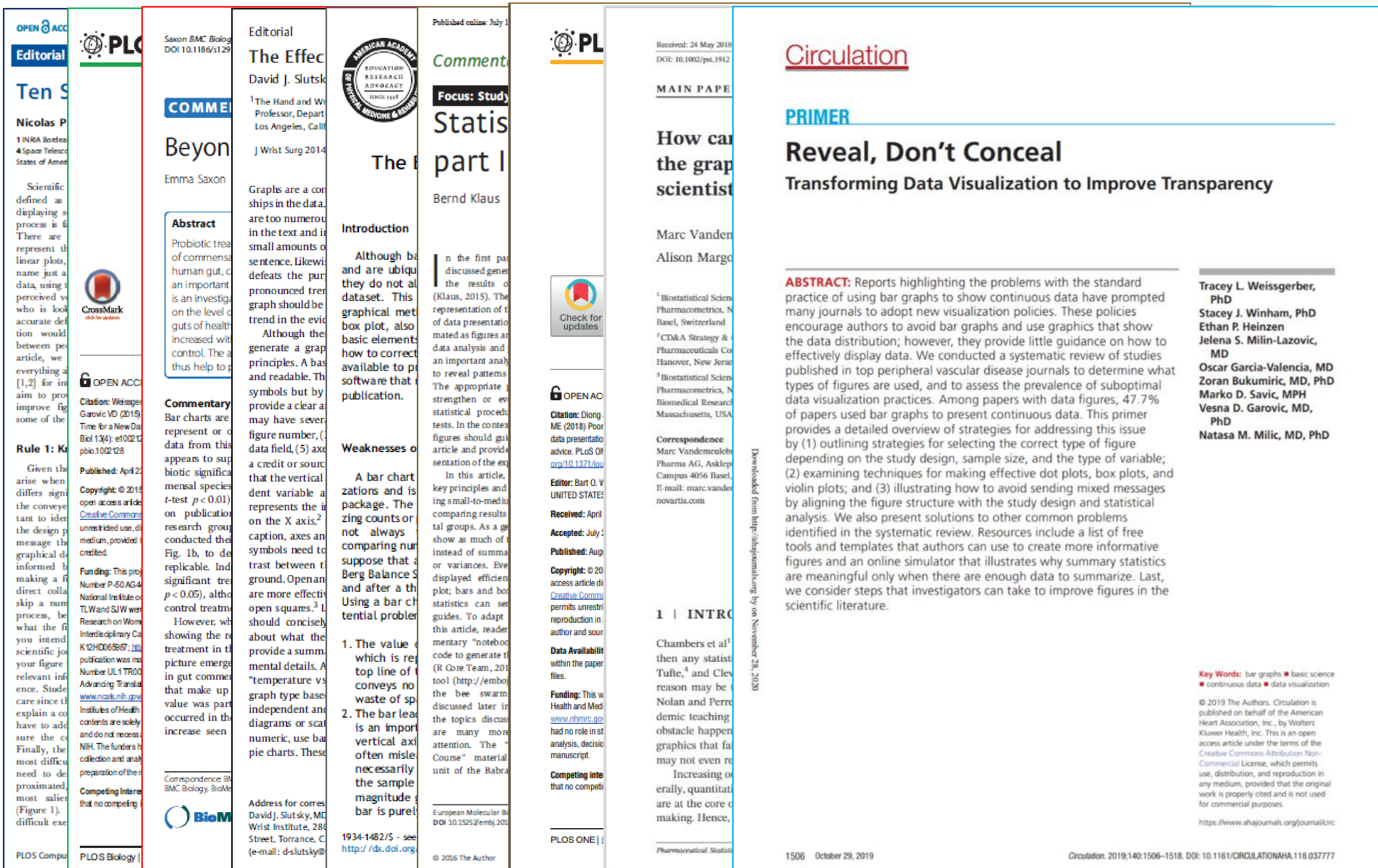




- Students will be making decisions about data analysis and representations
- However, more emphasis has been placed on interpreting graphs and not creating them (e.g. AP Biology up to 2012)
- Students are competent plotting points and identifying coordinates
- Students struggle with important concepts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

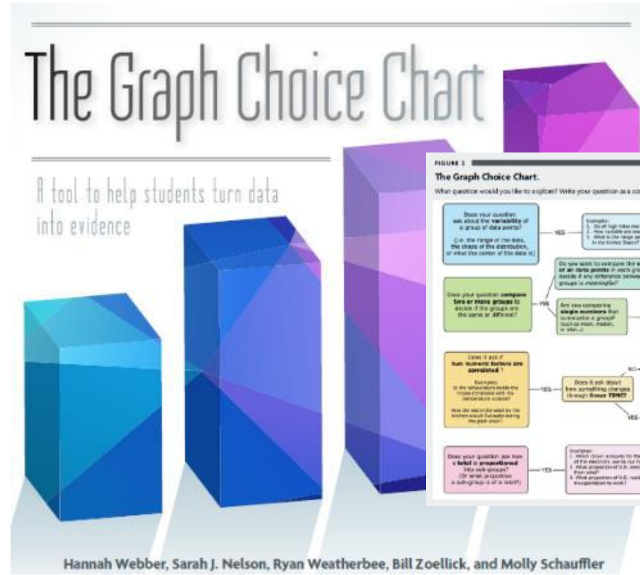
Padilla et al. ,1986; Brasell & Rowe, 1993; Berg and Smith, 1994; Ainley, 1995; Mevarech & Kramarsky, 1997; Mathewson, 1999; Grunwald and Garfield,2003; Bakker, 2004; Leonard & Patterson, 2004; Clase et al., 2010; Hartman, 2010; Tairab & Al Naqbi, 2010; Meletiou & Lee, 2010; Bray-Speth et al., 2010;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

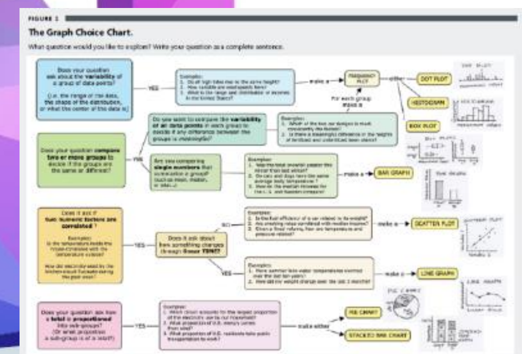
A call for better representations by practitioners (2014 -present)

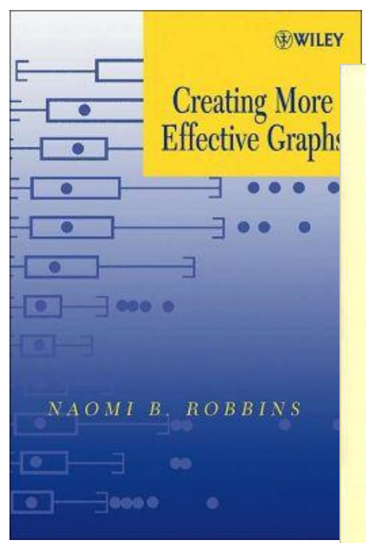


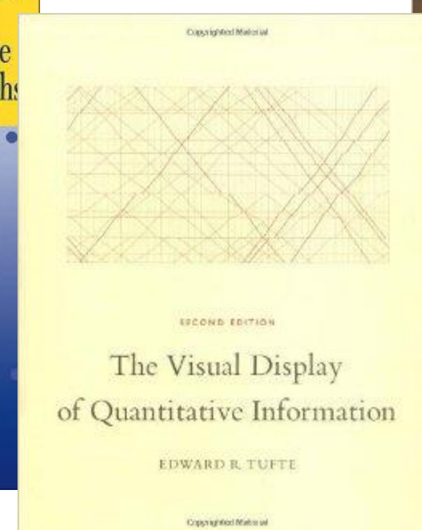
Existing graph choice and construction resources

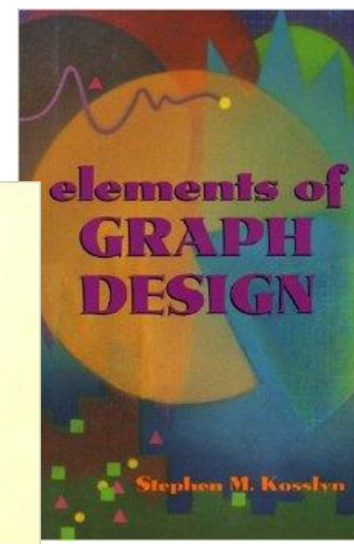
- Generic guidelines on proper graph choice and construction
- Materials not situated in the context of experiments for science/biology
- No empirical investigation on reasoning with graph choice and construction











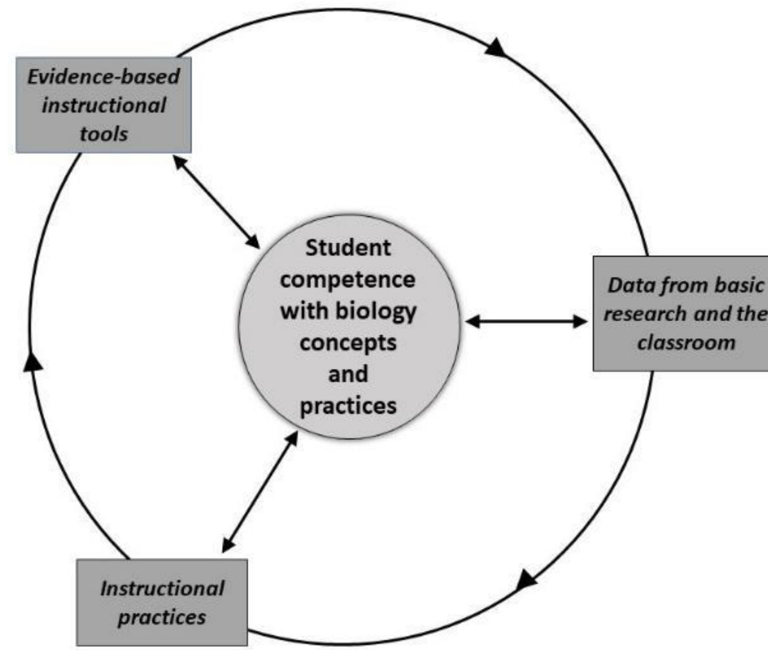
Motivating question and objectives

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).

Early insights from clinical interviews

Aakanksha Angra

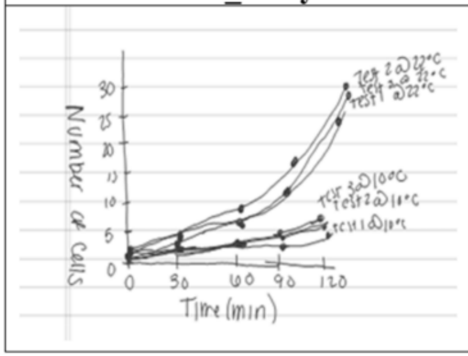
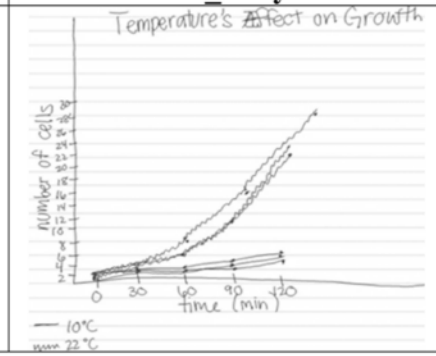
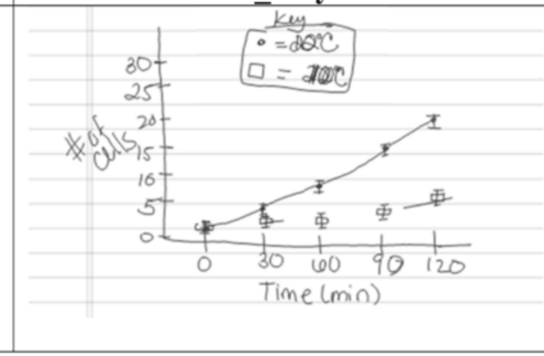
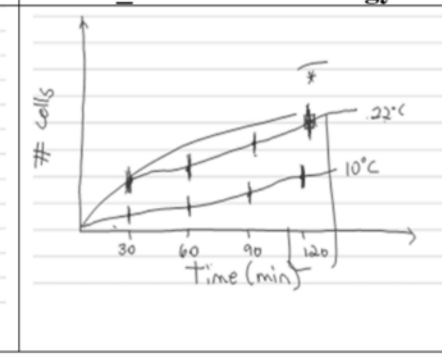
- Semi-structured think-aloud interviews
  - Pen-and-paper graph choice and construction task

- 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)

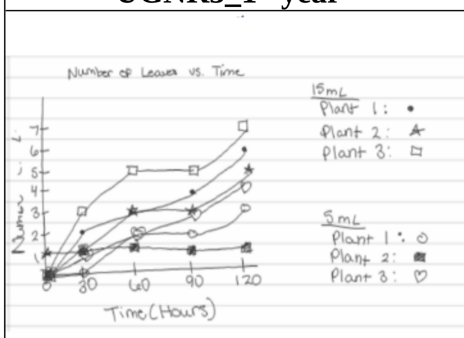
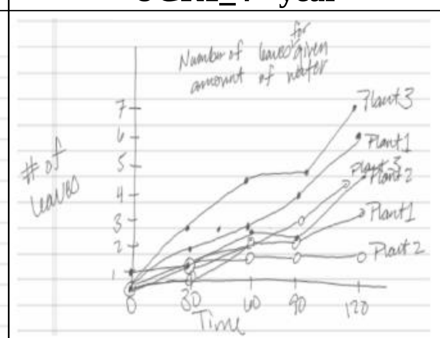
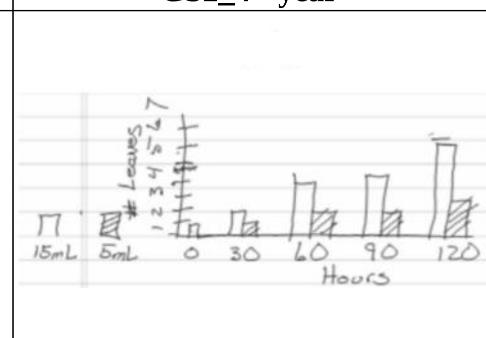
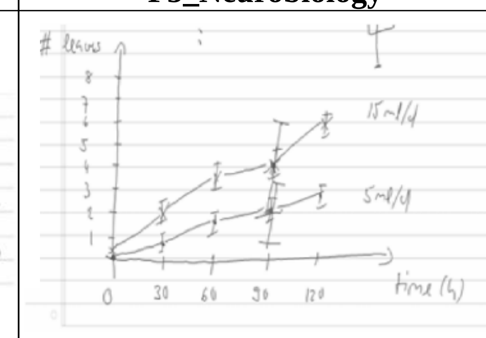
Angra and Gardner, 2016 Advances in Physiology Education
Angra and Gardner, 2017 CBE-Life Sciences Education

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
			

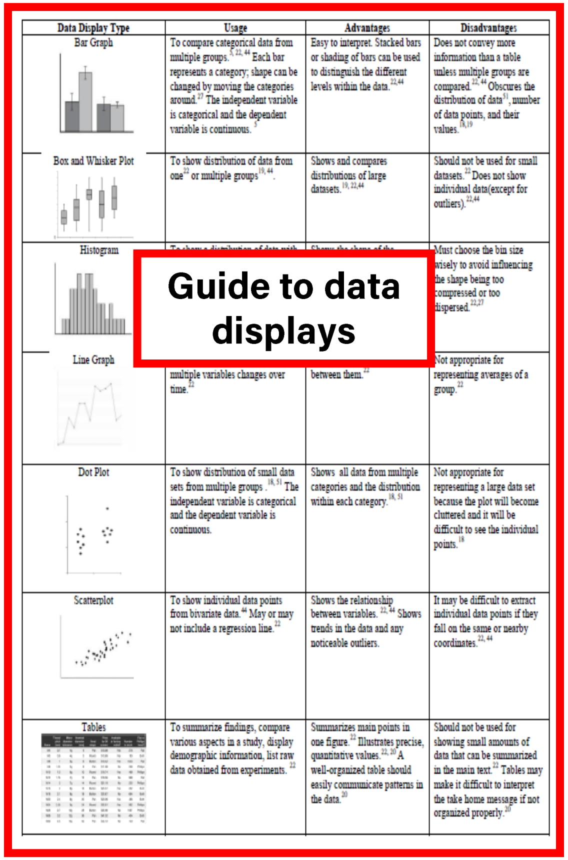
Angra and Gardner, 2017 CBE-Life Sciences Education

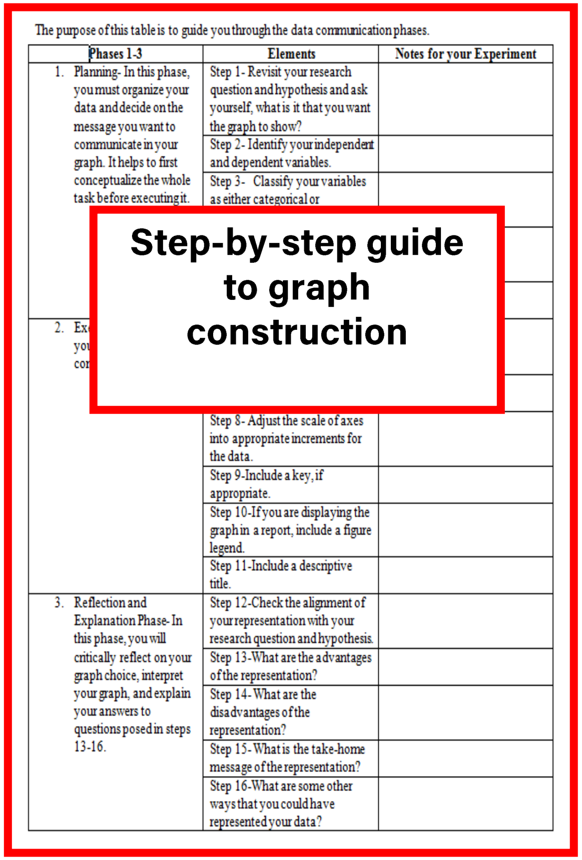
- *Minimal to no planning prior to construction*
- *Meticulous graph construction*
- *Raw data often used in graphs*
- *Data considered independent of experimental context*
- *Variability in experimental data not considered*
- *Graph choice not based on question and/or hypothesis*
- *Intuitive reasoning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

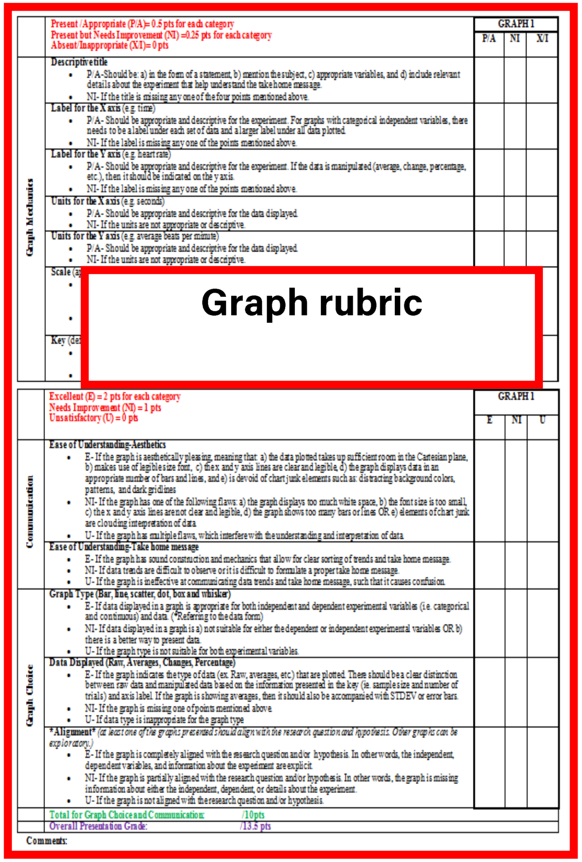
Angra and Gardner, 2017 CBE-Life Sciences Education

Guides to Improve Graph Choice and Construction

I use these with students in all my classes!







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

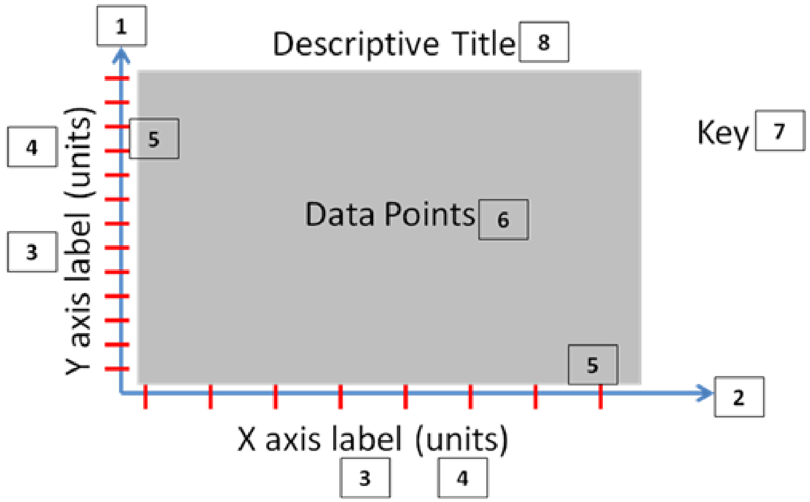
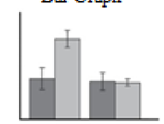
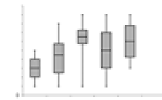
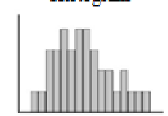
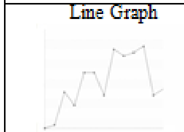
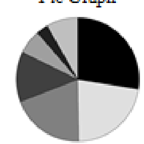
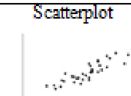
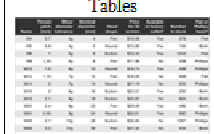


Table 1- Summaries the common types of graphs, their usage, advantages, and disadvantages.

Graph Type	Usage	Advantages	Disadvantages
Bar Graph 	To compare categorical data from one or multiple groups.[1,5] 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.

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.

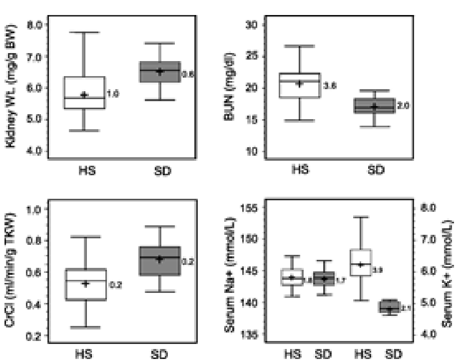


Figure 5. Example of a box plot (From: http://ajprenal.physiology.org/cgi/content-nw/full/298/6/F1484/F2)

The Step-by-step guide

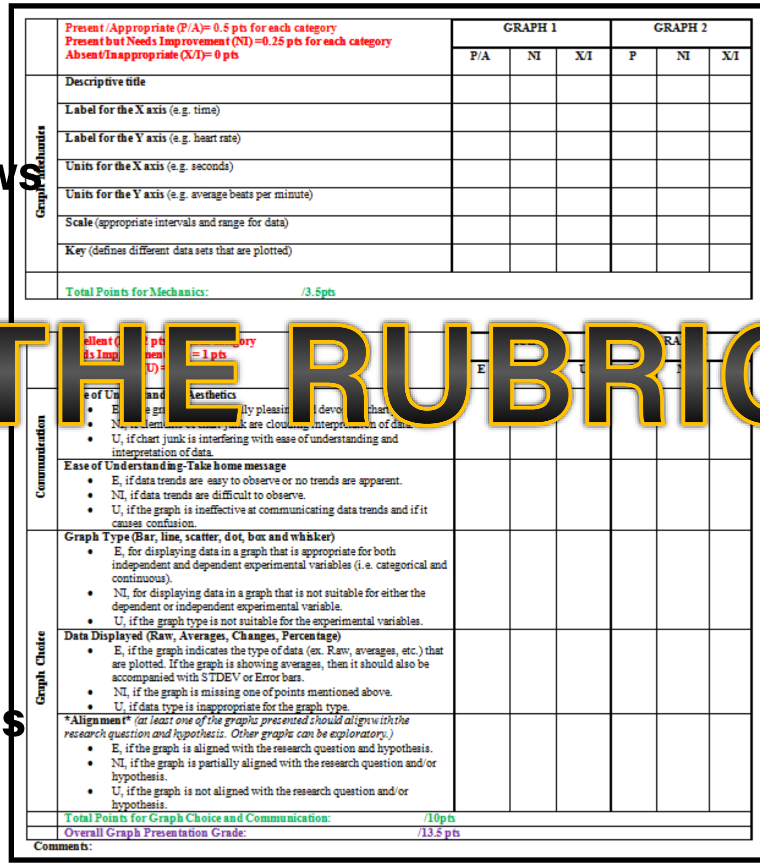
Phases 1-3	Elements	Notes for your Experiment
1. Planning- In this phase, you must organize your data and decide on the message you want to communicate in your graph. It helps to first conceptualize the whole task before executing it.	Step 1- Revisit your research question and hypothesis and ask yourself, 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

Angra and Gardner (2016) Advances in Physiology Education

Expert-Novice Interviews

THE RUBRIC

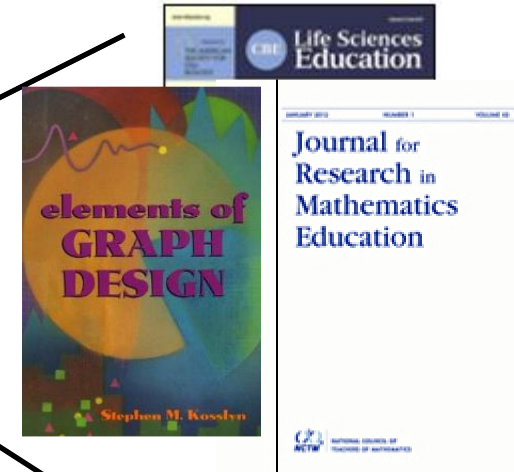


	Present /Appropriate (P/A)= 0.5 pts for each category Present but Needs Improvement (NI) =0.25 pts for each category Absent/Inappropriate (X/I)= 0 pts	GRAPH 1			GRAPH 2		
		P/A	NI	X/I	P	NI	X/I
Graph Mechanics	Descriptive title						
	Label for the X axis (e.g. time)						
	Label for the Y axis (e.g. 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						

	[illegible]ellent ([illegible] pts [illegible]ory [illegible]ds Imp[illegible]ent[illegible] =1 pts [illegible]U) =	[illegible]			[illegible]RA[illegible]		
		E	[illegible]	U	[illegible]	N[illegible]	[illegible]
Communication	[illegible]e of Un[illegible]and[illegible] Aesthetics • E[illegible] e gr[illegible]lly pleasi[illegible] devo[illegible] char[illegible] • N[illegible] k are clou[illegible] interp[illegible] of data • U, if chart junk is interfering with ease of understanding and interpretation of data.						
	Ease of Understanding-Take home message • 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) • 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) • 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 *(at least one of the graphs presented should align with the research question and hypothesis. Other graphs can be exploratory.)* • 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						

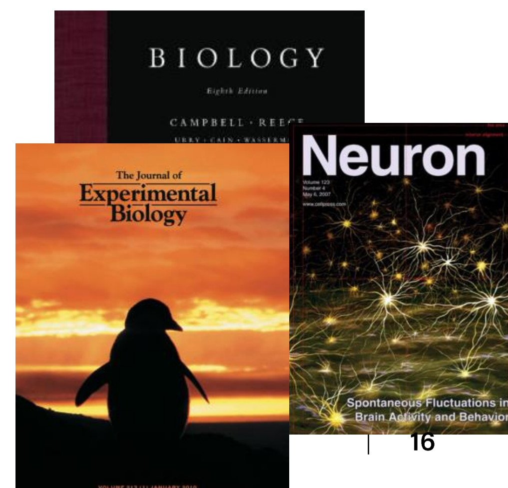
Comments:

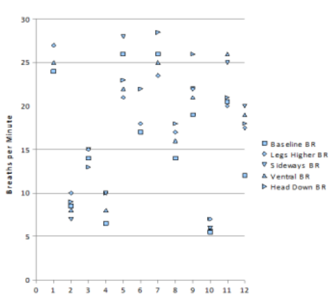
Graphing Literature



Student Reflections

Graphs in Textbooks And Primary Literature





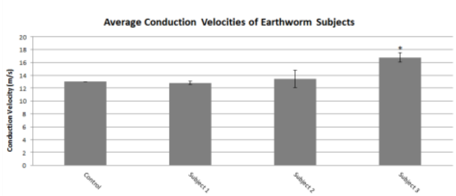


Fig. 1. The average conduction velocities of the median giant axon for the following earthworm subjects: Control at 20°C, and Subjects 1-3 at 25°C. Subject 3 has a significantly larger ($P < 0.05$) average conduction velocity when compared (Student T-Test) to the average conduction velocity of both Subject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

Student Graphs

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** • 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 • 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)** • 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) • 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 *(at least one of the graphs presented should align with the research question and hypothesis. Other graphs can be exploratory.)* • 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:

→ **Communication**

- Aesthetics
- Take home message

→ **Graph Choice**

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

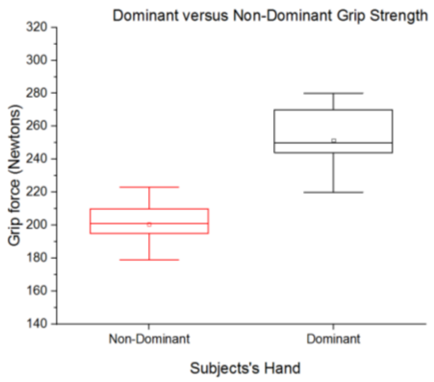
Angra A and Gardner SM (2018) The Graph Rubric: Development of a Teaching, Learning, and Research Tool. CBE-Life Sciences Education 17: ar65 doi: 1187/cbe.18-01-0007

Example activity: Graph evaluation using the graph rubric

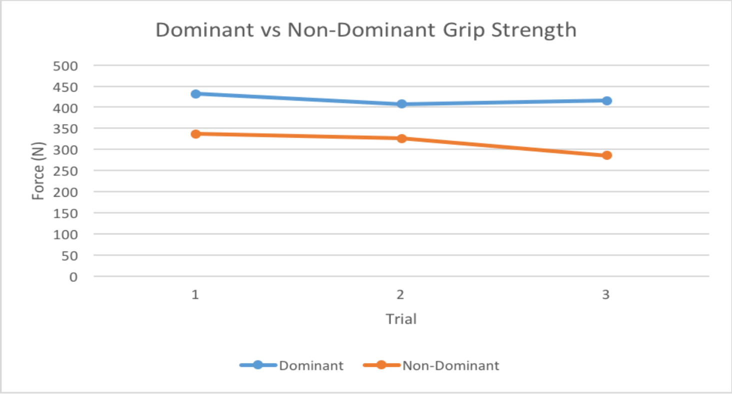
Goals of activity:
- Allow students to become familiar with the graph rubric by using it
- Reveal areas of competence and challenges
- Guided discussion allows for calibration of expectations, clarification of questions, and encouraging a critical and reflective approach

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

PURDUE UNIVERSITY

Joel Abraham, Co - PI

CALIFORNIA STATE UNIVERSITY FULLERTON 1957

Eli Meir, Co - PI

SimBio
Simulated Biology. Real Learning.

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

- 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)

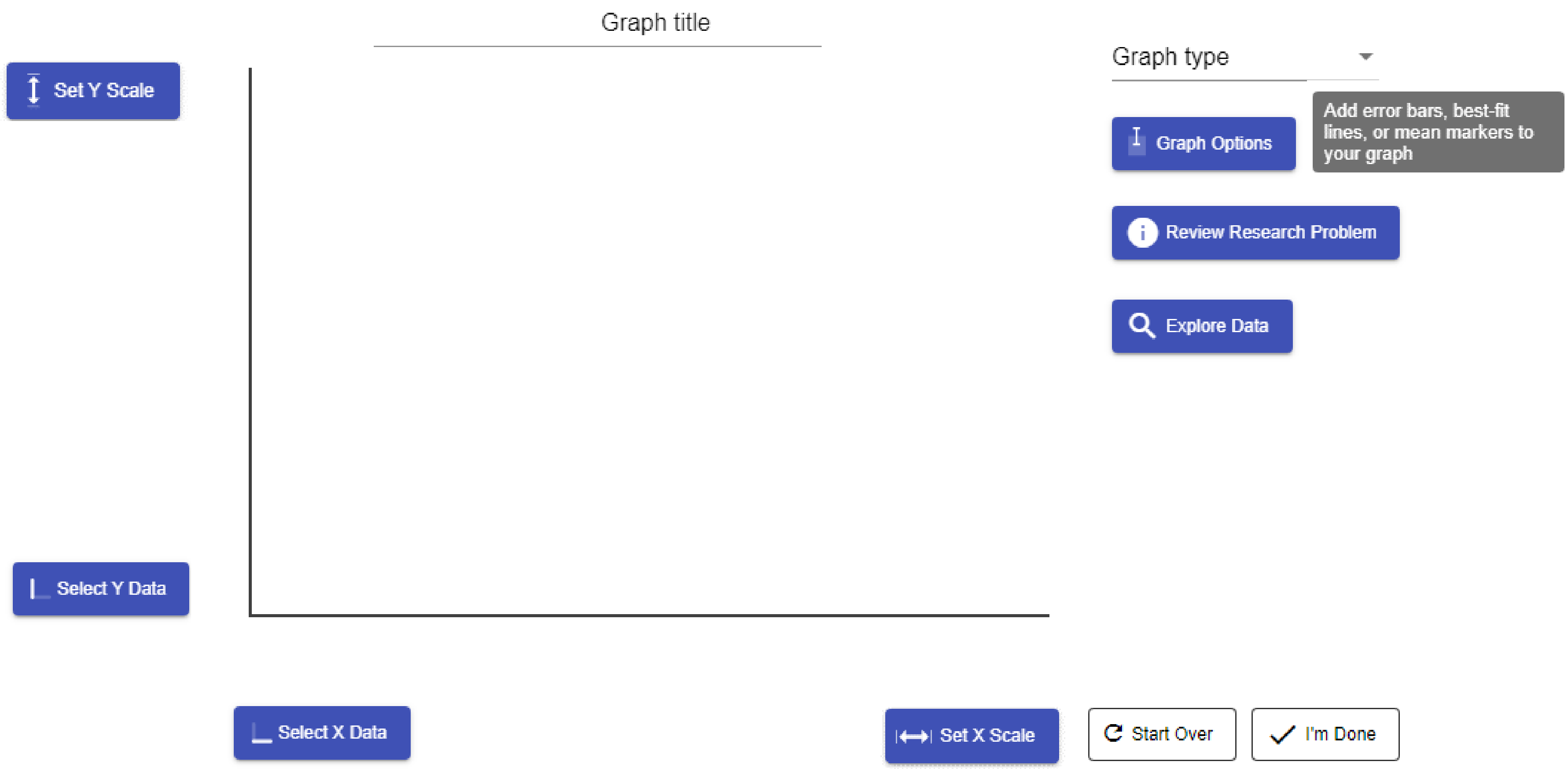
Category
Data Selection
Data Exploration
Graph Assembly
Graph Reflection

- 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

Abraham et al., (in preparation)

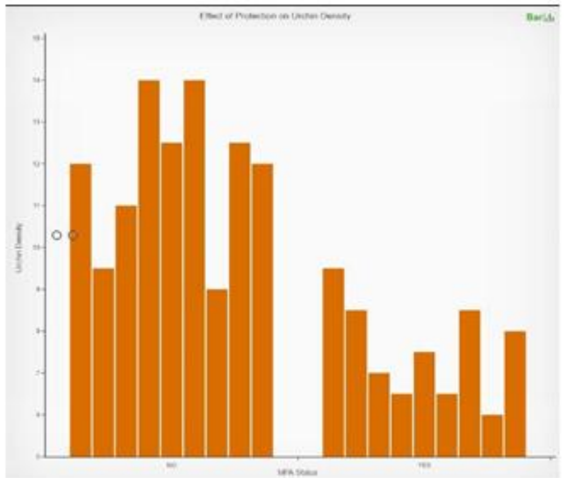
Graphing interface

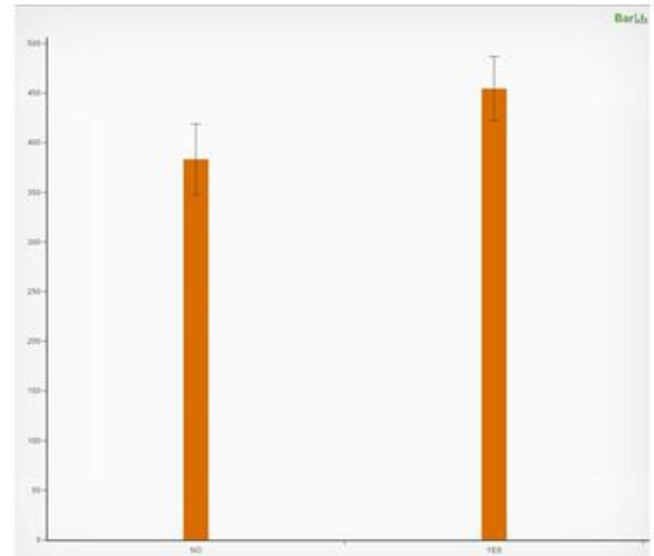
Graphing interface is simple and somewhat constrained



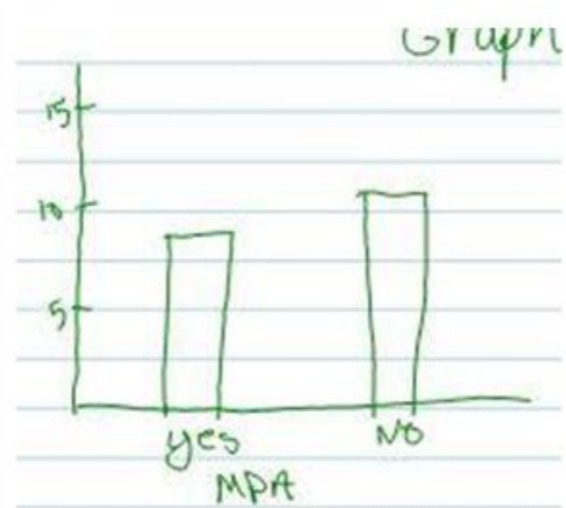
Types of graphs generated from GraphSmarts scenario

GraphSmarts





Pen-and-paper



Gardner et al., 2021 J. Science Ed and Tech

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

PURDUE UNIVERSITY

Biology Scholars

NSF

1346567
1726180

THANK YOU

Stephanie M. Gardner
Purdue University
sgardne@purdue.edu

PURDUE UNIVERSITY