Purdue University
Purdue e-Pubs

Open Access Theses

Theses and Dissertations

January 2016

DIFFERENCES IN CONTEXT: REVEALING EXPERT-NOVICE GRAPH KNOWLEDGE IN BIOLOGY

Mozhu Li
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Mozhu Li

Entitled
DIFFERENCES IN CONTEXT: REVEALING EXPERT-NOVICE GRAPH KNOWLEDGE IN BIOLOGY

For the degree of Master of Science

Is approved by the final examining committee:

Stephanie M. Gardner
Chair

Nancy J. Pelaez

Dennis J. Minchella

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Stephanie M. Gardner

Approved by: Daoguo Zhou 7/19/2016

Head of the Departmental Graduate Program Date

DIFFERENCES IN CONTEXT: REVEALING EXPERT-NOVICE GRAPH KNOWLEDGE IN BIOLOGY

A Thesis

Submitted to the Faculty

of

Purdue University

by

Mozhu Li

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2016

Purdue University

West Lafayette, Indiana

For my parents and husband

ACKNOWLEDGEMENTS

I would like to first express my deep gratitude to my thesis advisor, Dr. Stephanie M. Gardner, for her constant guidance and encouragement. Dr. Gardner has supported throughout my thesis with her patience and knowledge, while allowing me to work in my own way. She spent time communicating with me every week, and she was always willing to help whenever I had questions about my research. I could not have imagined having a better advisor and mentor for my Master’s study.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Nancy Pelaez, and Dr. Dennis J. Minchella, for their interest in my research, their insightful comments, and their encouragement.

My sincere thanks goes to my fellow student Aakanksha Angra, for introducing me to the topic as well for her support on the way. Also, I would like to thank my friends and colleagues in the Purdue International Biology Education Research Group (PIBERG) and Discipline-Based Education Research Group (DBER-GS), especially Matthew Lira, Maurina Aranda, and Richard Lie, for enlightening me with their precious suggestions on my work. I would like to thank Dr. Bruce A. Craig and Yunfan Li from the Purdue Statistical Consulting Service for helping me with the quantitative data analysis and visual presentation. I would like to thank the participants in our study, who have willing shared their precious time during the process of interviewing.

Finally, I would like to thank my parents, Hongbo Li and Jie Dong, and my husband, Jerry Chen, both by keeping me harmonious and supporting me throughout the way.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Li, Mozhu. M.S. Purdue University, August 2016. Differences in Context: Revealing Expert-Novice Graph Knowledge in Biology. Major Professor: Stephanie M. Gardner.

Graphs are typically defined as visual representations that depict and sometimes summarize quantitative data. Visual representation of quantitative data is broadly used in scientific textbooks, papers, and lectures as well as popular media seen in everyday life. Thus, understanding graphs and data became an essential skill for all students to master. However, correctly and fully using graphs requires a person to have multiple competencies (diSessa & Sherin, 2000). For an instance, during graph construction, variables need to be identified and characterized, data are screened and often reduced, and a graph type needs to be chosen that is appropriate for the data. Student have difficulties interpreting and constructing scientific graphs (Beichner, 1994; Mevarech, & Karamarsky, 1997; Shaw, Padilla, & Mckenzie, 1983; Speth et al., 2010); in spite of using some of the documented difficulties to improve instruction, difficulties persist for undergraduate science students (Speth et al., 2010; McFarland, 2010). We aim to compare the differences in graph knowledge among undergraduate biology students, graduate biology students, and biology professors. Using the results, we hope to better understand and define the role that graph knowledge plays in students' ability to choose and create appropriate graphs from data. This will be beneficial to instructors who teach analytical and graphical skills at school and to educators who design the curriculum with a purpose of effective teaching and learning.

CHAPTER 1: INTRODUCTION TO RESEARCH QUESTION

1.1 Importance of Graphs as Visual Representations

Graphs are visual representations that depict, and sometimes summarize, quantitative data. Rougier, Droettboom, and Bourne (2014) provided a more accurate definition of graphs by stating that scientific visualization is a graphical interface between people and data. Graphs are used in lectures by professors and in textbooks by authors to explain important concepts to students (Treagust & Tsui, 2013). They are also used frequently in scientific papers to present complicated data and to show trends and ideas (Treagust & Tsui, 2013). It is also very common to see graphs in our daily lives, such as weather and stock market reports.

Graphs are an essential part of scientific communication among scientists in research manuscripts in scientific journals. Cleveland (1984) measured the amount of graph usage in 57 scientific journals and found that around one-third of the space of science journals are devoted to graphs. He measured the graph area by the amount of text replaced by a graph, and a figure was judged to be a graph if it had scales and conveyed quantitative information. A more recent study targeting journals in the medical field showed a very similar result (Cooper, Schriger, and Tashman, 2001). In the study, the three authors performed a blinded review of all graphs published in Annals of Emergency Medicine from January 1998 to June 1999. Out of the 147 original research communications, 46% contained at least one graph.

From the results of these studies, one can see that graphical representations in science publications and presentations are very common and critically important, especially in scientific research with quantitative methods. With biology education focused more on students learning content within the context of the practices of scientists (see section, 1.3, below), it is particularly critical for college-level science students to be able to read graphs, understand data behind graphs, and interpret the messages that the graphs want to convey. However, neither constructing graphs nor interpreting graphs is an easy task, as even expert scientists can have difficulty. A series of study were conducted on scientists and professors who were experts in their field, mostly in life science and physical science (Roth & Thom, 2009; Roth, 2013). They found that even these experts had problems interpreting introductory graphs from other fields or graphs within their own disciplines but outside their immediate area of expertise. In a recent review of research articles that are published in top physiology journals, researchers assert that scientists need to improve data presentation using more complete representations. They especially emphasized a number of critical problems within the presentation of continuous data in small sample size studies (Weissgerber, Milic, Winham, & Garovic, 2015).

1.2 Graph Knowledge

Proper graph construction and evaluation requires knowledge and skills from various disciplines and practices. diSessa (2000, 2004) provided a new concept: MetaRepresentational Competence, or MRC. He pointed out that graphing is not dichromatic: it is not a “yes/no” question. MRC stands for the full range of capabilities that people need to have in order to construct and use external representations. In his opinion, in order to reach a deep, rich, and generative understanding of graphs, a student would need to be able to do such things as: “Invent or design new representations”, “Critique and compare the adequacy of representations and judge their suitability for various tasks”, “Understand the purposes of representations generally and in particular contexts and understand how representations do the work they do for us”, “Explain representations, i.e., the ability to articulate their competence with the preceding items”, and “Learn new representations quickly and with minimal instruction”. In light of the MRC framework, we believe that to correctly and fully use graphs, the graph creator not only needs knowledge of graphs, but also needs to be familiar with statistical knowledge, spatio-visuo knowledge, and disciplinary-related knowledge and experiences. *Graph knowledge is defined as the knowledge a person has about a type of graph including its name, its function, data that can be displayed with it, its affordances and limitations*.

In diSessa's study (2004), he pointed out that MRC should not be treated as "hard-wired" abilities. Instead, MRC is gradually developed through practices both in and out of school. Thus, one's graph knowledge should be highly related to one's inner knowledge, or knowledge one has gained from past experiences. In other words, MRC is not only a competence with representations, but also the reflective aspects about the representations, their creation, and usage. Implicit in the MRC framework is expert-like competence that comes with experience. In the case of graphs, this would be competence with disciplinary inquiry, statistics, and knowledge of graphical representations.

1.3 College-level Biology Students as Targets

It is both important and valuable to study college-level biology students' graph knowledge. Undergraduate biology education reforms have called for students to be involved in the practices of science in classes, course-based undergraduate research, and research apprenticeships (AAMC-HHMI, 2010; AAAS, 2011; PCAST, 2012). A number of studies also mentioned that an increasing number of undergraduate students were engaging in biological research in order to meet the increasing stringent academic criteria, to get into graduate or professional schools, to become competitive in employment upon graduate, or for a variety of other reasons (Dasgupta, Anderson, & Pelaez, 2014; Wei, & Woodin, 2011, Laursen, Hunter, Seymour, Thiry, & Melton, 2010). Getting involved with research requires the students to be familiar with data and data representations, or graphs. While students across disciplines and education-levels are having various difficulties with graphing, we decided to particularly target college-level students with biology majors. We believe that studying and uncovering the role of graph knowledge in using graphs appropriately will help us have a more complete understanding of the reasons behind biology undergraduate students' graphing difficulties, thereby providing targets to help instructors to improve students' performance in graphing.

1.4 Research Objective

In our study, we have a main research question: What are the differences in graph knowledge among undergraduate biology students, graduate biology students, and biology professors?

This study will bring benefits to students, teachers, and other educators. Understanding the reasoning behind students' difficulties with graphs will not only help students learn better in their STEM classes, but will also help teachers and professors to improve their teaching methods to help students succeed. It will also provide valuable information for educators who develop and arrange curriculum as well as teachers and professors who use graphs in their teaching.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview

Graphs are important representations that are very commonly used in scientific communication, but creating effective graphs from data involves a wide range of skills and knowledge. In spite of years of instruction on visualizations and experiences with graphing, students in higher education (and sometimes even professors) still have difficulties in graph construction, interpretation, and evaluation. To improve instruction to increase students' competence with graphing, we first need to understand the difficulties they have. The literature review that follows includes theoretical perspectives, recommendations for undergraduate biology education, and data on graphing difficulties along the novice-to-expert continuum.

2.2 Meta-Representational Competence

In order to study the factors that are needed for a person to read and use graphs appropriately, we consult the components of Meta-Representational Competencies. The term “Meta-Representational Competencies”, or “MRC”, represents the full range of capabilities that a person has when constructing, understanding, and evaluating eternal representations. The term was first developed by diSessa and Sherin in their paper in 2000; instead of concentrating on uncovering students’ misconceptions on graphs, the researchers wanted to focus on what students “knew” about graphs. They claimed that MRC did exist in students, as students were shown to have a deep, rich and generative understanding of external representation (diSessa, Hammer, Sherin, & Kolpakowski, 1991). There are four main components in the MRC:

1) Invention: the ideas and skills that a person needs to have in order to construct new graphs;
2) Critique: the knowledge that a person needs to have in order to compare and judge the quality of graphs;
3) Functioning: the knowledge about the "why" and "how" of graphs, i.e. the function, purpose, advantages, and limits of graphs;
4) Learning: the knowledge that a person needs to have in order to foster their own learning of new graphs.

diSessa and Sherin pointed out in their paper that delimiting a "list" of knowledge that students have about representations was difficult, as the knowledge were more than simple facts to be memorized. The researchers decided to study the knowledge by understanding how it developed, and they believed that MRC developed from students' previous experiences with representations, i.e. the production and evaluation of graphs as well as the communication using graphs.

The learning and teaching of MRC was plausible and valuable for improving instruction of scientific representations. On one hand, students had a rich and deep basis of MRC for instructors to build on, and they found MRC-related experiences very engaging and sense-making. On the other hand, MRC was frequently used by scientists and mathematicians to design their representations, and the increasing use of technology had put an increasing premium on MRC. In addition, MRC tasks might help instructors to attract students who were less engaged in mathematics and science, due to its rich and often continuous nature which is different from current mathematics and science instruction.

A critical element within MRC is graph knowledge. We define *graph knowledge* as the knowledge a person has about a type of graph, including the graph's name, function, type of data that can be displayed with the graph, as well as the graph's affordances and limitations. Graph knowledge is deeply inter-connected with the four components of the

MRC: Invention, Critique, Function, and Learning. In addition to MRC, the spatio-visuo knowledge, mathematical and statistical skills, discipline-special knowledge and the meta-cognition of knowledge also interact with graph knowledge. We could say that graph knowledge includes parts of the MRC components, but we do not intend to use graph knowledge to represent all the aspects of MRC. In our study, we will only focus on studying the role of *graph knowledge* in constructing and understanding graphs, and we believe studying and uncovering the role of graph knowledge in using representations will help educators to understand students' difficulties with graphs and help teachers to improve their instructions.

2.3 Education Standards and Recommendations

So why are graphs so important? They are used in many places (lectures, textbooks, papers, reports) and they have a variety of functions, such as to present data, to show trends, to support claims, to communicate ideas, and so on (Treagust & Tsui, 2013; Weissgerber, Milic, Winham, & Garovic, 2015). Several calls for science education reform include an increased emphasis on students engaging in the practices of science as a means to increase engagement and learning of disciplinary content. These practices include working with data and applying mathematical and quantitative approaches to its analysis and interpretation, including graphing. In HHMI's report: Scientific Foundations for Future Physicians, one of the competencies emphasizes on students' abilities to integrate data, modeling, computation, and analysis. Specifically, the students need to be able to apply basic mathematical tools, including functions, graphs, measurement and scale, to reach a basic understanding of problems (AAMC-HHMI Committee, 2009). In "Vision and change in undergraduate biology education: a call to action", the report points out that developing and interpreting graphs is one of the core competencies for students to master in order to use quantitative reasoning. These contents indicated that researchers and educators are calling for emphasis on graph education and assessment. Therefore, it is increasingly important for college-level science students to be able to read graphs, understand data behind graphs, and interpret the messages that the graphs want to convey.

2.4 Students' Difficulties with Graphs

While calls to undergraduate biology education reform suggest that students work with, analyze and interpret data, undergraduate students have difficulties with several concepts and skills related to graphing. Although this study only targets higher-education students, we felt necessary to do a literature review on younger students' difficulties with graphs as one of our participant populations included lower division students who are recent high school graduates. In the U.S., students start to get in contact with graphs in K-12 (NGSS Lead States, 2013), and their graph knowledge and experiences began to build up since then (Novick, 2004). We also observed in our study that our participants recalled graph knowledge that they obtained from primary schools, middle schools and high schools. Knowing the students' difficulties with graph since they've started to learn graphs would give us a big picture on the types of graph knowledge that out participant might be lacking and aligns with the novice to expert continuum approach to our study.

Graphing in K-12 Education

Graphs should not be a new tool to students, since they start to learn about graphs at an early age (NGSS Lead States, 2013). Bryant & Somerville (1986) claimed that young students did not find the spatial demands of graphs difficult at all. The two researchers presented a study targeting 32 students from the same school, in which 16 are six-year-olds and 16 are nine-year-olds. Their goals were to determine whether children can find the y-axis value if given the x-axis value on a graph, and whether the fact that they have to extrapolate non-perpendicular lines in graphs causes them difficulty in reading graphs. In the first part of the study, a position was given on one axis, and the student had to find the corresponding position on the other axis by extrapolation. In the second part of the study, Children were shown two different graph-like displays in each of which a straight line was drawn through the origin, one at an angle of 56 degrees and the other 34 degrees. The results indicated that, although six-year-olds were significantly less accurate than nine-year-olds in their extrapolations of imaginary straight lines, the two groups of students can easily cope with the spatial aspects of graphic information.

However, there are several studies dealing with students' problem with graphing in their later stage of education and with graph construction in addition to graph interpretation. Shaw, Padilla, and Mckenzie (1983) claimed that students in Grades 7 through 12 demonstrated an inadequate ability to construct and interpret line graphs. They target 625 middle school and high school students, asking them to provide baseline data of line graphing skills to examine their graphing ability. Their results showed that seventh-grade and eighth-grade students were significantly less successful in graphing basic line graphs than high school students, which have more experiences with scientific graphs, demonstrating the suggestion that fundamental graphing skills are developing over this time frame, but could be introduced and emphasized in earlier grades.

In a similar and more recent study, Mevarech and Kramarsky (1997) also targeted middle school students on their graphing abilities. Their goal was to investigate student's conceptions and misconceptions relating to the construction of graphs. In their study, 92 grade 8 students were randomly selected from two different middle schools. The participants were asked to construct graphs representing each of four given situations representing increasing, constant, curvilinear, and decreasing functions. The students were given pencils and four sheets, each sheet with one problem printed on the top. On the pre-test, only 26 (27%) students constructed all four graphs correctly, which is a relatively low percentage. 40% of the students failed to construct even one graph correctly. Three major categories of problems were also identified, including constructing an entire graph as only one point (i.e., when some students constructed correctly the x and y axes, but they marked only one point, one bar, or one histogram), constructing a series of graphs with each representing only one factor from the given data, and conserving the form of an increasing relationship between variables under all four conditions.

Undergraduate Students Difficulties with Graphs

Students continue to experience difficulties with graphical representations in colleges. Meletiou-Mavrotheris and Lee (2010) conducted a study targeting only college-level introductory statistics students, and they found out that these students have difficulties in graph reading and interpretation, graph construction, and graph evaluation. Their study aimed to investigate students' ability to reason about variation in histograms. The site of the study was an introductory statistics course and there were 35 students in the class. The students were assessed by 10 tasks related to histograms, and the questions were related to the construction, interpretation, and application of histograms. Some of the students were also observed and videotaped while solving problems. Their results showed that, at the beginning of the course, students have very limited understanding of graphical representations. For example, a majority of students were not able to correctly distinguish between histograms and bar graphs. Some of the students' difficulties persisted despite the course's continuous efforts to help them improve their understanding of histograms.

A similar study was published in 2009 by Bruno and Espinel, where they targeted primary education major undergraduate students. In the study, 29 primary education majors were given a written test with two questions which was designed by the researchers. The test was developed to test students' ability to construct, read and interpreting a variety of quantitative representation of data, including frequency polygon and histogram. A descriptive analysis of the tests then was performed along with a study of students' answers of interest to the research objective. The results showed that, in the 29 students, only 1 student correctly drew graphs in both questions, and the rest of them made mistakes in either the histogram or the frequency polygon (a line graph showing frequencies of groups).

Two studies with undergraduate biology students highlight the fact that difficulties with knowledge and skills related to graphing still persist, even with some instruction on data and graphing. A study that was published by Bray-Speth et al. in 2010 studied the quantitative literacy skills of undergraduate students with life science majors or pre-health students. The researchers assessed the quantitative skills that students had and concluded that the students experienced difficulties in representing data graphically. In the study, a biology-related situation was given to the students, and the students were asked to perform a simple calculation and to create a graph from the resulting data. The answers were then evaluated by researchers based on choosing the right type of graph for the data, appropriately labeling the axes, and correctly graphing the calculated data. Analysis of students' responses revealed that undergraduate students with life science and pre-health majors experience difficulties with representing data on a graph, labeling the axes properly, and formulating complete arguments from data. McFarland discussed common graphing problem that college biology students made in her paper (2010). She stated that, from her own experiences with students, college-level biology students sometimes failed to present appropriate labels for graphs, chose wrong type of graph to represent data, had frustration with scaling, and had problems identifying the relationships between variables.

The results of these studies demonstrated the fact that college-level undergraduate students, both in and outside of the discipline of biology, were experiencing difficulties in representing graph construction and interpretation. These studies were all done in the United States.

Learning Graphs under the Current Education System

According to the constructivism theory, students' graph knowledge is actively constructed with their existing concepts and models, and it is modified with new learning experiences, in which science education plays a critical and essential role (Duffy, & Jonassen, 1992).

It is necessary for us to become aware of the effect that formal instruction on graphs can have on students. In Mevarech and Karamarsky's study (1997) in which they targeted middle school students on their graphing abilities, on the pre-test only 27% of students constructed all four graphs correctly. However, after being taught by an instructor four times a week for three weeks, the number of students who constructed all four graphs correctly increased to 45%. This significant increase in correctness indicated that, following instruction, students could overcome the difficulties and improve performance on graphing.

Other than the role of teachers and instructors, the design of the curriculum on graphing skills is also critical to students. Picone, Rhode, Hyatt, and Parshall (2007) targeted 240 college-level students in ecology and environmental science courses. In their study, they assessed graphing skills of the 240 students from four colleges and universities. Over the course of a semester, the researchers integrated graph education and scientific data analysis throughout the lab and lecture courses using an active-learning method that they developed. Students graphing skills were assessed before, during, and after the courses. Compared to pre-tests, a significant increase in bar graph and scatterplots interpreting skills were detected. There is also a considerable increase in their abilities on making graphs from raw data, i.e. graph construction. On the other hand, it is noticeable that very little improvement was detected in their ability to understand independent and dependent variables. More than half of the students still have difficulties in summarizing overall trends from data with variation. Students also failed to improve their abilities to interpret complex bar graphs with interactions between variables.

As we discussed earlier in Chapter 2.1, a crucial skill for appropriate graphing is discipline-related knowledge. Within the discipline of biological science, constructing and reading graphs using experimental data are critical skills for students to have (Dasgupta, Anderson, & Pelaez, 2014). In a recent study, Shi, Power, & Klymkowsky (2011) conducted a study targeting undergraduate students on their thinking of experimental design. The authors claimed that a well-designed control group was a key component of a scientific experiment, thus they tested college-level students on their understanding of the roles of control experiments. To their surprise, a high percentage of students still experienced difficulties identifying control conditions in experiments after completing three college-level laboratory courses.

2.5 Expert-Novice Studies in Science

A goal of undergraduate education is to move novice students forward along a continuum of knowledge and skill closer to the expert-like state. A number of previous studies were carried out to discover and analyze the differences between the experts and the novices on their ability to categorize information, understand situations, and solve problems in science. By doing so, the difficulties that novices face in science could be identified (Ericsson & Smith, 1991).

In Bransford, Brown, and Cocking's book "*How people learn: Brain, mind, experience, and school*", the authors researched a number of expert-novice studies in various fields, including studies in areas of mathematics, physics, history, computer science, chess, teaching, etc. They pointed out that the differences between experts and novices in processing knowledge and solving problems were complex and on different levels. For instance, the extensive knowledge that experts had differentiates them from novices in the ways they acquire, organize, and interpret information from outside environment (Bransford, Brown, & Cocking, 1999). Experts do better than novices in recognizing features and meaningful patterns of knowledge, and they were more aware that knowledge is conditionalized on various circumstances (Bransford, Brown, & Cocking, 1999). When solving scientific problems, experts tended to first acquire an overall understanding of the problems (i.e. thinking in terms of "big ideas") while novices were more likely to look at problems by fitting them into formulas and vying for answers that they experienced in their everyday lives (Bransford, Brown, & Cocking, 1999). In addition, the authors mentioned that experts had the abilities to retrieve knowledge that were relevant to problems with little attentional efforts; i.e. they could link new problems with their previous knowledge and experiences relatively effortless than novices (Bransford, Brown, & Cocking, 1999).

Other than Bransford, Brown, and Cocking's work, a number of other expert-novice studies also demonstrated critical characteristics of expertise. A study related to graphing was performed by Carter and his colleagues in 1988. They carried out an examination on the expert-novice differences in processing visual information in classrooms. The results suggested that experts appeared to be better at forming connections among various pieces of information and representing situations into meaningful units when compared to novices (Carter, Cushing, Sabers, Stein, & Berliner, 1988).

In the discipline of biology and biological science, Boshuizen and Schmidt (1992) examined and analyzed the role of biomedical knowledge in clinical reasoning by experts, novices, and participants at intermediate levels at expertise. Using a combined think-aloud interviewing and post-hoc explanation methodology, they showed that experts have more in-depth biomedical knowledge, which generally support a three-stage model of expertise development in medicine (acquisition, practical experiences, and integration).

The above studies are highly related to my research question in terms of expert-novice framework, context (visual representations) and discipline (biological science). Using expert-novice as the framework and connecting the ideas of MRC and the theory of constructivism, I hypothesize that experts and novices have different abilities and patterns in processing biological graphical information and solving related problems, specifically differences in their graph knowledge. My findings will provide science educators with additional targets for instruction to help students increase their competence with data and graphing.

CHAPTER 3: METHODS AND DESIGN

3.1 Study Design

This study is a sub-project of a larger project developed by Stephane M. Gardner and Aakanksha Angra. The procedures of the whole project are: modified Pre-interview PPI (Clase, Gundlach, & Pelaez, 2010), Graphing Protocol, Background Information, Stage 1 - Graph Construction, Stage 2 - Graph Evaluation, Stage 3 – Graph Knowledge, and Post PPI. My study only focuses on Stage 3, which is a sub-project that focuses only on studying subjects' graph knowledge.

Participants

There are 58 participants in this study. Eight of them are faculty members of Purdue University's Biological Science Department; 13 of them are graduate students of Purdue University's Biological Science Department; 13 of them are upper-level undergraduate students (juniors and seniors) with a Biology major at Purdue University, 24 of them are lower-level undergraduate students (freshmen and sophomores). All participants in this study were recruited via email invitation from the researchers directly (professor and graduate student pools) or indirectly through course mailing lists. Participants were given a $20 Gift card at the end of the study to compensate them for their time. The recruitment and procedures were done in accordance with IRB protocol No. 1210012775, "Investigating the Reasoning Involved in Creating Graphical Representations of Data in Biology".

Interview Procedure

In Spring 2012, we sent 15 graphs to 12 professors at Purdue University and asked them to rank them in terms of graphs that they felt all undergraduate students should know and graphs that they covered the most frequently in class. The 12 professors resembled a variety of sub-discipline in biology, including Environment and Ecology, Development and Disease, and Molecular Biology. They also have a mixture of teaching experiments, including a variety of teaching targets (i.e. biology undergraduate students, biology graduate students, and non-biology major undergraduate students) and teaching format (normal lectures and laboratory). The 15 graphs we sent to them, and the graphs consisted of only axes and data, i.e. without labels, scales, or titles. The information of professors and the results were shown in *Table 1, 2 and 3.*

The five graphs (a bar graph, a line graph [growth curve], a scatterplot, another line graph [variation], and a histogram) chosen by the 12 professors were used in the interview. Again, these graphs have no scale or text on them. During the interview, the five graphs were given to the subject in the order of: the bar graph, the line graph (growth curve), the scatterplot, the second line graph (variation), and the histogram. After each graph was given, the interviewer asked five prompts/questions in order:

Q1: Please examine the graph and tell me your first impression.

Q2: Now from your past knowledge and experiences, please describe to me what type of data can fit this graph.

Q3: Can you think of a specific scenario to fit the graph?

Q4: Why did you choose that specific scenario to fit this graph?

Q5: What type of trends does this type of graph helps to convey?

When answering those five questions, the participants were asked to “speak aloud their thoughts”. The interviews were recorded using Smartpen, and then transcribed verbatim into text document by playing back the audiofile using Echo Desktop. We then used the coding scheme that we developed to code the data.

Table 1 Professors who voted on the graphs to be used in the interview were identified with a number. UUG: Upper-level undergraduate students. LUG: Lower-level undergraduate students. EE: Ecological and Evolutionary Biology. DD: Development and Disease. MB: Molecular Biology.

Number	**Cluster**	**Targeting Students**	**Teaching Form**
1	EE	Biology Freshman	Lecture
2	EE	Biology Freshman, Senior	Lecture
3	DD	Biology Sophomore	Lecture
4	DD	Biology Sophomore	Lecture
5	DD, MB	Biology UUG, Graduate	Lecture
6	DD	Biology UUG	Lecture
7	EE	Biology UUG, Graduate	Lecture and Lab
8	EE	Biology Sophomore	Lecture
9	DD	Biology UUG	Lecture and Lab
10	EE	Biology Sophomore, Senior	Lab
11	EE	Non-Biology LUG	Lecture
12	MB	Biology Senior	Lab

Table 2 The types of graphs and the results.

15 Original Graphs	Type	Chosen?	Graph
1	Line graph (curved, with dots)	No	
2	Bar graph (adjusted to 100%)	No	
3	Scatterplot	**Yes**	See *Table 3*
4	Radar Graph	No	
5	Log-log Plot	No	
6	Line graph (multiple, w/SD)	No	
7	Line graph (variation)	**Yes**	See *Table 3*
8	Histogram (w/bars)	**Yes**	See *Table 3*
9	Line graph (growth curve)	**Yes**	See *Table 3*
10	Histogram (w/lines)	No	
11	Dot plot (categorical)	No	
12	Dot plot	No	
13	Box Graph	No	
14	Bar graph (w/ error bar)	**Yes**	See *Table 3*
15	Line graph (curved, w/o dots)	No	

Table 3 The five graphs that were chosen and used in the interview. This form was modified from Angra, A., & Gardner, S. M., 2016.

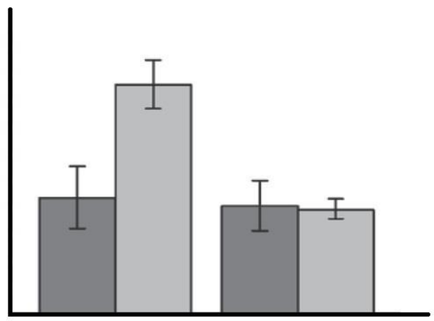
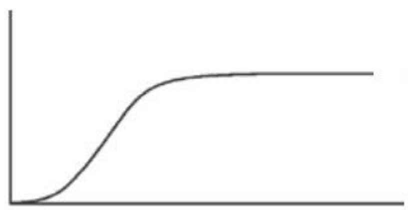
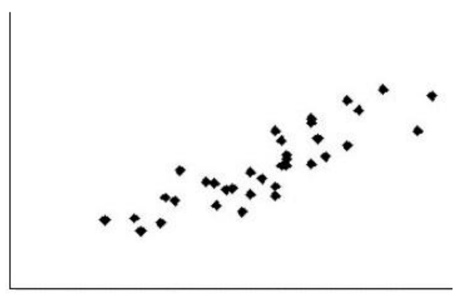
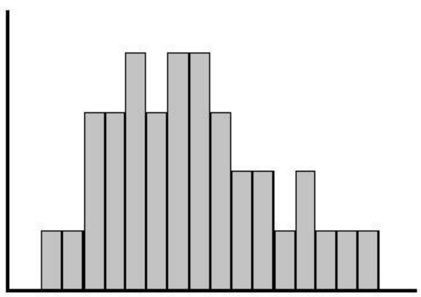
Graph Number and Graph Type	Graph Usage
1. Bar Graph w/ error bar 	• To compare categorical data, percentages, or summary statistics from multiple groups. (Schriger D.L., & Cooper R.J., 2001) • Each bar represents a category; shape canbe changed by moving the categories around. (Humphrey P.B., Taylor S., & Mittag K.C., 2014)
2. Smooth Line Graph 	• To show how a single variable or multiple variables changes over time or to show how a variable deviate from a set baseline. (Few S. 2004) • X axis portrays categories while the Y axis portrays quantitative values. (Few S. 2004)
3. Scatterplot 	• To show individual data points from bivariate data. (Schriger D.L., & Cooper R.J., 2001)

Table 3 Continued

Graph Number and Graph Type	Graph Usage
4. Varied Line Graph 	• To show how a single variable or multiple variables changes over time or to show how a variable deviate from a set baseline. (Few S. 2004) • X axis portrays categories while the Y axis portrays quantitative values. (Few S. 2004)
5. Histogram 	• To show a distribution of data with the independent variable as continuous. (Humphrey P.B., Taylor S., & Mittag K.C., 2014) • Uses numerical data instead of categorical data. (Humphrey P.B., Taylor S., & Mittag K.C., 2014)

3.2 Methods

Inductive and Deductive Coding

In order to summarize our qualitative data into a brief summary format, we used a hybrid process of inductive and deductive coding to analyze our data. “Coding” is a process of encoding qualitative information, usually the text, into explicit “codes” which can be organized into themes and categories to help reveal trends and patterns in the data (Boyatzis, 1998). While a deductive analysis involves using previously-outlined patterns or “coding schemes” to help organize data (Crabtree and Miller, 1992), an inductive analysis refers to the approaches reading raw data, i.e. the transcripts, to derive patterns, themes, or concepts (Thomas, 2006).

Table 4 The theoretical constructs that were consulted in the process of deductive coding

Papers	Theoretical Constructs
Postigo & Pozo, 2004	Three levels of depth in which graphic information are processed: 1) Explicit Information: a superficial reading of basic graph elements 2) Implicit Information: reading of graphics materials beyond isolated elements; "decoding" information 3) Conceptual Information: establishing relationships between different graph elements; presenting overall analysis of information
Novick, 2004	Six types of knowledge required for diagrammatic competence: Implicit Knowledge (non-verbal performance assessment), Construction Knowledge (rules for graph constructing), Similarity Knowledge (similarity of a situation to other situations), Structural Knowledge (structure of a particular type of graph), Metacognitive Knowledge (monitor the comprehension), and Translational Knowledge (transfer information from one representation to another).
Friel, Curcio, & Bright, 2001	Three levels of graph comprehension: 1) Elementary: extracting data and information from graph, such as locating 2) Intermediate: interpolating and finding relationships in the data in a graph, such as integrating and interpreting 3) Overall: moving beyond the data and analyzing the relationships implicit in a graph, such as generating and predicting
Carswell, 1992	Evaluation levels of graph comprehension: 1) Point reading or attention to a single specifier 2) Local and global visual comparison of data and feature in graphs 3) Synthesis and integration of most of or all the graphic features

Each of the two approaches has its unique benefits, and we decided to use a mixture of the two approaches: we first started with inductive coding, reading through raw transcripts and trying to identify patterns in the data, by which we developed and established categories. In further steps, some theoretical constructs (*Table 4*) were consulted to explain and evaluate the categories. After a coding scheme (*Appendix A*) was constructed, we used the coding scheme to code the remaining data.

Developing the Coding Scheme

I started analyzing data by reading through the transcripts to get a general idea of what our participants were talking about. While reading, I also used a pen to memo on the transcripts. I noticed some similar patterns (the same words, phrases or similar short sentences) among multiple transcripts.

> P8124: "I mean it's close to a normal distribution but it's skewed a little on the high end."
>
> G4235: This seems to be like one data set in which there is like a normal distribution of the data.

In the above example, the two participants both said the phrase "normal distribution". I highlighted this phrase in these two transcripts, and when later I noticed a third person saying "normal distribution" I also highlighted the phrase in that transcript.

> G1906: "The x axis and the y axis need to be labeled with units."
>
> G6092: "My first impression is that the x axis and the y axis is not labeled."
>
> P8124: "There is nothing on the axis. There's no units, there's no key."

In the above example, although there is not a specific word or phrase mentioned, all three participants talked about a graph that lacked labels on the axes, which I considered as similar patterns. I highlighted these sentences in the transcripts, and I wrote down "lacking labels" next to them. Later when I noticed other participants talking about graphs lacking labels, I also wrote down the same thing next to the sentences.

If a same word or phrase or a similar sentence was mentioned in more than one transcripts, I took notes down saying that this is a "common pattern". In this way, I read all transcripts and highlighted all similar words, phrases, and sentences. I called them "codes". The number of different codes were growing as I read more and more transcripts, and I kept modifying and refining the name and meaning the codes.

After I read all the transcripts, I had a list of codes. A code may be a word, such as "average", "median", "conditions", "comparison"; a phrase, such as "independent variable", "control group", "standard deviation", "normal distribution"; and my note for short sentences, such as "lacking labels", "take home message", "learned from class". My next step was to put them into different categories according to their identity or meaning. For example, "average" and "standard deviation" both belong to statistical terms; the participants should be thinking about statistical analysis when they talked about "average" or "standard deviation". So, I put "average" and "standard deviation" under the category "statistical terms and analysis", which later was joined by "standard error", "median", "R^2", "degree of correlation", and other terms that I thought should be included in the same category.

1. Statistical Terms and Analysis
   - 1.a Average
   - 1.b Standard Deviation
   - 1.c Median
   - 1.d Degree of correlation (i.e. tight correlation/strong correlation)
   - 1.e Significant Difference

The "Statistical terms and analysis" category now looked like this. I found out that in these five codes, only the first three codes "average", "standard deviation" and "median" belonged to defined statistical terms, whereas the last two codes "degree of correlation" and "significant difference" were more likely referring to statistical analysis: they were more than a term. So I added a layer to this category, introducing two sub-codes: "Statistical Terms" and "Statistical Analysis", and this category now looked like this:

1. Statistical Terms and Analysis
    1.a Statistical Terms
        Average
        Standard Deviation
        Median
    1.b Statistical Analysis
        Degree of correlation (i.e. tight correlation/strong correlation)
        Significant Difference

The next step was to add definitions to the categories and the sub-categories, not only for me to define and understand the terms better, but also for other researchers to learn the meaning of the terms. After adding definitions to the terms and refining the details, the "Statistical Terms and Data Analysis" category now looked like this:

1. Statistical Terms and Data Analysis: The mention of specific statistical terms or functions

    1.a Statistical Terms: The mention of statistical terms in the scenario/example, such as: average, standard deviation, median, standard errors, mode, range, variance, etc.

    1.b Statistical Analysis: The mention of types of statistical analysis in the scenario/example, such as: degree of correlation, statistical significance, trendline or best-fit-line, R^2, etc.

Most of the time, I categorized a code based on its own meaning. For instance, "experimental group" and "control group" were categorized under "experimental design", and "naming x axis with a variable" and "adding a title" were put under "graph construction". However, under some special circumstances, a code might be categorized or defined based on the question the participant was asked when answering the question. For instance, we asked five questions during the interview, one of which is "Can you think about a specific scenario to fit this graph?" This question required the participant to think through their learning experiences to find out an example that could fit in the graph, or “populate the graph with data.” In this case, i.e. when the participant was answering this particular question, I put their answers under the category “Example Type”.

2. Example Type
    - 2.a Academic/Research Related
    - 2.b Personally Experienced in Life (e.g. weather reports, phone apps, news)
    - 2.c Moth Comparison
    - 2.d Couldn’t think about an example
    - 2.e Bacteria Growth
    - 2.f Leave Example
    - 2.g Vague, Short Example

Then, I added sub-categories under this category. “Personal Example” referred to examples that the participant personally experienced before, such as lab-related examples, real-life examples, and textbook examples. The example below is categorized as a “personal example” for the participant, a professor, learned this scenario from his or her research project.

> P1562: “Talking about actin… so spindle of actin, the actual length. And sometimes also shown as a percentage of total… of actin… It reflects the change in production of the F-actin and it reaches… which must be below 100% because there must be some monomeric actin to keep the process going. So basically it means that you have 80%, 90% F-actin and 10% or 20% G-actin that still at single and is being removed.”

Earlier in Chapter 3, I mentioned that my thesis project is a part of a larger research project. My thesis project is the fourth task of the whole project, and in the first three tasks, participants were given a couple of pre-designed, detailed scenarios. These scenarios, including the bacteria growth example, moth comparison example, and leaf growth example, were classified under the sub-category “previous example”. I wanted to distinguish this sub-category from “personal example”, because using scenarios that were given in the previous interview questions might indicate a lack of personal graph knowledge and experiences. An example of “previous example” is presented below.

> LUG6788: "You could use the moth example that we did in the previous example… the darker colored bars could represent the dark moths, and then the lighter colored are the white ones, and then each of the groupings represent the different time periods."

There are also cases that participants failed to think of a specific example. When a participant gave us a very vague example or could not think of an example at all, these scenarios were classified under the sub-category "vague example".

> LUG7358, Graph3: "… In high school, plotting or something, but really… I am not sure."

After adding sub-categories and description, the category of "Example Type" looked like this:

2. Example Type: The type of the specific example/scenario the participant generate; this is not to be confused with the type of graph source
   - 2.1 Personal Example: The type of example/scenario is based on participant's personal knowledge, such as: personally experienced in life, research-related, teaching, etc.
   - 2.2 Previous Example: The type of example/scenario is based on examples that were given in previous tasks in the interview, including: bacteria growth, moth comparison, and leaves growth examples.
   - 2.3 Vague Example: The given example/scenario is very vague, lack of details, or the participant failed to give an example/scenario

I modified and refined all the categories using the above processes. Then, three "main categories" in our coding scheme were developed and all categories were put under the three main categories: 1) Graph Description (basic description of the graph without further interpretation), 2) Graph and Data Analysis (local or global interpretation of the graph), and 3) Instantiation (concepts and reasoning that subjects engage in while populating the graph with data). For example, "graph type" belonged to Graph Description, because the subject did not need to do any interpretation to recognize the type of a graph. A category could sometimes be included in more than one of the main categories; for instance, "experimental design" was included in both "Graph and Data Analysis" and "Instantiation". In this case, a code would be put in "Graph and Data

Analysis" if the subject was interpreting the given graph, or it would be put in "Instantiation" if the subject was talking about the specific scenario with which they came up.

The idea and the development of these three main categories came from both our observation and the patterns summarized by other studies (see *Table 4*). This is a combination of inductive coding (our observation from the transcripts) and deductive coding (structured and formatted patterns by other researchers). While each coding methods have their advantages and disadvantages, they often come together in qualitative research (Schadewitz and Jachna, 2007).

Under these three main categories, there are sub-categories that state precisely the definitions and the criteria that we were following. In further steps, some theoretical constructs (*Form 1*) was consulted to explain and evaluate the categories. After a coding scheme was constructed, we used the coding scheme to code the remaining data. The final coding scheme was included in *Appendix A*, with all details attached to the categories.

Using the Coding Scheme to Code

To better demonstrate the using of the coding scheme, an example is given below. In the example, a participant was given a histogram by the interviewer and was asked to answer the five questions. Using the coding scheme, we read through the transcript and found codes in it. In the example, codes in "Graph Description" category were highlighted in yellow; codes in "Graph and Data Interpretation" category were highlighted in green; and codes in "Instantiation" category were highlighted in blue.

> I: So here is your last graph. Please examine the graph and tell me your first impression.
>
> 2212: This is a basically a bar graph. And because all the bars look at the same, this probably just represent just a single plot. You are just comparing it as different sets of your independent variables. So should I come up with a scenario?

I: Yes.

2212: Okay. So I just chose a similar scenario with time and money, but this time I chose a bill statement. So this could be… let's say at the first month, or maybe first two months, you don't own anything. Let's say January and February, and then March… Let's say 1000, 2000, 3000, 4000, 5000. So let's say for the first two months, you don't use that credit card at all. And then you began using it. So let's say for the first two months you spend 1000 dollars, and then you spend around 3500 dollars, and then up to 4500 dollars, and then you began to decrease the money you are spending on that card. So that would be an example of the scenario.

I: And why did you choose that specific scenario to fit this graph?

2212: I feel like it would be a practical way representing this type of data. Because when people begin spending a certain amount of money they begin to decrease using it. So like save money and keep a steady budget. So maybe like kind of levels off there, like a 1000 dollars.

I: So have you encountered this scenario or this graph before?

2212: Maybe not the exact same graph before, but the scenario is practical.

I: So you've seen this graph before? Where?

2212: I have seen this type of graph. Mainly textbooks. You don't mainly deal with bar graph much so just general science graphs and textbooks you see bar graph a lot.

I: And can you describe to me why you label your axes that way?

2212: Sure. Because I am doing a credit card statement, you only see the amounts of each months. And like the time is the independent variable. And um, the amount of money on the bill is the dependent variable. So that depends on how much money you spend and you see at the end of each month.

I: What type of trends does this type of graph helps to convey?

2212: It generally convey the trend of a single plot sort of increases and decreases again.

I: What is the take home message of your scenario?

2212: So when you begin to use your card you are increase your spending rapidly, and then you reach the point past the mid-year and then you begin to decrease your spending.

I: Anything else you want to tell me about this graph?

2212: No.

After reading the transcript and finishing the coding process, we counted the number of codes in each main category, which gave us:

LUG2212 Graph: Histogram	**Graph Description**	**Graph and Data Interpretation**	**Instantiation**
Number of Codes	2	6	6

After getting the number of codes, the proportion of codes in each category could also be calculated:

LUG2212 Graph: Histogram	**Graph Description**	**Graph and Data Interpretation**	**Instantiation**
Proportion of Codes	0.14	0.43	0.43

This was an example of only one graph from one participant. Overall, we had 58 participants and each of them had five graphs. All graphs from all participants were coded and calculated in the methods described above.

The Rules of Counting the Same Code

When we used the coding scheme to code the transcript, it was fairly common that the same code appeared more than one time. For instance, this participant (LUG8308) was talking about error bars and uncertainty:

> Interviewer: What do you mean by uncertain?
>
> LUG8308: It means… like I mentioned earlier, when we make measurement, say if you are measuring the mass of the leaves instead, so I am changing this… if you are measuring the mass of the leaves, you are using a measurement to measure the mass, you have to include the error bars to say that the mass is approximately in this range. I am sorry can you repeat the question? (Interviewer: So, what do you mean by uncertain?) So the mass is not known to be this value. It can be in this range. The uncertainty is the measure of a range that it could be truly located in. It's larger here than it is over here. So in this case the uncertainty varies nearly the entire number of leaves. However, that is a bad application. If they are measuring the mass of the leaf, I am much better at it. So if you are measuring the mass of the leaf, this error bar says that there's much more uncertainty in it. There is no method knowing the exact measure, however, it says that the general mass could line within this range. However, here it is smaller, it is known that the actual mass lies in a much more precise range.

Since the participant talked about collecting data by taking measurements, I highlighted the words "measurement" and "measure" in this part of the transcript and put this code into the "Instantiation > Graph Construction > Type of Data" category. However, in the above transcript, we observed that the participant mentioned the same code nine times. Does talking about the word "measurement" and "measure" repeatedly indicate that the participant knew more about data collecting methods? In the above example, the answer was most likely no, because the participant was only repeating the words without developing the idea into a deeper stage.

The basic role we used for counting the same code was simple: If the same code appeared more than once, count the code only one time, *unless* the code was in a different context. Like in the above example, instead of counting the code nine times, we counted the code only one time.

There were occasions that we needed to count a same code more than once: when the repeated code was in a different context. A good example is when a participant mentioned a same code in another graph type. For instance, if a participant talked about "measurements" when he or she saw the bar graph, and talked about "measurements" again when he or she looked at the histogram, then we counted the code twice: once in bar graph, once in histogram.

In the same type of graph, a participant might talk about a same thing but within different contexts. For example, a participant (P8124) talked about treatment groups and conditions when she looked at the bar graph, telling us her first impression of the graph.

> (Interviewer: What is your first impression of this graph?)
>
> P8124: "…It looks like you've got, um, 2 measures for 2 different treatment groups."

Later, after answering several other questions, in a different context she talked about conditions and treatment groups again:

> (Interviewer: so why did you use the scenario?)

> P8124: "…I don't know, you had different colors, you had them grouped in two clearly separated… so that I could tell that these were two paired, two different pairs rather than just 4 bars of...equally....4 different conditions."

In this case, although the participant talked about the same code, she talked about it under a different context (answering another question). In addition, the participant did not simply repeat what she said when answering question 1 but developed the idea more deeply and provided the reasoning. We counted these two codes twice.

Inter-Rater Reliability

The inter-rater reliability was carried out by two other researchers. Due to the huge amount of data, the two raters did not go through all the transcripts. Instead, one rater coded two transcripts of each of the five graphs, and the other rater coded two transcripts of the bar graph and the growth curve. After comparison and discussion, the degree of agreement reached 80%. Thus, we concluded that the coding scheme was reliable.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Qualitative Analysis of Transcripts

The purpose of this qualitative study was to examine the differences in graph knowledge of experts and novices. Our analysis focused on the number and the identity of codes falling in different categories, and we identified patterns and trends which will be explained below.

All of the 58 participants of our study were recruited from a large Midwestern university on a voluntary basis. Within the 58 participants, 8 were biology professors (P), 13 were biology graduate students (G), 13 were upper-level undergraduate students (UUG), and 24 were lower-level undergraduate students (LUG). The professors were a group of research-active scientists who had acquired extensive knowledge on representations, and thus they were considered as “experts” due to their expertise of the field. The lower-level undergraduate students, i.e. freshmen and sophomores, were assumed to have had the least experiences on scientific representations, thus were considered as the “novices”. The upper-level undergraduate students, i.e. juniors and seniors, were assumed to have had more experiences with biology graphs then the UUGs, since they took more biology classes in college and had more opportunities to learn biology graphs in lecture or lab. The graduate students had finished all the undergraduate courses and were conducting their research projects, but they still had less experience with scientific graphs than the professors. Therefore, they were considered as between the professors and the undergraduate students, or the “intermediates”. As such our population of participants provides us with an expert-novice continuum; The four groups, ordered from the most expert to the least, are P, G, UUG, and LUG. We did not include other variables, such as age, gender, race and ethnicity, etc. in our analyses.

In the previous chapters, we already described how the codes were generated and classified into different categories. All codes emerged from participant responses to different prompts in the interview that were designed to reveal their graph knowledge. *Table 5* is a summary of the three main categories and sub-categories that are under them. A detailed summary of the codes with definitions and examples could be found in *Appendix A.*

Table 5 The 3 main categories of the coding scheme with their definitions and the sub-categories. The bolded sub-categories (i.e. Experimental Design and Graph Construction) were the two sub-categories that we did extra analysis on.

	Definitions	**Sub-Categories**
1. Graph Description	Explicit knowledge, or what people can get directly from the graph without further interpretation	1.a Description of Graph 1.b Type of Graph
2. Graph and Data Interpretation	Implicit knowledge, interpreting parts of a graph or see overall trend/function of the whole graph without a specific scenario	2.a Experimental Design 2.b General Conclusion 2.c Statistical Terms and Data Analysis 2.d Trends 2.e Variables
3. Instantiation	Populating graphs with data and conceptual understanding of the graph with linking the reasoning with previous personal knowledge or experiences.	3.a Example Type **3.b Experimental Design** 3.c Statistical Terms and Data Analysis 3.d Metacognition or Metacognitive Monitoring **3.e Graph Construction** 3.f Mention of Other Graph 3.g Source of Graphs 3.h Trends

To give the readers a better idea of what the answers from the participants of different education levels were like, an example of a professor's answer for interview questions No.1 is presented below.

> Interviewer: [Showed the bar graph to P1562] "What is your first impression of this graph?"
>
> P1562: "Well I see a bar graph with two axes. And I see no indication what the axes are and also no units. The bar graph representing what values… error bars that are most likely standard deviations or standard errors of the mean. And I see two groups, darker gray and lighter gray, probably comparing with… and whether there's difference between the two pairs, on whether what these bars or axes belong. Looks like the first pair, there's a difference between the dark bars and the lighter bars. It's about two-fold increase and significant. For the second pair, the main values are similar and looks like these two groups are not significantly different. If I would compare, two darker grey bars are similar, while two lighter grey bars seem to be different."

The answer provided by Participant P1562 represented what a typical answer from a professor was like. In the answer, we noticed that after seeing "a bar graph with two axes", the professor soon realized the lacking of important components of a bar graph: "I see no indication what the axes are" and "also no units". Then, the professor not only saw "error bars", but they also pointed out the meaning of those error bars: "that are most likely standard deviations or standard errors of the mean". The professor started talking about the data in the graph ("I see two groups, darker gray and lighter gray"), followed by pointing out the trend or the function of the graph ("probably comparing with… and whether there's difference between the two pairs"). Then, the professors noticed that the difference in the first pair is "about two-fold increase and significant", and the second pair "are not significantly different." After comparing within pairs, the professor then did a comparison between the two pairs, saying that the "two darker grey bars are similar" while "two lighter grey bars seem to be different". In this detailed answer to the first question, we could see that the professor talked about a lot of things: graph type, graph shape, graph components that are missing, statistical terms, graph trend, graph function, and statistical analysis.

In contrast with the answer from the professor, we present the below answers to the same question of the same graph, from a graduate student, an upper-level undergraduate student, and a lower-level undergraduate student. These answers represented what typical answers from their education level were like.

> Interviewer: [Showed the bar graph to the participant] “What is your first impression of this graph?”
>
> G7476: So it’s comparing two different things shown by different color of the bars. And both of them depend on what’s on the x or on the y, it looks like whatever this lighter color bar is, it’s decreasing with respect to the x axis. Whereas the darker one is staying the same. Possibly a control experiment. It’s not changing, versus the lighter one is changing with respect to the x axis.
>
> UUG7222: Okay, so you are comparing two different items. And there are two different times or concentrations that you are comparing them in. For example, in physics they would call it an error bar but I am not sure.
>
> LUG8535: It doesn’t really tell me anything because it doesn’t have label on either axis. It’s probably some kind of statistical analysis with outliers. I’ve seen this type of graph in AP stats.

In the answer, the graduate student stated that the graph was “comparing two different things” which “depend on what’s on the x or the y (axis)”. He also stated that “the darker one is staying the same… possibly a control experiment”. The upper-level undergraduate student also talked about “comparing two different items” and “there are two different times or concentrations that you are comparing them in”. He also mentioned that he had seen “an error bar… in physics” but he was “not sure”. The lower-level undergraduate student said the graph “doesn’t really tell anything” because “it doesn’t have label on either axis”. He said the graph was “probably some kind of statistical analysis with outliers”, but instead of pointing out what type of statistical analysis, he only said that he had “seen this type of graph in AP stats”.

4.2 Quantitative Analysis of Qualitative Patterns

Coding to reveal data patterns: The Number of Total Codes

We coded all 58 transcripts using the coding methods described in Chapter 3. The codes for each participant were extracted from the verbatim transcripts and the number of codes were counted under each main category: 1) Graph Description, 2) Graph and Data Interpretation, and 3) Instantiation. We first added the number of codes in each main category together to get a number of total codes from each participant. There are four groups: Professors, or P (n=8); Graduate students, or G (n=13); Upper-level undergraduate students, or UUG (n=13); and Lower-level undergraduate students, or LUG (n=24). *Figure 1* presents the number of total codes (i.e. all codes from a participant, including codes from all 5 graphs and all 3 main categories) of the four education levels.

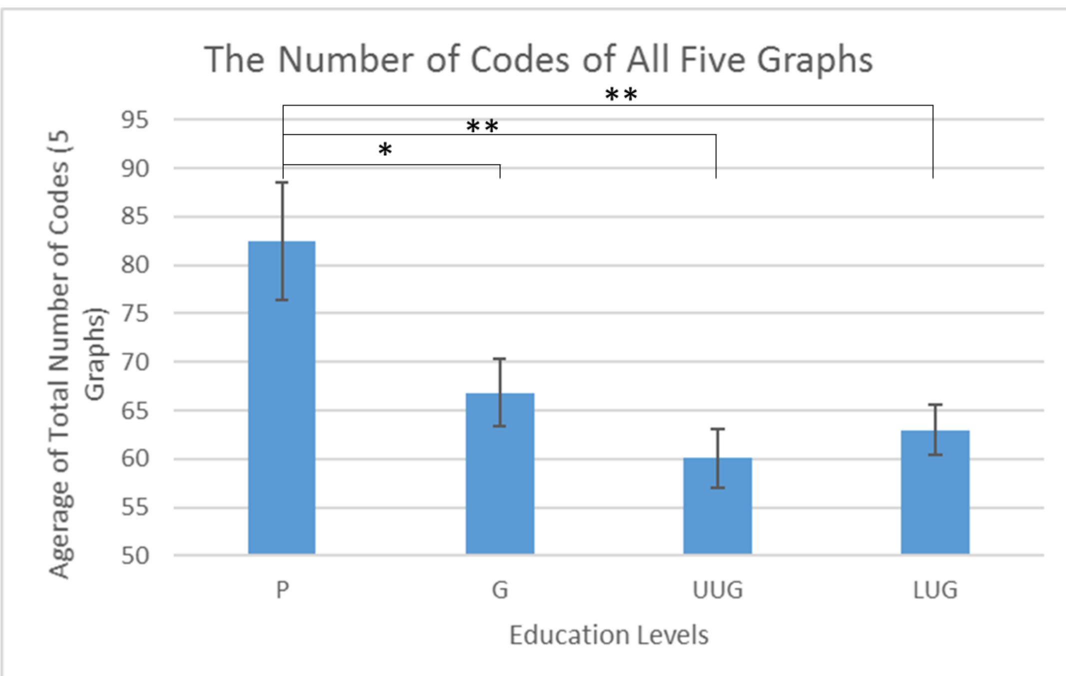


Figure 1 The raw number of codes of all five graphs of professors (n=8), graduate students (n=13), upper-level undergraduate students (n=13), and lower-level undergraduate students (n=24). The error bars represent one standard error. *p<0.05; **p<0.01, unpaired t test.

There was a significant difference in the number of total codes for professor (M=82.5, SD=17.12) and graduate students (M=66.85, SD=12.33), $t(19)=2.44$, $p=0.025$. Significant differences were also found between professors and upper-level undergraduate students (M=60.08, SD=10.99), $t(19)=3.68$, $p=0.002$; as well as between professors and lower-level undergraduate students (M=63.00, SD=12.45), $t(30)=3.49$, $p=0.002$. The code number of the three student groups do not have significant difference (for G and UUG, $t(24)=1.48$, $p=0.15$; for G and LUG, $t(35)=0.900$, $p=0.37$; for UUG and LUG, $t(35)=0.709$, $p=0.48$).

The 3 Main Categories: Patterns graph knowledge codes across graph types.

The three “main categories” that the codes were organized under were: 1) Graph Description (basic description of the graph without further interpretation), 2) Graph and Data Analysis (local or global interpretation of the graph), and 3) Instantiation (concepts and reasoning that subjects engage in while populating the graph with data) (Refer to *Table 5* and *Appendix A* for detailed definitions and sub-categories). To look at the graph knowledge from the expert-novice perspective, we wanted to look at the patterns of codes for across the participant groups for the five graphs: the bar graph, the smooth line graph, the scatterplot, the varied line graph, and the histogram. The data of the five graphs (bar graph, smooth line graph, scatterplot, varied line graph, histogram) were collected and analyzed separately, and the results were presented in separate graphs. The independent variable was the education level of the participants. *Figure 2* presents the average number of codes under each of the three main categories, Graph Description, Graph and Data Interpretation, and Instantiation (*Table 5*).

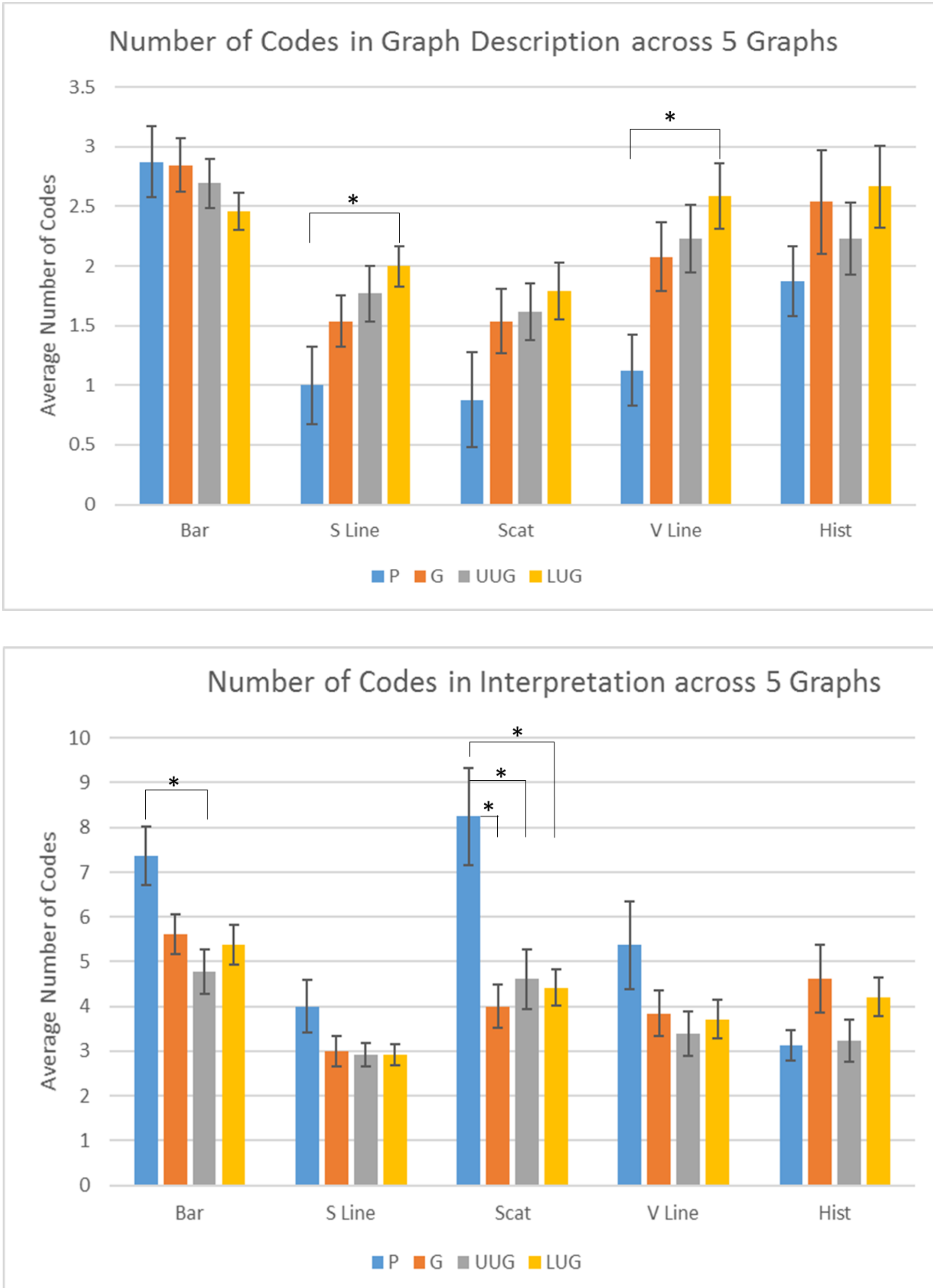
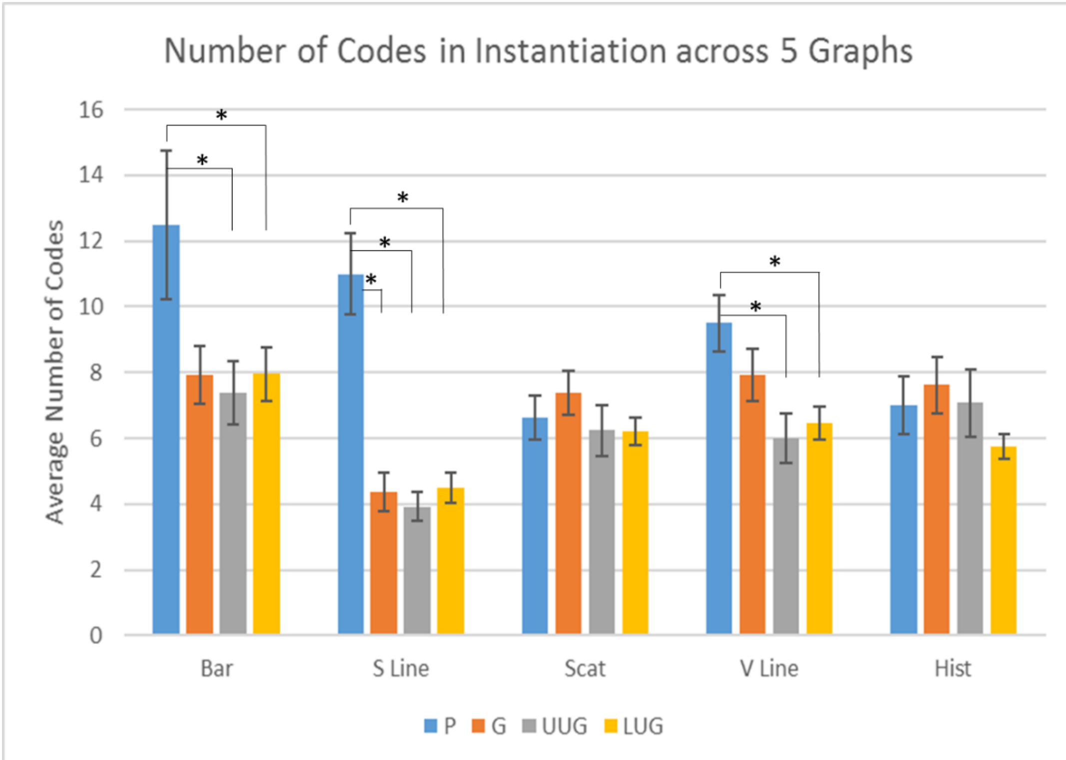


Figure 2 The average number of codes in the three main categories of the four education-level groups of the five graphs. The error bars represent the standard error. *p<0.05, calculated using 1-way ANOVA with a Tukey's post-hoc test.

Figure 2 Continued



Patterns of code numbers across the different graphs and different education groups were revealed in *Figure 2*. Both 1-way ANOVA with a post-hoc Tukey's test (*Table 6*; also indicated by asterisks in *Figure 2*) and 2-way ANOVA were carried out to test the statistical significance of differences in the results.

From the results of the 1-way ANOVA (see *Table 6*), we could see that professors are significantly different from the three student groups in many areas. In contrast, there is no significant difference among the student groups. The patterns of differences are unique for each graph. For the bar graph and the two line graphs, most of the significant differences are found in Instantiation category. For the scatterplot, it is the Graph and Data Interpretation that contains significant differences among different education levels. For the histogram, there is no significant differences in number of codes among education levels at all.

Table 6 Differences in number of codes in the three categories among different education levels. Calculated using 1-way ANOVA with a Tukey's post-hoc test. Yellow background indicates significant difference. P=Professors, G=Graduate students; UUG=Upper-level undergraduate students; LUG=Lower-level undergraduate students.

		P vs G	P vs UUG	P vs LUG	G vs UUG	G vs LUG	UUG vs LUG
Bar Graph	Description						
	Interpretation						
	Instantiation						
Smooth Line Graph	Description						
	Interpretation						
	Instantiation						
Scatterplot	Description						
	Interpretation						
	Instantiation						
Varied Line Graph	Description						
	Interpretation						
	Instantiation						
Histogram	Description						
	Interpretation						
	Instantiation						

Using the 2-way ANOVA, we first looked at the influence of participants' education level on their code numbers in the three main categories (regardless of graph types). The results from a post-hoc Tukey's test indicated that professors have significantly fewer codes than lower-level undergraduate students in Graph Description category, but that they have significantly more codes in the Graph and Data Interpretation and Instantiation categories. Professors also have significantly more codes in Graph and Data Interpretation and Instantiation categories when comparing to graduate students and upper-level undergraduate students. There is no significant difference among the three student groups in any of the categories.

When looking at the influence of graph type on the participants' code numbers in the three main categories, the most significant differences are between bar graphs and the other four graphs. Participants talked about more things in bar graphs comparing to the other graphs, especially to the smooth line graph. In all three main categories, participants talked more about bar graphs then the smooth line graph.

Comparing across the 4 education-level groups in each graph, we could see that the trends were different according to the different categories. In the "Graph Description" category, the overall trend was that the number of codes increased from the most expertise group (professors) to the least expertise group (lower-level undergraduate students). On the other hand, in the "Graph and Data Interpretation" and the "Instantiation" categories, the overall trend was that the number of codes decreased from the most expertise group to the least expertise group.

The Code Proportion

During the interviewing and transcribing process, we noticed that the length of the talking and the number of codes might not represent the amount, variety, and distribution of the graph knowledge within the three categories appropriately. Some participants tended to "talk more" than the others, and did not necessarily have more graph knowledge. For example, a participant might talk a lot when he/she did not know what the graph meant but simply were finding all the different terms they could think of, hoping one of them would "make sense". In other cases, some participants simply preferred to talk a lot, even including things that are relatively irrelevant to the questions.

Thus, in order to take out the influence of speaking habits and to see if there were patterns in the distribution of codes across the three code categories, we decided to use "code proportion" rather than "code number" when we looked at the three categories. To get the "code proportion" of a category of a participant, we simply took "code number of a category" and divided it by "the total code number."

$$P1 = \frac{Code\ \#\ of\ \text{"}Graph\ Description\text{"}}{Total\ Code\ \#}$$

$$P2 = \frac{Code\ \#\ of\ \text{"}Graph\ and\ Data\ Interpretation\text{"}}{Total\ Code\ \#}$$

$$P3 = \frac{Code\ \#\ of\ "Instantiation"}{Total\ Code\ \#}$$

Since all codes should fell into the three main categories, the sum of the three “code proportion” should equal to one.

$$P1 + P2 + P3 = 1$$

In order to present the data for better visualization, we developed a new type of graph – the “triangle graph”. As shown in *Figure 3*, there are three axes in a triangle graph, and each of them represent one of the three main categories: the bottom axis is Graph Description, the right axis is the Graph and Data Instantiation, and the left axis is the Instantiation. The scale on the axes runs from 0 to 1, corresponding to the proportion of codes in that category.

In *Figure 3*, all 58 participants each had their own dot. Given the code proportion of the three main categories, which should add up to 1, a dot could be fixed at one single point on this graph. The four different education-level groups were represented by different symbols for comparison: The blue dots were professors (n=8), the red dots were graduate students (n=13), the green diamonds were upper-level undergraduates (n=13), and the gray squares were lower-level undergraduates(n=24).

At the beginning of the analysis, we predicted that the professors (the most expert group) would have a larger proportion of their codes in Instantiation category compared to the undergrads (the least expertise group), since we observed the pattern in *Figure 2* when we were looking at number of codes. We also predicted that the undergrads (novices) would have a larger proportion of their code in Graph Description category compare to the professors (experts), since the novices tend to notice what was presented on the “surface” of a graph whereas the experts tend to look into the information that were not presented directly but were implied.

However, what we observed seemed to be different from our prediction. *Figure 3* is a visualization of the code proportions of the three main categories of the Bar Graph. We observed that all dots, excepted for the one at the very right of the graphic, were clustered

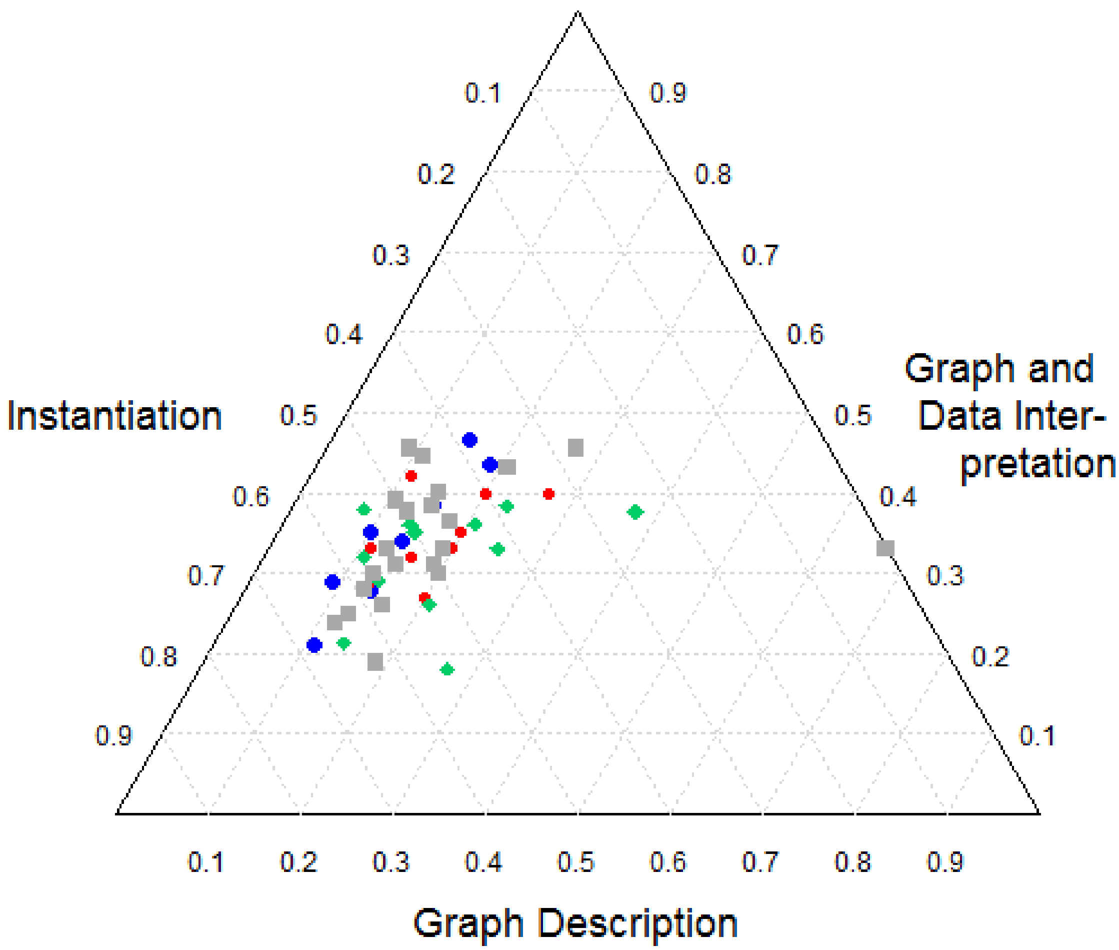


Figure 3 Visualization of the code proportions of the three main categories of the Bar Graph. The blue dots were professors (n=8), the red dots were graduate students (n=13), the green diamonds were upper-level undergrads(n=13), and the gray squares were lower-level undergrads(n=24).

together at the left of the triangle. This area corresponds to low proportion in Graph Description (0.1~0.3), relatively low proportion in Graph and Data Interpretation (0.2~0.45), and relatively high proportion in Instantiation (0.35~0.7). This means that the majority of the participants, regardless of their education-level groups, spent most of their time talking about codes in Instantiation. The four education-level groups were largely overlapping with each other, indicating that the intra-group difference was larger than the inter-group difference. However, we noticed that the dots representing the professors were more closely clustered together comparing to the three other groups, which might indicate a similar reasoning pattern shared by these experts.

The triangle graphs of the three main categories of the other four graphs (smooth line graph, scatterplot, varied line graph, histogram) expressed similar trend with the Bar Graph with few exceptions. These trends were 1) intra-group difference was larger than inter-group difference; 2) participants spend most of their time talking about codes in Instantiation; and 3) dots representing the professors were very closely clustered together.

The 2 Sub-Categories of Instantiation: Experimental Design and Graph Construction

Given our research focus, which was to "examine the difference of graph knowledge in experts and novices", we decided to dig deeper into the sub-categories to find out trends and patterns in areas that have been associated with expert practices and areas of competence. In Chapter 2, we mentioned that experts are different from novices in their abilities to appropriately construct graphs. We also talked about the role of experimental design in biology research and data representation in Chapter 2. Both of these are related to critical MRC components with graphical representations: *Invention*, or students' abilities to construct or design new representations; and *Functioning*, or students' knowledge on the "why" and "how" of graphs. As such, we looked at the Instantiation sub-categories "Graph Construction" and "Experimental Design" (*Table 5*) to see whether the proportion of these sub-categories were different among experts and novices, as we would predict based on experience and expertise.

Thus, to get the "code proportion" of a sub-category in category "Instantiation", we take "code number of a sub-category" and divided it by "the total code number of Instantiation".

$$\text{PE} = \frac{Code\ \#\ of\ \text{"}Experimental\ Design\text{"}}{Code\ \#\ of\ \text{"}Instantiation\text{"}}$$

$$\text{PC} = \frac{Code\ \#\ of\ \text{"}Graph\ Construction\text{"}}{Code\ \#\ of\ \text{"}Instantiation\text{"}}$$

$$\text{PO} = \frac{\textit{Code \# of all Other sub-categories in "Instantiation"}}{\textit{Code \# of "Instantiation"}}$$

$$\text{PE} + \text{PC} + \text{PO} = 1$$

PO, in this circumstances, represents the proportion of codes in all other sub-categories in Instantiation. The sum of PE, PC and PO should again equal to one.

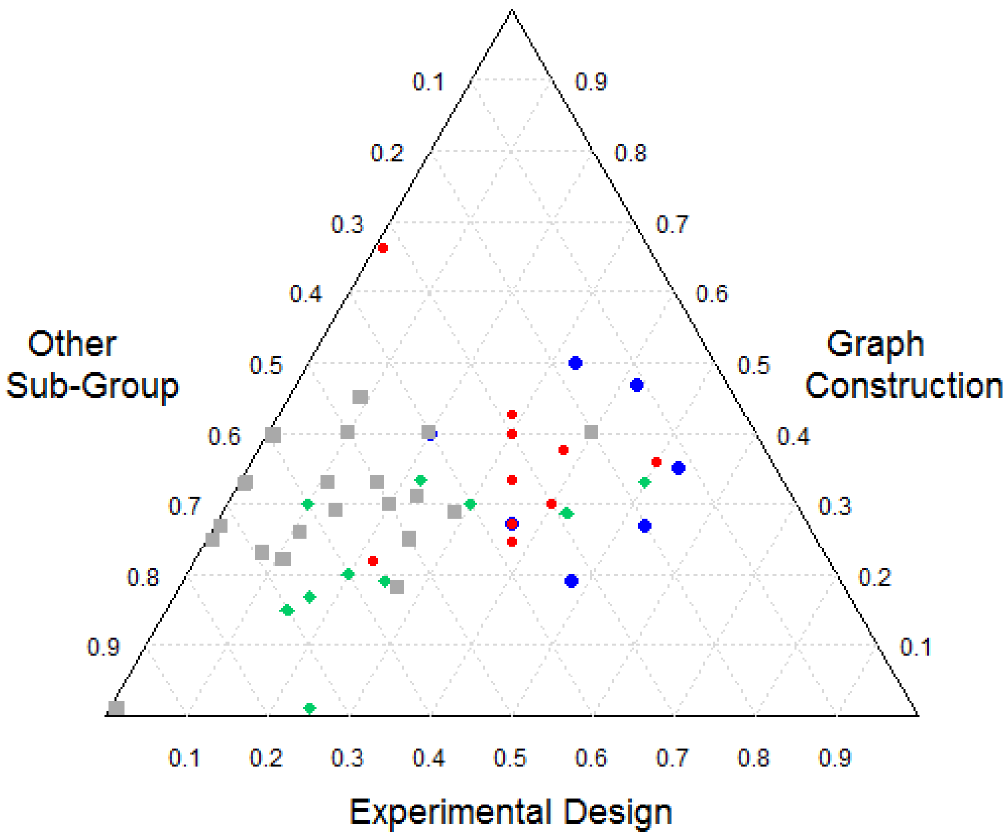


Figure 4 Visualization of the code proportions of the sub-categories within the Instantiation of the Bar Graph. The blue dots were professors (n=8), the red dots were graduate students (n=13), the green diamonds were upper-level undergrads(n=13), and the gray squares were lower-level undergrads(n=24).

Figure 4 is a visualization of the code proportions of sub-categories within Instantiation for Bar Graph. This time, a different trend was observed with inter-group differences becoming apparent; four education-level groups no longer overlapping with each other. While the proportion of codes in Graph Construction were similar among the four education-level groups, the proportion of codes in Experimental Design were different among the four groups. There is a shifting of the dots, from the right of the graph to the left of the graph, when the expertise level of the participants shifting from high to low: professors had proportions of 0.3~0.6 in Experimental Design, while LUGs had proportions of only 0~0.3. The other two groups fell in the middle of these two extreme groups.

To explore any statistically-significant differences between the four participant groups, a statistical model was built treating all participants as one single group, and another model was built treating the four different education-level groups as four groups. The two models were then compared with each other to see which one fit the data significantly better using a chi-square test. The test of "difference in likelihood" followed the equation:

$$2 * Difference\ in\ Likelihood\ \propto \chi^2{}_{df}$$

In the combined model, there were 3 parameters (the three instantiation subcategories); in the separated model, there were 12 parameters (4 participant groups * three instantiation subcategories). Thus the degree of freedom, or df, equaled 12 – 3, or 9. We then calculated the likelihood of the two models (83.74 for the combined model and 56.55 for the separated model), and that gave us the difference in likelihood, which is 83.74-56.55, or 27.19. Following the above equation, we calculated the *p* value under the chi-square formula. The *p* value was less than 0.001, and we could say there were significant differences among the four groups.

The triangle graphs of the three main categories of the other four graphs (smooth line graph, scatterplot, varied line graph, histogram) did not express trends similar with the Bar Graph with few exceptions. Instead, it was difficult to find a clear trend in those graphs, except that the dots representing the professors were more closely clustered together then the other three groups. *Figure 5* includes the visualization of the code proportions of the sub-categories within the Instantiation of the four graphs.

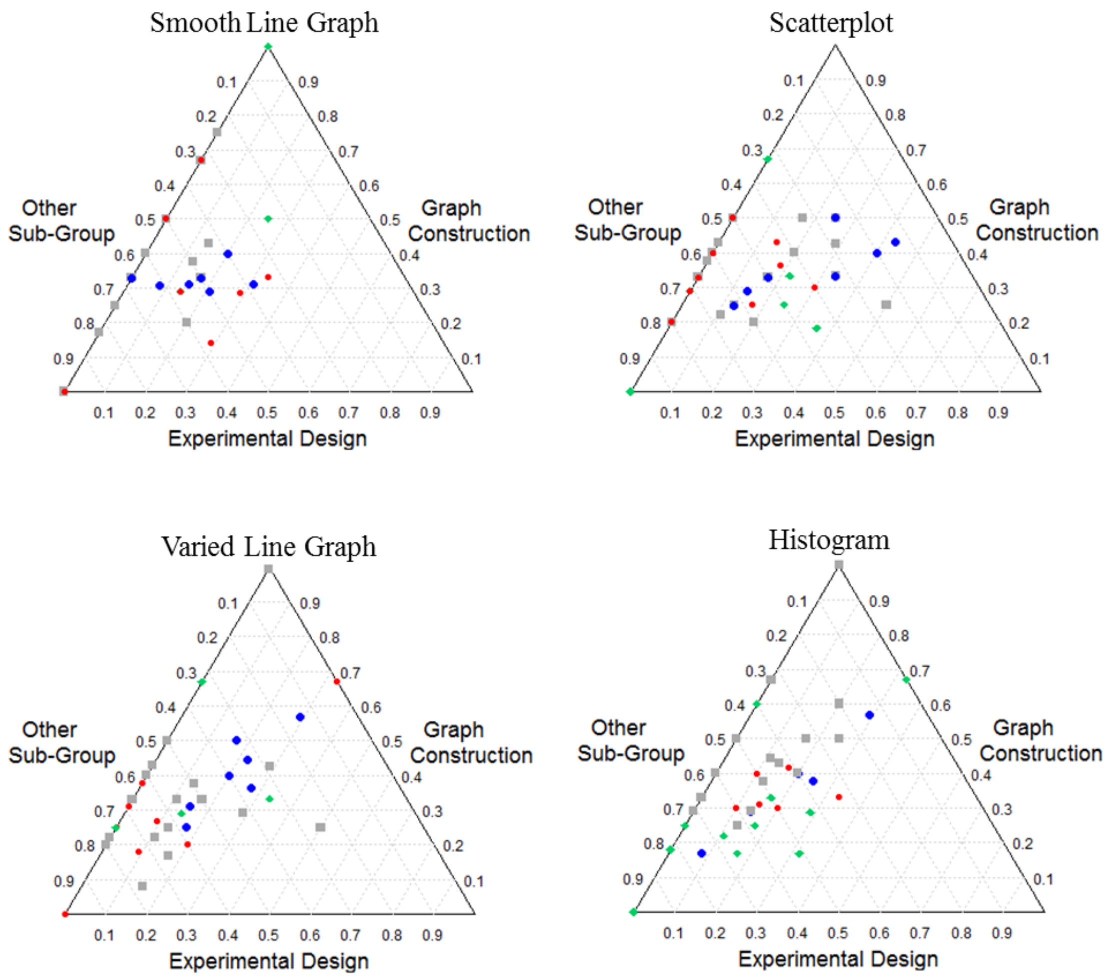


Figure 5 Visualization of the code proportions of the sub-categories within the Instantiation of the Smooth Line Graph (upper left), the Scatterplot (upper right), the Varied Line Graph (lower left), and the Histogram (lower right). The blue dots were professors (n=8), the red dots were graduate students (n=13), the green diamonds were upper-level undergrads(n=13), and the gray squares were lower-level undergrads(n=24).

Appropriateness Levels of Specific Scenarios

Within the three main categories (Graph Description, Graph and Data Analysis, and Instantiation), the participants spent the most of their time talking about the last category: Instantiation. When we looked into the transcripts, we saw some interesting trends when diving into the specific scenarios that the participants gave during interviews. For instance, a great proportion of scenarios that were given by professors were highly related to their research and were full of details. On the other hand, students tended to describe scenarios that came from their day-to-day life, and some of them even tried to force one scenario into multiple graphs. We decided to extend our investigation of graph knowledge by evaluation the quality and attributes of the scenario examples given by the participants, for we not only wanted to evaluate whether the participants' scenarios were appropriate for each graph type but also urged to explore the degree to which aspects of experimental concepts were incorporated, as appropriate, in the example scenarios.

We reviewed all the transcripts and developed a 3-scale level system for the evaluation of the appropriateness of scenarios.

Complete (the highest level): The scenario given by the participant a) contains the correct type of data; b) is appropriate for the graph type; c) includes important experimental concepts. Examples of each graph type are given below:

Bar Graph: Categorical Independent Variables, Comparison, Treatment, Control/Experimental Conditions, Multiple trials, Error bars, etc.

Line Graph (both): Continuous Variables, Change, Prediction, etc.

Scatterplot: Multiple individuals, Association, Relationship, etc.

Histogram: Continuous Independent Variables, Distribution, Normal Distribution, Skewed, etc.

(Also see *Table 3* for reference of the expected type of data and trends for the five graphs.)

Incomplete (the medium level): The scenario given by the participant is mostly correct but lacking one or two of the three components of the Complete level.

Inappropriate (the lowest level): The scenario given by the participant lacks all three components of the Complete level; or the scenario contains serious misconception; or the participant failed to give a scenario.

In *Figure 6*, the number and distribution of the three appropriateness level from all 58 participants is illustrated in a heat map. Each participant group was asked to provide a scenario for each of the 5 graphs types. From the figure, we could see that the majority of the professors' examples were complete (green). Only 2 scenarios were incomplete (yellow), and none of them were inappropriate (pink/red). There are more scenarios that were incomplete and inappropriate in graduate students, and even more in upper-level undergraduate students. At the bottom of the graph, we could see that almost a half of the scenarios from the lower-level undergraduate students were classified as incomplete or inappropriate. The number of scenarios that were incomplete and inappropriate seemed to increase downward: from the most expertise group to the least expertise group.

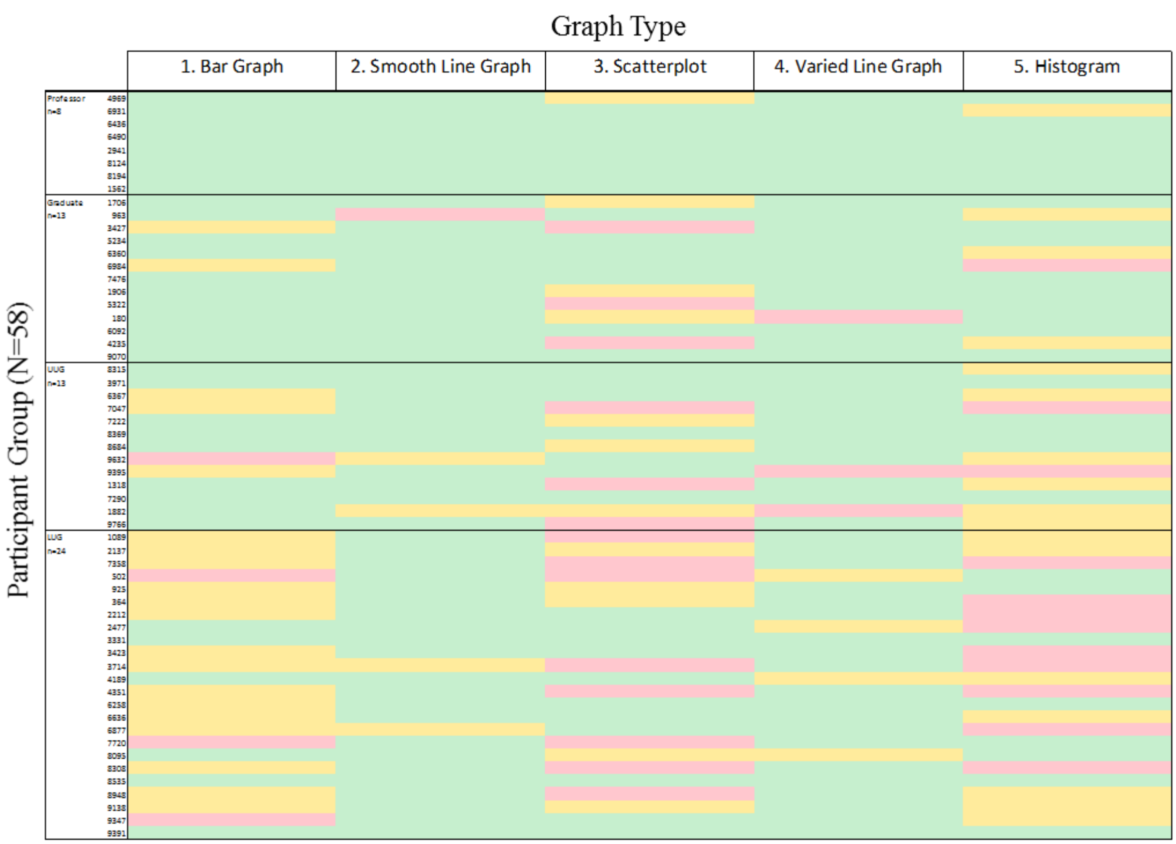


Figure 6 The appropriateness level of scenarios from all 58 participants. From the top to the bottom: professors (n=8), graduate students (n=13), upper-level undergrads(n=13), lower-level undergrads(n=24). Complete scenarios were colored in green; Incomplete scenarios were colored in yellow; Inappropriate scenarios were colored in red.

We could also make comparison across the graph type. While few participants made incomplete and inappropriate examples for the two line graphs (the smooth line graph and the varied line graph), more participants made incomplete and inappropriate examples for the bar graph, the scatterplot, and the histogram. The "top-ranked mistake" that the participants made in the three type of graphs were:

> Bar Graph: failed to mention experimental concepts, i.e., the error bars and multiple trails, the different conditions, the control and experimental groups, etc.;
>
> Scatterplot: had the misconception that the variable on the y axis should change with the continuous variable (time) on the x axis;
>
> Histogram: had the misconception that this was a bar graph and should have qualitative or categorical variables on the x axis.

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

Graphical representation in scientific communication is very common and extremely important, especially when reporting the quantitative results from scientific experiments. An increasing number of undergraduate students are now engaging in research according to a number of studies and recommendations (AAMC-HHMI, 2010; Dasgupta, Anderson, & Pelaez, 2014; Auchincloss et al., 2014; Wei, & Woodin, 2011, Laursen, Hunter, Seymour, Thiry, & Melton, 2010). It is thus crucial for college-level biology students to become familiar with representations, to be able to construct graphs from data, and to read and use graphs with experimental data appropriately. Although students from a variety of education-levels and disciplines are experiencing difficulties with graphing, we chose college-level biology students as our research targets particularly.

In the Meta-Representational Competencies framework which was developed by Sherin and diSessa (2000), a number of competencies, such as disciplinary-related experiences, spatio-visuo knowledge, statistical knowledge, and so on, were necessary for a person to use graphs fully and correctly. In our study, we only focused on exploring people's graph knowledge, which is defined as the knowledge a person has about a particular type of graph, including the graph's name and type, the graph's function, data and trend that can be displayed with the graph, the graph's affordances and limitations. Using expert-novice comparison as our theoretical framework, our main research question was: What are the differences in graph knowledge among undergraduate biology students, graduate biology students, and biology professors?

Using think-aloud interviewing as our method, we asked 58 participants from different education-levels in a department of biological sciences questions about five type of commonly used graphs in biology to test their graph knowledge. This chapter will present the summary of findings from the study and the answer the research question. This chapter will also discuss the scope and the limitation of this study, talk about the implications, and address a number of future research directions.

5.1 Summary of Findings

RQ: What are the differences in graph knowledge among undergraduate biology students, graduate biology students, and biology professors?

Within the 58 participants, 8 were biology professors (P), 13 were biology graduate students (G), 13 were upper-level undergraduate students (UUG), and 24 were lower-level undergraduate students (LUG). In terms of total number of codes in the entire interview, i.e. including their answers for all 5 questions and all 5 graphs, professors showed their expertise in analyzing graphs and creating scenarios. They had the largest number of codes compared to the other three groups, and this could be explained by looking at example answers from professors and from students.

By analyzing example answers from the professors and the students (provided in Chapter 4), it is easy to tell that professors' answers were longer compared to those of students. The codes in professors' answers were also more diverse compared to those of students in terms of number of different things they mentioned that relate to our definition of graph knowledge, such as graph's name and type, the graph's function, data and trend that can be displayed with the graph, the advantages and disadvantages of the type of the graph. This is the explanation of the result that professors have more codes overall than the three student groups (see *Figure 1*). Thus, we could say that there are expert-novice differences in graph knowledge. However, in order to analyze the differences and to identify the patterns in them, we need to take a step forward to look at the number and identity of codes across different levels.

The Three Main Categories: Expert-Novice differences in code categories

The three main categories in coding were discussed in Chapter 3 and Chapter 4: 1) Graph Description (basic description of the graph without further interpretation), 2) Graph and Data Analysis (local or global interpretation of the graph), and 3) Instantiation (concepts and reasoning that participants engage in while population the graph with data). At the beginning of the study, we predicted that the professors would spend more time talking about the Graph Interpretation and Instantiation, and that the students would spend more time talking about Graph Description compared to professors. Our reasoning included, but was not limited to, that experts tend to explore meaningful patterns and features when analyzing data, and they tend to recognize big ideas and core concepts when solving problems (Bransford, Brown, & Cocking, 1999). Experts are also capable of retrieving important relevant knowledge relatively effortless compared to novices (Bransford, Brown, & Cocking, 1999).

Our results of the number of the codes showed that in the "Graph Description" category, the overall trend was that the number of codes increased from the most expertise group (professors) to the least expertise group (lower-level undergraduate students); and in the "Graph and Data Interpretation" and the "Instantiation" categories, the overall trend was that the number of codes decreased from the most expertise group to the least expertise group (see *Figure 2*). Since we only counted unique codes, the number of codes also represented the diversity of things that participants mentioned. In Chapter 2, we mentioned that compared to the novices, the experts tend to look less at what were already presented in the graph but to recognize meaningful features and patterns from the graphs and to retrieve previous relevant experiences relatively effortless (Bransford, Brown, & Cocking, 1999). The previous part of results matched our prediction of the differences in the graph knowledge of the four education levels.

Our results also showed that the four education levels exhibited the same distribution of number of codes in the three main categories. Around half of the codes fell into the Instantiation category; some codes fell into the Graph and Data Analysis category; and the least codes fell into the Graph Description category (see *Figure 3*). This distribution of codes is not entirely unexpected because the five interview questions were developed to examine the graph knowledge and previous experiences of the participants with graphs. Questions 3 (Can you think of a specific scenario to fit the graph?) and Question 4 (Why did you choose that specific scenario?), especially, were asked to direct the participants to think of a specific example. Thus, it was not surprising that most of the codes under these two questions fell into the Instantiation category. Question 5 (What kind of trend can be represented by this type of graph?) directed the participants to think about the trend and function of the graph, and Question 2 (What type of data can fit this graph?) asked the participants to examine the type and the pattern of the data. Thus, it could be predicted that most of the codes under these two questions would fall into the Instantiation category and the Graph and Data Interpretation category, regardless the amount of graph knowledge the participant actually had. While the similar distribution of codes across the categories between the participant groups is not unexpected, the identity of those codes was different for some of the graphs.

Expert-Novice differences in Instantiation

We were not able to distinguish trends in graph knowledge at the main category level; the professors and the students had a similar distribution of the number of codes falling into the three main categories. We then decided to move forward to the sub-category level.

The relationships among experiments, data, and representation in biology are strong: the data that are represented in graphs come from experiments. Undergraduate students have difficulties with scientific experiments had already been documented in a number of studies (Wei & Woodin, 2011; Shi, Power, & Klymkowsky, 2011; Dasgupta, Anderson, & Pelaez, 2014). Could the students' lack of experiences with scientific experiments be one of the causing factors of their difficulties in graphing? Particularly, could the difficulties that the students had were due to difficulties in linking data representation with experimental design, or that students were unable to plot experimental data appropriately in graphs? To explore this topic, we chose two sub-categories within the Instantiation category to study the pattern of the code distribution: “Experimental Design” and “Graph Construction”.

While no difference in Graph Construction is observed, our analysis reveals interesting patterns in Experimental Design. In Bar Graph only, while all four education groups spent a lot of time talking about Instantiation, the four groups spent different proportion of their codes talking about Experimental Design, which is the description of experiment-related design in the participant’s example or scenario of the graph (See *Figure 4*). This sub-category includes, but was not limited to, multiple trials, different conditions, treatments, measurements, experimental/control groups, observational data, or the participant simply mentioning “experiment” or “experimental design” (See the full coding scheme in *Appendix A* for reference). There are significant differences in the proportion of experimental design: professors spent the most of their codes talking about things related to experiments, followed by the graduate students, then the upper-level undergraduate students; the lower-level undergraduate students spend the least of their codes talking about experiments-related stuff.

From the above results from Experimental Design sub-category of the bar graph, the professors appear to have been very familiar with creating and analyzing graphs with experimental data. This could be due to the considerable amount of experiences they had with their own research, from reading graphs in journal papers and textbooks, such as teaching their students about graphs, and so on. Professors understood that the graphs were used for summarizing data, conveying trends, and efficiently communicating, thus they were trained for years to read and create graphs and they knew that the graph could not exist without data from experiments or observations.

The graduate students also presented a relatively good understanding of the importance of experimental design. This could be because they also spent a lot of time doing their research and studying papers, although not as much trained as professors. Given that most of the graduate students were also teaching assistant, they might also be reinforced the importance of graphs in their preparation of teaching.

The two groups of the undergraduate students were the two groups spent the least of their codes of experimental design, which not necessarily indicated their unfamiliarity with experiments but strongly suggested their unawareness of the linkage between data representation and scientific experiments. They were ‘not thinking about where the data came from’ when they were looking at the graph with data in it; even with the lead of the question “Can you think of a specific scenario to fit this graph?”, most of them still had difficulties talking about the source of the data or the scenario. This phenomenon could be due to that the undergraduate students were lacking one or more components in the MRC; particularly, they were having difficulties in understanding the purpose of the representations and why we use them. The students in LUG and UUG groups, especially the former, are the ones that had the least experiences with biological science compared to the professors and the graduate students. Thus, they were relatively unfamiliar with disciplinary-related knowledge and experiences, as well as statistical knowledge and spatio-visual knowledge, all of which were necessary for appropriately using graphs.

We need to point out here, that the above pattern of codes in Experimental Design only existed in the Bar Graph. The other four graphs did not have a clear pattern in distribution compared to the Bar Graph. We were not surprised to find it out, for in the five graphs, the bar graph with error bars is the one that was used the most in biology experiments and, thus, should be easiest for the participants to talk about (Weissgerber, Milic, Winham, & Garovic, 2015. Also see *Table 3*). The advantages of bar graphs are to compare categorical data, percentages, or summary statistics from multiple groups (Schriger & Cooper, 2001), which is suitable for scientific experiments with control groups and treatment(s). Thus, some of the experimental terms such as “experimental group and control group” and “significant difference”, is best aligned with the bar graph but not the other four graph types.

Understanding the Graphs: Appropriateness of Scenarios

After analyzing the identity and distribution of graph knowledge of the participants, we extended our investigation by evaluation the quality and attributes of the scenario examples given by the participants, not to see certain codes were or were not present but were they appropriate and was the scenario aligned with a specific graph type.

In Chapter 4, we presented the results of the scenario evaluation. Three appropriateness levels were created and used: Complete (the scenario given by the participant a) contains the correct type of data; b) is appropriate for the graph type; c) includes important experimental concepts); Incomplete (the scenario given by the participant is mostly correct but lacking one or two of the three components of the Complete level); and Inappropriate (the scenario given by the participant lacks all three components of the Complete level; or the scenario contains serious misconception; or the participant failed to give a scenario). See *Chapter 4* and *Table 3* for reference of the expected type of data and trends for the five graphs.

From the results, we concluded that the ability to analyze graphs correctly and make appropriate scenarios increased with the expertise level (*Figure 6*). I.e., the professors (experts) were more likely than the undergraduate students (novices) to be able to read and understand a graph, to link the graph with previous personal experiences, to think of a type of experimental data that could be presented in the graph, and to come up with specific data that could fit the graph.

This pattern existed in all five graphs. Thinking back to the analysis of the previous results, we could conclude that having graph knowledge in Graph and Data Interpretation and Instantiation would largely contribute to appropriately understanding graphs and data, particularly an awareness of the linkage between scientific graph and experimental data.

When looking at scenario appropriateness across the *graph type* instead of the participants' education-level (see *Figure 6*), we could see that the participants did well with the two line graphs (the smooth line graph and the varied line graph). Only a few of them gave incomplete or inappropriate scenarios. On the contrast, all groups did not as well in the other three graphs: the bar graph, the scatterplot, and the histogram. Almost a half of participants gave incomplete or inappropriate scenarios, and the reasons behind each graph were unique. For the bar graph, many participants failed to provide detailed explanation for their scenarios, particularly when it came to experimental concepts, such as treatments, conditions, control and experimental groups, multiple trials, etc. The participants named variables and put them on the axes without giving any further explanation of their scenarios, and this could be because the lack of awareness of the relationship between graphs and experiments. For the scatterplot, lots of participants had misconceptions on the relationship between the two variables: they gave inappropriate types of variables and believed that the variable on the y axis should change with the variable on the x axis. A number of participants tried to fit previous examples (from what they gave for the smooth line graph) into the scatterplot by putting time on the x axis. For the histogram, many participants think it as a bar graph and gave inappropriate independent variables, such as qualitative or categorical variables. This could be due to an incomplete or partial understanding of the type of graph and the data that can be displayed in it.

5.2 Scope, Limitation, Future Directions, and Implications for Instruction

Scope, Limitations and Future Research Directions

Our study provides an explanation of students' difficulties of biology graphs in terms of graph knowledge. Our results also have major implications for biology teaching and course design. However, this study only targeted the higher-education students, collecting data from college-level students, graduate students, and professors. Although students have learned about graphs and obtained graph knowledge since K-12, we did not have the chance to look at the graph knowledge of younger students, which limited the scope of this study. This could serve as a future direction of the study of biology graph and graph knowledge.

Another limitation could be that we only used five commonly used graphs to test our participants' graph knowledge: a bar graph, two line graphs, a scatterplot, and a histogram. Compare to other studies of external representations (such as Cox, Romero, du Boulay, & Lutz, 2004, which used 90 different representations in the experiments), this could be a limitation of scope. Nevertheless, we want to point out that the five graphs we used were carefully chosen by twelve biology faculty members from within the department that the student populations were recruited, according to the importance of knowing the graph and the frequency of using the graph in teaching. We believed that, although we only used five graphs, these graphs appropriately represented the most important types of graphs that college-level students should become familiar with and know how to use.

In Chapter 4, we used the triangle graphs to provide the visualization of the results of code distribution. The triangle graph is developed with the help of Purdue Statistic Consulting Program, and the graph is good at showing distribution of the 3-dimension positions of groups of individuals. It is able to show the inter-group differences and the intra-group differences clearly, especially when there are multiple individuals in each groups *and* there are multiple groups. On the other hand, outliers in triangle graphs could get unwanted attention from readers; also, if there are points overlapping with each other, the readers might not notice the overlaps and could get misled from the data. Overall, the triangle graph serves as a great tool to visualize multiple data points that fall in multiple groups, especially in a 3-dimension environment.

Another limitation exists in the using of expert-novice comparison. In our study, we treated the results from professors as if they were experts who had acquired excessive amount of graphing experiences that they represented the highest expertise level. However, although these faculty members showed expertise in graph construction and evaluation, a few of them gave incomplete or inappropriate answers to our questions. This fact is actually consistent with a number of Roth's studies, in which he pointed out that even professors had difficulties communicating using graphs. In Roth's most recent study (2013), he conducted an ethnographic study of a science laboratory aimed at studying the absorption of light in the eyes of salmonid fish. He found out that, when these science experts used graphs that are generated many steps downstream from their study, even if the graphs show the results of their own data, these scientists started to experience difficulties understanding and interpreting the graphs correctly. Roth and Thom's earlier study (2009) also indicated a similar trend: 17 physicists and 16 biologists were asked to interpret graphs from biology introductory courses, and the results showed that only 27% of the scientists were able to give correct answers on a graph that is similar to the oxygen-shrimp frequency graph. He took one more step to look at these 27% of scientists, and noticed that out of these 9 people, 7 of them were biologists who were teaching at undergraduate levels.

In an effort to improve the quality of science communication at the highest level, scientific research journals have begun to advocate for more transparent and appropriate graphing of data featuring editorials and regular pieces on data displays. A paper by Rougier, Droettboom, and Bourne (2014) targeted at scientists who used graphs to visualize their data. The researchers called for an improvement of figure and graph design and explained some common pitfalls in using graphs in communication for scientists. In a paper by Weisenberger et al. (2015), the authors also suggested that scientists urgently need to improve their usage of appropriate representations to present the data, and they strongly recommended training investigators in data presentation, especially the selection of graph types according to data types. Specifically, they suggest replacing graph types such as the bar graph to graphs which display all of the data such as categorical dot plots.

Finally, BioMed Central has a regular series called 'What's Wrong with this Picture' that aims to educate its readership on the potential misrepresentations of data. Future studies could be directed to solve this issue by studying and exploring the relationship between scientists' issues with graphs and their graph knowledge.

Implications for Instruction

The results of our study align with the MRC framework, indicating that multiple competences *are* required for students to use graphs fully and correctly. Instructional design of graphing should be targeted at improving students' *graph knowledge*, to help students perform better in science curriculum. Scientific graphing should be incorporated into courses, in which students should be taught common graph knowledge that would help them perform better in graph construction, interpretation, and evaluation. Specifically, the implication for future biology instruction from our study is to improve students' graphing competency by emphasizing the linkage between scientific representations and experimental data. For instance, when students are involved in research and lab work, instructors should encourage the students to collect the data, to make representations using the data, *and* to draw conclusions from the representations. When instructors use graphs in lectures to convey ideas or theories to students, they should talk about the resource of the data in the graphs and the experimental or observational settings to help students understand the graph. By showing students that scientific experiments, data, and graphs are interconnected, students should be able to get a big picture of the process of scientific research, which would support their future career as scientists.

LIST OF REFERENCES

LIST OF REFERENCES

AAMC-HHMI Committee. (2009). Scientific foundations for future physicians. *Washington, DC: Association of American Medical Colleges*.

Angra, A., & Gardner, S. M. (2016). Development of a framework for graph choice and construction. *Advances in physiology education*, *40*(1), 123-128.

Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., ... & Towns, M. (2014). Assessment of course-based undergraduate research experiences: a meeting report. *CBE-Life Sciences Education*, *13*(1), 29-40.

Berg, C. A., & Phillips, D. G. (1994). An investigation of the relationship between logical thinking structures and the ability to construct and interpret line graphs. *Journal of Research in Science Teaching*, *31*(4), 323-344.

Berg, C. A., & Smith, P. (1994). Assessing students' abilities to construct and interpret line graphs: Disparities between multiple-choice and free-response instruments. *Science Education*, *78*(6), 527-554.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American journal of Physics*, *62*(8), 750-762.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.

Brewer, C. A., & Smith, D. (2011). Vision and change in undergraduate biology education: a call to action. *American Association for the Advancement of Science, Washington, DC*.

Bruner, J. S. (1966). *Toward a theory of instruction* (Vol. 59). Harvard University Press.

Bruno, A., & Espinel, M. C. (2009). Construction and evaluation of histograms in teacher training. *International Journal of Mathematical Education in Science and Technology*, *40*(4), 473-493.

Bryant, P. E., & Somerville, S. C. (1986). The spatial demands of graphs. *British Journal of Psychology*, *77*(2), 187-197.

Carswell, C. M. (1992). Choosing specifiers: An evaluation of the basic tasks model of graphical perception. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *34*(5), 535-554.

Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal of teacher education*, *39*(3), 25-31.

Clase, K. L., Gundlach, E., & Pelaez, N. J. (2010). Calibrated peer review for computer-assisted learning of biological research competencies. *Biochemistry and Molecular Biology Education*, *38*(5), 290-295.

Cleveland, W.S. (1984). Graphs in Science Publications. *The American Statistician*, 38:4, 261-269.

Cooper, R. J., Schriger, D. L., & Tashman, D. A. (2001). An evaluation of the graphical literacy of Annals of Emergency Medicine. *Annals of emergency medicine*, *37*(1), 13-19.

Cooper, R. J., Schriger, D. L., Wallace, R. C., Mikulich, V. J., & Wilkes, M. S. (2003). The quantity and quality of scientific graphs in pharmaceutical advertisements. *Journal of general internal medicine*, *18*(4), 294-297.

Crabtree, B. F., & Miller, W. F. (1992). A template approach to text analysis: developing and using codebooks.

Dasgupta, A. P., Anderson, T. R., & Pelaez, N. (2014). Development and validation of a rubric for diagnosing students' experimental design knowledge and difficulties. CBE-Life Sciences Education, 13(2), 265-284.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and instruction*, *22*(3), 293-331.

diSessa, A.A., & Sherin, B. L. (2000). Meta-representation: An introduction. *The Journal of Mathematical Behavior*, *19*(4), 385-398.

Duffy, T. M., & Jonassen, D. H. (1992). Constructivism: New implications for instructional technology. *Constructivism and the technology of instruction: A conversation*, 1-16.

Few, S. (2004). Show me the numbers. *Analytics Pres*.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education*, 124-158.

Humphrey, P. B., Taylor, S., & Mittag, K. C. (2014). Developing consistency in the terminology and display of bar graphs and histograms. *Teaching Statistics*, *36*(3), 70-75.

Laursen, S., Hunter, A. B., Seymour, E., Thiry, H., & Melton, G. (2010). Undergraduate research in the sciences: Engaging students in real science. John Wiley & Sons.

Leivas Pozzer, L., & Roth, W. M. (2003). Prevalence, function, and structure of photographs in high school biology textbooks. *Journal of Research in Science Teaching*, *40*(10), 1089-1114.

McFarland, J. (2010). Teaching and Assessing Graphing Using Active Learning. *MathAMATYC Educator*, *1*(2), 32-39.

Meletiou-Mavrotheris, M., & Lee, C. (2010). Investigating college-level introductory statistics students' prior knowledge of graphing. *Canadian Journal of Science, Mathematics and Technology Education*, *10*(4), 339-355.

Mevarech, Z. R., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, *32*(3), 229-263.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press.

Novick, L. R. (2004). Diagram literacy in preservice math teachers, computer science majors, and typical undergraduates: The case of matrices, networks, and hierarchies. *Mathematical thinking and learning*, *6*(3), 307-342.

Olson, S., & Riordan, D. G. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. *Executive Office of the President*.

Orgill, M., & Bodner, G. (2007). Locks and keys. *Biochemistry and Molecular Biology Education*, *35*(4), 244-254.

Picone, C., Rhode, J., Hyatt, L., & Parshall, T. (2007). Assessing gains in undergraduate students' abilities to analyze graphical data. *Teach. Issues Exper. Ecol*, *5*, 1-54.

Postigo*, Y., & Pozo, J. I. (2004). On the road to graphicacy: The learning of graphical representation systems. *Educational Psychology*, *24*(5), 623-644.

Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. Cambridge University Press.

Roth, W. M., & Thom, J. S. (2009). Bodily experience and mathematical conceptions: From classical views to a phenomenological reconceptualization. *Educational Studies in Mathematics*, *70*(2), 175-189.

Roth, W. M. (2013). Undoing decontextualization or how scientists come to understand their own data/graphs. *Science Education*, *97*(1), 80-112.

Rougier, N. P., Droettboom, M., & Bourne, P. E. (2014). Ten simple rules for better figures. *PLoS Comput Biol*, *10*(9), e1003833.

Schadewitz, N., & Jachna, T. (2007). Comparing inductive and deductive methodologies for design patterns identification and articulation.

Schriger, D. L., & Cooper, R. J. (2001). Achieving graphical excellence: suggestions and methods for creating high-quality visual displays of experimental data. *Annals of emergency medicine*, *37*(1), 75-87.

Shaw, E. L., Padilla, M. J., & McKenzie, D. L. (1983, April). An examination of the graphing abilities of students in grades seven through twelve. In *meeting of the National Association for Research in Science Teaching*.

Shi, J., Power, J. M., & Klymkowsky, M. W. (2011). Revealing student thinking about experimental design and the roles of control experiments. *International Journal for the Scholarship of Teaching and Learning*, 5(2), 8.

Speth, E. B., Momsen, J. L., Moyerbrailean, G. A., Ebert-May, D., Long, T. M., Wyse, S., & Linton, D. (2010). 1, 2, 3, 4: infusing quantitative literacy into introductory biology. *CBE-Life Sciences Education*, *9*(3), 323-332.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, *27*(2), 237-246.

Treagust, D. F., & Tsui, C. Y. (Eds.). (2013). Multiple representations in biological education. Springer Science & Business Media.

Wei, C. A., & Woodin, T. (2011). Undergraduate research experiences in biology: alternatives to the apprenticeship model. CBE-Life Sciences Education, 10(2), 123-131.

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol*, 13(4), e1002128.

APPENDIX

APPENDIX

The Coding Scheme with Examples

1) Graph Description: Explicit knowledge, or what people can get directly from the graph without further interpretation
    1.a Description of Graph: The description of components or shape of a graph without any interpretation
    1.b Type of Graph: The mention of the type of graph

Code No.	Code	Definition and Example
1.a.1	Description about the axes	The description of the axis of the graphs, such as the name on the axis, the scale and the unit of axis, etc. Example: [P1562, Graph1] "Well I see a bar graph with two axes… no indication what the axes and also no units on axes."
1.a.2	Description about the graph	The description of the data and/or the elements in the graphs, such as data points, bars, lines, error bars, etc. Example: [LUG6788, Graph1] "There's just two different colored bars in what looks like two different sections."
1.a.3	Lack of resources	When a participant point out that the graph is lacking a resource, such as best fit line, the key, the labels, the title, etc. Example: [LUG3423, Graph1] "It doesn't have either axis labeled, it doesn't have a title…"
1.b.1	Right type of graph	The mention of the right type of graph. Example: [G3427, Graph3] "It's a scatterplot."
1.b.2	Wrong type of graph	The mention of the wrong type of graph. Example: [G6984, Graph5] "It is a bar graph."

2) Graph and Data Interpretation: Implicit knowledge, interpreting parts of a graph or see overall trend/function of the whole graph without a specific scenario
    2.a Experimental Design: The mention of experimental design using information provided by the graphs, not from their own scenario

2.b General Conclusion: A general conclusion about the function of the graph or the take home message of the scenario
2.c Statistical Terms and Data Analysis: The mention of specific statistical terms or functions
2.d Trends: The description of the partial or overall trend of a graph
2.e Variables: The direct reference to a variable

Code No.	Code	Definition and Example
2.a.1	Multiple Trails and Measurement	The mention of using multiple trails/times to get an average and/or error bars Example: [G0963, Graph1] "They have error bars, shows that they've done a lot of trails."
2.a.2	Treatment Groups	Difference between treatment groups, such as mention of a control group and/or experimental group Example: [LUG0364, Graph1] "…the two bars are about the same height, it's probably a control group; your experimental condition… is probably this one that is significantly different from the other three."
2.a.3	Conditions	Mention of different conditions without pointing out specific name or type of the conditions Example: [UUG8315, Graph1] "Like I said, probably something with two conditions here and here."
2.b.1	Graph Function	The mention of the function of a type of graph Example: [UUG9632, Graph5] "This one can show like if it's a bell curve or if it's left skewed or right skewed."
2.b.2	Take Home Message	The description of the summary or the take home message of a specific scenario Example: [G5322, Graph1] "That *betf* expression at 6 hour is much higher in activated cells than non-activated cells, and that the *betf* expression is higher at 6 hours than at 12 hours."
2.c.1	Statistical Terms	The mention of statistical terms, such as: average, standard errors, mode, median, range, variance, etc. Example: [UUG7290, Graph1] "So… these bars are not discrete individuals but some kind of average of individuals and we are looking at the variation… in these individuals."
2.c.2	Statistical Analysis	The mention of types of statistical analysis, such as: trendline or best-fit-line, R^2, degree of correlation, statistical significance Example: [P6436, Graph3] "A good correlation, it's very tight."

2.d.1	Association	The description of the trend of a graph as association between variables, such as correlation and/or regression Example: [G4235, Graph3] "This is a classic example of a regression, or a linear regression, so you have a lot of data and you can see the trend in here…"
2.d.2	Comparison	The description of the trend of a graph as comparison Example: [G6984, Graph1] "Any data which is a comparison of… two groups, where same measurement if being made…"
2.d.3	Difference or Change	The description of the trend of a graphs as difference between variables, changes over time, or growth rate, etc. Example: [G0180, Graph2] "This graph shows the rate of something. It's time for quantitation…"
2.d.4	Distribution	The description of the trend of a graph as showing distribution Example: [P4969, Graph5] "It's the probability of something, like a distribution of students' grade in the class… kind of a like a bell-shape curve."
2.e.1	Independent Variables	The direct mention of the phrase "independent variable" Example: [LUG6788, Graph1] "If you want to, but they're being graphed against the same independent and dependent variables."
2.e.2	Dependent Variables	The direct mention of the phrase "dependent variable" Example: [G4235, Graph4] "I mean a first phase in which the independent variable increase as the dependent variable increases."
2.e.3	General reference to a variable	Mention of "variable" without specifying variable type Example: [LUG9391, Graph3] "They're not related any other way besides the two variables that you are looking at."
2.e.4	Variable type	Mention of variable type, such as: categorical, continuous, numerical, observational, etc. Example: [G1906, Graph5] "Again I would say this is for continuous variables."

3) Instantiation: Populating graphs with data and conceptual understanding of the graph with linking the reasoning with previous personal knowledge or experiences.

    3.a Example Type: The type of the specific example/scenario the participant generate; this is not to be confused with the type of graph source

3.b Experimental Design: The description of experiment-related design in the participant's example/scenario of the graph
3.c Statistical Terms and Data Analysis: The mention of specific statistical terms or functions
3.d Metacognition or Metacognitive Monitoring: The "knowing about knowing"; students' reflection on their own knowledge and thought processes in real time
3.e Graph Construction: The description of graph components when drawing their graphs according to their examples/scenarios
3.f Mention of Other Graph: When participants mention other graphs
3.g Source of Graphs: The mention of the previous experiences with the graph
3.h Trends: The description of the partial or overall trend of a graph

Code No.	Code	Definition and Example
3.a.1	Personal Example	The type of example/scenario is based on participant's personal knowledge, such as: personally experienced in life, research-related, teaching, etc. Example: [P1562, Graph2] "Talking about actin… so spindle of actin, the actual length. And sometimes also shown as a percentage of total… of actin… It reflects the change in production of the F-actin and it reaches… which must be below 100% because there must be some monomeric actin to keep the process going. So basically it means that you have 80%, 90% F-actin and 10% or 20% G-actin that still at single and is being removed."
3.a.2	Previous Example	The type of example/scenario is based on examples that were given in previous tasks in the interview, including: bacteria growth, moth comparison, and leaves growth examples. Example: [LUG6788, Graph1] "You could use the moth example that we did in the previous example… the darker colored bars could represent the dark moths, and then the lighter colored are the white ones, and then each of the groupings represent the different time periods."
3.a.3	Vague Example	The given example/scenario is very vague, lack of details, or the participant failed to give an example/scenario Example: [LUG7358, Graph3] "… In high school, plotting or something, but really… I am not sure."
3.b.1	Experimental Function	The description of experimental design in specific scenarios, such as: multiple trails, different conditions, treatment groups, etc.

		Example: [G0963, Graph1] "So I would've done is, the first two is before treatment with drug, and this is the mutant, treatment with drug."
3.c.1	Statistical Terms	The mention of statistical terms in the scenario/example, such as: average, standard errors, mode, median, range, variance, etc. Example: [P6490, Graph1] "…these are most likely standard deviations or standard errors of the mean."
3.c.2	Statistical Analysis	The mention of types of statistical analysis in the scenario/example, such as: trendline or best-fit-line, R^2, degree of correlation, statistical significance Example: [G6092, Graph1] "…You'll find the whole group who didn't wash their hands, the bacteria number increase significantly, but not will the group who have washed hands."
3.d.1	Appropriateness	Participant evaluate their own example/scenario, without providing any correction Example: [G1706, Graph5] "…I know it's a very bad explanation but that's the best I can do now."
3.d.2	Correction of Example	Participant evaluate their own example/scenario and provide correction of the same example/scenario Example: [G0963, Graph2] "Growth initially with time… it shouldn't be like this. It goes up. Okay, I'll rewrite. So this time, growth on the x axis, and time on the y axis."
3.d.3	Provide Another Example	Participant evaluate their own example/scenario and provide a better example/scenario Example: [G6092, Graph5] "I think they have to be something has equal difference. Oh! There might be a better example… okay now I found out that in the center the values are high… looks like something follow normal distribution, so…"
3.e.1	Naming Axis	When participants attribute specific names to axes Example: [UUG8369, Graph3] "x is the growth rate, y is the ROS, so that works better, it's hard to get a third thing here."
3.e.2	Attribute General Variables to Axis	When participants attribute a type of variable to axes Example: [LUG8308, Graph1] "So just use a general practice, placing the dependent variable on the y axis, and x axis with the independent variable."
3.e.3	Naming Variable	When participants attribute specific names to types of variables Example: [LUG2212, Graph1] "…For most of graphs I dealt with, time is a very comment independent variable."

3.e.4	Type of Data	The description of the types of data used in specific examples/scenarios Example: [UUG6367, Graph5] "Maybe some data that are not continuous… They have separate categories."
3.e.5	Adding a Title	When participants attribute a title to the graph
3.f.1	Another Type of Data	The mention of a different type of graph other than the current graph Example: [LUG2212, Graph1] "It can't be percentage, because it would then be a pie chart."
3.f.2	Ideal Graph	The mention of an ideal graph of the current graph Example: [G5322, Graph5] "A normal distribution – it might be a little bit skewed. But it's pretty much a normal distribution."
3.f.3	Same Type of Graph in Different Shapes	The mention of the same type of the graph as the current graph but in different shapes Example: [LUG0364, Graph1] "Lot of graphs lie this are in the paper s that I read, and… actually most of them are usually opposite… so you see a decrease in one group relative to the control."
3.g.1	Cannot Remember Source	When the participants fail to remember their previous experience with the type of graph Example: [G0180, Graph4] "(So have you seen this graph before?) I might have… but I don't remember anything."
3.g.2	Learning in Class	The mention of encountering/using the type of graph in classes Example: [UUG7290, Graph2] "Professor (Name) made us do so many of these… drilled into me."
3.g.3	Personal Experienced in Life	The mention of encountering/using the type of graph in personal life Example: [LUG3423, Graph4] "Well I was looking at the weather this morning, and it was a graph like this. That's what I've seen it most used for."
3.g.4	Previous Examples	The mention of encountering/using the type of graph in previous examples in the same interview, including: bacteria growth, moth comparison, and leaves growth examples. Example: [UUG1318, Graph1] "Just because the entire bacteria situation was fresh in my mind, and it was easy to take it from the previous example and change it a little bit to fit this graph."
3.g.5	Research or Laboratory Related	The mention of encountering/using the type of graph in research or laboratory related experiences

		Example: [P1562, Graph1] "Yes, in my research. And in other people's research."
3.g.6	Teaching or Mentoring	The mention of encountering/using the type of graph in teaching or mentoring students Example: [P6490, Graph2] "I think it's because I teach this. When I am teaching population dynamics and I have the students look at this."
3.g.7	Textbook or Scientific Papers	The mention of encountering/using the type of graph in textbook or scientific papers Example: [P6490, Graph4] "I think it's mainly because I've seen data in textbooks and lectures that are plotted this way to show how population numbers change over time."
3.h.1	Association	The description of the trend of a graph as association between variables, such as correlation and/or regression Example: [UUG9397, Graph3] "(What is the take home message?) So test 1 and test 2 scores are positively correlated."
3.h.2	Comparison	The description of the trend of a graph as comparison Example: [LUG8095, Graph1] "It's really just trying to show a side by side comparison for each trial, so tube 1 and tube 2 you just have a comparison between the two."
3.h.3	Difference or Change	The description of the trend of a graphs as difference between variables, changes over time, or growth rate, etc. Example: [LUG2477, Graph2] "It shows the trend of the rate of reaction over time…? Like how the rate changes over time."
3.h.4	Distribution	The description of the trend of a graph as showing distribution Example: [P6931, Graph 5] "It tells you about how evenly distributed the species are in the community."