POINT LOMA NAZARENE UNIVERSITY

An Exploration of Graphing Software Program Selection and Relative Proficiency Among Undergraduate Biology Students

A thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Science

in General Biology

by

Elizabeth Lynn Ferguson

Committee in charge:

Dr. April Maskiewicz, Chair

Dr. Dianne Anderson

Dr. Lori Carter

2014

The thesis of Elizabeth Lynn Ferguson is approved, and it is acceptable in quality and form for publication:

____

____

____

Chair

Point Loma Nazarene University

2014

Table of Contents

List of Figures and Tables

Acknowledgement

I would like to express my appreciation for the guidance and support of my committee members, Dr. April Maskiewicz, Dr. Dianne Anderson, and Dr. Lori Carter. In particular, I am profoundly grateful to my committee chair, Dr. Maskiewicz for the tireless efforts, edits, creativity, and engagement through the learning process of this master thesis, without which this effort might not have come to fruition. I would also like to thank the participants for their willingness to share their extended time and attention throughout the entirety of this research project. I am very grateful to my family and friends, who have supported me throughout entire process with their patience and reinforcement. Finally, I am eternally appreciative for the love, encouragement and hardworking inspiration provided by my dear husband, James.

Abstract of the Thesis

An Exploration of Graphing Software Program Selection and Relative Proficiency Among Undergraduate Biology Students

by

Elizabeth Lynn Ferguson

Master of Science in General Biology

Point Loma Nazarene University, 2014

Dr. April Maskiewicz, Chair

Although the use of technology in biological career paths is increasing globally, studies indicate that United States undergraduate biology students are not prepared for this technologically advanced workforce. One potential element of this issue is that students may not have adequate training with technology-based tools as it applies to undergraduate biology curriculum. To address this problem, my exploratory study examines student proficiency for graphing with Excel, and aptitude for an industry standard graphing program, SigmaPlot 12.0. Student preference was included in this study to gauge their overt and underlying perspectives of each software program. Four Junior and six Sophomore students participated in a teaching experiment that consisted of simple and complex graphing instruction using both Microsoft Excel and SigmaPlot 12.0. At the close of the teaching experiment, students created graphs using the program of their choice, and then participated in an interview. Interview results indicate a preference for Excel with simple graphing tasks and SigmaPlot 12.0 for more complex graphing tasks; yet, 80% of students indicated that Excel is burdensome. Analysis of the steps students took to generate each graph in their final session showed an overall lower proficiency level with Excel versus SigmaPlot for five out of six tasks. Analysis of students' graph products also suggest that biology students overwhelmingly find Excel-based graphing challenging, and use of alternate software found in the biology industry is recommended for addressing this issue in educational settings.

Introduction

The science, technology, engineering and math (STEM) workforce in the United States is profoundly dependent on the success of the country's undergraduate students. However, issues of undergraduate skill level, student retention, and the inability of recent graduates to address complex interdisciplinary problems in their careers has caused widespread concern that undergraduate biology students are unprepared to meet the needs of a growing industry (National Research Council, 1996; National Research Council, 2003). Undergraduate reform in STEM programs launched in the late 1990's focused initially on the inability of students to solve ill-defined problems inherent in research and deficits in group communication and teamwork skills (Department of Education, 2000; National Research Council, 1996; National Research Council, 2003). In addition to finding college graduates who were entering the workforce unprepared, the National Research Council (1996) identified problems with retaining students in science and technology undergraduate programs. Approximately half of the population entering college with an interest in pursuing a STEM career dropped out of programs after an introductory class (Seymour & Hewitt, 1994). Studies found that these problems were attributable to course structure, which focused on rote memorization instead of the process of science through hands-on experiences and problem-driven critical thinking (Lindwall & Ivarsson, 2004; National Research Council, 1996). The problems identified during the turn of the century including unprepared graduates and a lack of STEM student retention worried industry leaders that the U.S. could potentially lose its foothold in an internationally emergent research-based commerce (Department of Education, 2000).

A decade later, the problems with STEM undergraduates remain (Johnstone, 2012; Labov, Reid & Yamamoto, 2010; Woodin, Carter & Fletcher, 2010), only now the stakes are higher. Global high-technology industries comprise a large percentage of the top producing workforce sectors, including communications equipment, semiconductors, pharmaceuticals, scientific instruments, and aerospace engineering (Singer, Nielsen & Schweingruber, 2012). Within biology, the same problems of inadequately trained undergraduates and decreased retention of potential future scientists exists (Goldey, Abercrombie, Ivy, Kusher, Moeller, Rayner, Smith & Spivey, 2012; Wood, 2009; Woodin et al., 2010). The report by the President's Council of Advisors on Science and Technology found that fewer than 40 percent of students who enter college intending to major in a STEM field complete college with a STEM degree (President's Council of Advisors on Science and Technology, 2012). Discipline-based education research (DBER) in science, particularly biology, indicates that undergraduate students are not provided with opportunities to practice technological skills. This research also finds students are in further need of experience with authentic, complex problems like those they would encounter in the workplace (American Association for the Advancement of Science, 2011; Hoskinson, Caballero, & Knight, 2013). In order to prepare biology undergraduates for the current workforce, colleges and universities need to further develop programs to incorporate modern industry techniques, tools and interdisciplinary methods (Labov et al., 2010; Woodin et al., 2010).

Technology continues to play an increasing role across all STEM disciplines, and undergraduate biology education should reflect that shift. Several universities and colleges across the nation have made efforts to incorporate more hands-on laboratory approaches and team-based exercises in their curricula (Goldstein & Flynn, 2011; Gross, 2004). Despite these changes, computer-based programs have played a limited, if not absent, role in undergraduate biology education reform. The scientific industry increasingly touts the use of technology in product development and research through news and scientific publications. The addition of modern, interdisciplinary methods to the duties required of many biological jobs suggests the importance of incorporating computer-based activities within undergraduate education curricula. In order to assess the effect of incorporating technology in biology programs, more research investigating relevant tools, programs and activities should be conducted.

Purpose

The purpose of this research was to evaluate undergraduate biology students' understanding of, perspective towards and aptitude with computer-based graphing. This study assessed the proficiency levels of each student after they were guided through graphing modules with both Microsoft Excel and SigmaPlot 12, and compared and contrasted student performance for each type of graphing task. It also determined if students were more likely to choose a professional computer-based program, such as SigmaPlot, in place of a spreadsheet-centric graphing program, such as Excel, when guided through activities with each type of software. I hypothesized that if students are presented with a professional computer program that provides them with a powerful means of graphing their ideas in an efficient manner, it would stand to reason that they would choose to use SigmaPlot over Excel. Students' preference for a computer-based tool that is used in biology careers would suggest they are able to easily understand and use this tool prior to induction to the workforce. Additionally, providing students with an easier method to graph data allows beginning students to learn graphing practices quickly and to focus on the graph interpretation rather than steps for creation.

Theoretical Perspective

Scientists exist within an interdependent community that exhibits social practices, has in-depth knowledge of specific scientific content, and employs unique linguistics practices. This type of social environment often necessitates that new members first participate peripherally in community practices before succeeding at immersion within the society (Lave & Wegner, 1991). The situated learning perspective is rooted in the theory of contextualized learning, which suggests that meaningful learning is only accomplished in the prevailing environmental or social context within which an individual exists. Enculturation into practice, including the use of tools, social dialect and authentic activities is crucial to providing an environment that immerses students in the procedures and methodologies encountered in a scientific setting. This learning perspective, derived from Lev Vygotsky's (1979) early thoughts of social and cultural influence on learning, was first clearly identified by Lave and Wegner in 1991, who suggested that student learning wasn't singularly accomplished in the learner's head, but as a function of a social process. Similarly, Brown, Collins and Duguid (1989) describe situated learning by introducing the concept of a "cognitive apprenticeship," in which students are enculturated into genuine practices similar to a craft apprenticeship. Due to the nature of the scientific community, using methods outlined by the situated learning perspective provides an effective means for educating undergraduate science students.

Tools play a pivotal role in situated learning, and consist of any culture-specific objects that are either physical (i.e. hammer to a carpenter) or mental (i.e. models or heuristics). As students regularly use a tool, they learn to both understand the function of that tool and develop an implicit understanding of the "world" in which it is used (Brown et al., 1989). However, tools within a "world" also evolve over time, making it critical for students to be aware of contemporary physical and mental devices used within their field of practice. Allan Collins (Jones, 1990) suggested that computer-based situated learning enlists many of the characteristics desired of a "cognitive apprenticeship." Examples of the characteristics employed by a "cognitive apprenticeship" include modeling and explaining (wherein a process is demonstrated to explain a concept), and exploration (wherein students are able to test different hypotheses and methods). Collins states that "the computer allows us to create environments that mimic situations in the real world that we cannot otherwise realize in a classroom (or home)." Undergraduate biology students may thus benefit extensively from regular use of modern and complex tools such as industry software. Through awareness of these current and relevant tools, students are better prepared to develop an identity as a member of a community, which is just as critical as becoming knowledgeably skillful in that community (Lave, 1993).

Literature Review

The inclusion of technology within STEM education was initially instituted in the late 1970's and successively gained traction within K-12, undergraduate and graduate curricula during the early to late 1990's (Nemirovsky, Tierney & Wright, 1998; Summerlin & Gardner, 1973; Thornton, 1987; Weller, 1996). The Microcomputer-Based Laboratory Project (MLP) was an early example of the success of technology integration within science education. Using a microcomputer to collect, store, process, graph and analyze data collected from sensors, MLP demonstrated a variety of concepts to science students in an engaging manner. Research established that the introduction of real-time graphing equipment into K-12 classrooms advanced students' graphing skills (Brasell, 1987; Jackson, Edwards, & Berger, 1993; Linn, Layman & Nachmias, 1987; Mokros & Tinker, 1987). Research conducted by Mokros & Tinker indicated that microcomputer-based labs are successful because "they provide a genuine scientific experience and [they] eliminate the drudgery of graph production." To date, this engaging effect of technology applies to many of the popularly used computer-based tools such as videos, 3-D graphical representations, animations and simulated labs (Smetana & Bell, 2012).

While many colleges and universities made strides towards reforming programs to better prepare students for workforce participation, these efforts seem to have "stalled" over the past decade, and have not effectively met the needs of technology-based science industries (Johnstone, 2012; Labov, Singer, George, Schweingruber & Hilton, 2009). However, among the continual propositions of a need for undergraduate science, engineering and math program reform, there are several success stories that exemplify the incorporation of technology and context specific tasks within curricula. Several motivated reform educators implemented learner-centered, inquiry-based techniques in their curricula, the results of which indicate successful learning among undergraduates (Derting & Ebert-May, 2010; Dori & Sasson, 2008; Goldstein & Flynn, 2011; Gross, 2004; Plass, Milne, Homer, Schwartz, Hayward, Jordan, Verkuilen, Florrie, Wang & Barrientos, 2012; Smetana & Bell, 2012). Successful curricular revisions incorporate critical thinking and collaborative work with tasks such as hands-on activities, case studies and computer-based visual representations and simulations. Drawing direction from the National Research Council's (1996) suggestion for increasing experiences in quantitative and analytical methods, several undergraduate programs adapted laboratory curricula to include more hands-on, professionally relevant technological skills. These efforts succeeded in providing students with biology specific experiences that allow them to understand authentic complex problems, better preparing them for workforce entry (Gross, 2000; Hoskinson et al., 2013; Labov et al., 2009; National Research Council, 1996; Shubert, Ceraj & Riley, 2009). The success of these endeavors provides inspiration and guidance for increasing the technological skills of biology undergraduate students. However, critical computer-based data analysis activities (such as graphing) using industry-based software are not commonly encountered in biological undergraduate program reform and should be discussed with respect to their significance in curriculum. Two aspects of this topic identified in the literature include graphing difficulties and documentation of previously used tools; these topics are discussed below as they provide further context for the objective of this research.

Graphing Difficulties

Biological data is often voluminous and complex in nature, and graphs are an effective tool used to facilitate data representation. Graphing allows scientists to effectively communicate trends and anomalies in their data to a larger audience. Although data manipulation is a key skill for biologists, it can be a challenging task even for a professional scientist. Research shows that both scientists and students can misinterpret graphs by incorrectly using salient graph features to answer a question rather than incorporating all potentially relevant data (Roth & Bowen, 2001). However, the problems of interpretation by professional scientists are exclusive to graphs that do not exist within the context of their specific discipline (Bowen, Roth & McGinn, 1999;

Kozma, 2003). These studies also show that professional experience (i.e. enculturation within the field) through real world scenarios assists with the scientists' ability to interpret graphs from their own field.

Students' difficulties with graph comprehension, production and interpretation are well documented across all grade levels by science education researchers (Åberg-Bengtsson & Ottsson, 2006; Bransford, Brown & Cocking, 2000; Friel, Curcio & Bright, 2001). Studies show that issues with fluency and analytical problems are major challenges for students. Fluency, or an individual's ability to accomplish a task quickly and with expertise, is often lacking in student graphing practices. Instilling graphing fluency in all biology students is critical for providing them the confidence to carry out problem-based tasks which they are likely to encounter in research or workplace settings (Airey & Linder, 2008; Carlson, 1999; Ebenezer, Kaya & Ebenezer, 2011). Analytically, students often see graphs as static images that represent a literal picture of a situation as opposed to an abstraction of quantitative data (Glazer, 2011). Perhaps as a result of this conceptualization of graphs, students tend to analyze a graph by focusing on extreme data points as opposed to incorporating all data in their interpretation. Additionally, students have difficulty identifying the independent and dependent variables, as well as the slope and y-intercept (Bowen et al., 1999; Hattikudur, Prather, Asquith, Alibali, Knuth & Nathan, 2012; Picone, Rhode, Hyatt & Parshall, 2007). Assigning meaning to the graphs, identifying trends, and interpreting interactions between the variables are additional analytical tasks that students find complicated (Kozma, 2003; Picone et al., 2007; Speth, Momsen, Moyerbrailean, Ebert-May, Long, Wyse & Linton, 2010).

In spite of the intent and rigorous coursework of many undergraduate biology courses, students often struggle with interpretation and comprehension of abstract graphs used in various scientific disciplines (e.g. population ecology). Studies show that while students are exposed to graphs through lecture, discussion and testing, they are often unable to develop proficient graph construction and interpretation skills for several reasons (Glazer, 2011; Guthrie, Weber & Kimmerly, 1993; Singer et al., 2012). The linguistic resources (i.e. familiarity with topical terminology and practice in discussing graphs) available to scientists are often unavailable to students due to their lack of authentic scientific experiences. This inexperience results in both an inability to understand terminology assigned to a graph (within a caption or axis label) and a failure to consistently describe and discuss graphs with peers or instructors (Bowen et al., 1999). Undergraduate students often tend to be unfamiliar with unique symbols or representations used in discipline specific graphs, and are unable to proficiently interpret symbol meaning (Aberg-Bengtsson & Ottosson, 2006; Kozma, 2003; Roth & Bowen, 2001). Additionally, accurate graph interpretation involves fluency, or experience in graphing data, which is often difficult to acquire without recitation (Bowen, et al., 1999). The combination of these problems elicits a challenge for some undergraduate biology students to obtain proficient graphing skills. Research suggests that undergraduate science programs could benefit from extensive technology-based exercises that provide skills required to solve real world problems, as is often done in industry training (Martin, 2006).

Graphing Tools

Computer-based software and technology applications are increasingly represented in science education as tools for improving statistical analysis and graphing skills (Aberg-Bengtsson, 2006; Barton, 1997; Bransford et al., 2000; Goldey et al., 2012; Hudnut, 2007; Jackson et al., 1993; Lindwall & Ivarsson, 2004). Graph construction is typically performed in one of three ways: by plotting values manually, through the use of spreadsheet-centric software such as Microsoft Excel, or through the use of statistics and graphical software. Of the computer-based methods used to teach graphing practices, spreadsheet-centric software is most commonly used in undergraduate science program curricula (Baker & Sugden, 2007; Serra & Godoy, 2011). The 1993 release of Microsoft Excel incorporated Visual Basic for Applications (VBA) and allowed users to write macros for repetitive tasks, which resulted in its widespread use internationally. However, there are limitations to the capabilities of spreadsheet-centric programs such as Microsoft Excel. Examples include Excel's inability to easily generate certain types of graphs (e.g. box-and-whisker plots require extensive steps), complicated software design that leads to operator errors, and the inability to construct graphs from two-variable equations (Abramovich, Nikitina & Romanenko, 2010; Baker, 2004; Lim, 2004). Abramovich et al. recommends that spreadsheets be used jointly with graphing specific software to facilitate the construction of complex graphs.

Several studies related to concept learning in physics, mathematics and computational biology have identified the benefits of using dynamic software such as GeoGebra (Hohenwarter, 2007) and the statistical program R (Eglen, 2009; Peterlin, 2010). Although the use of software for the provision of authentic experiences in data manipulation is becoming more common in the fields of physics and mathematics, the efficacy of graphical software with respect to providing undergraduate biology students with a meaningful skill has not been evaluated. One potentially promising program for use with biology students is SigmaPlot 12.0, a technical graphing and data analysis software package commonly used in the biology industry. This package offers the capacity to easily integrate spreadsheets from Excel while providing the user with a seamless interface that removes the "drudgery" of graphing complex datasets.

Experience with industry standard graphing tools may aide undergraduate biology students in practicing the skills required of them in a research environment. Additionally, incorporation of practice with these tools including an emphasis on their industry applications during introductory biology courses may help alleviate high STEM attrition rates. In response to the call for undergraduate biology program reform, Hoy (2004) suggested that commercial software that marries a dynamic visual approach with graphical capabilities should be regularly implemented in biology education. His sentiment reflects the notion that students are more apt to be engaged in scientific processes when a user-friendly, technologically advanced program is made available to them. The National Research Council (Singer et al., 2012) recognizes and suggests the need for science education research that assesses the effectiveness of technology-based educational techniques in undergraduate science education. Therefore, for this research project, I compared two computer-based methods for graphing—the spreadsheet-centric program Excel and the technical graphing software SigmaPlot 12.0—to explore students' perceptions of the impacts afforded by the different tools, as well as assess their proficiencies with each tool. My research questions are:

What level of proficiency do the students attain for graphing complex datasets when guided through a short-term graphing course using each software package: Microsoft Excel and SigmaPlot 12.0?

When graphing simple to complex datasets, which software package do sophomore and junior undergraduate biology students choose to use and why? How does their choice relate to their proficiency with each program?

Methodology

Software Used in Study

This exploratory study compared and contrasted undergraduate biology students' use of and preferences for two computer-based graphing programs. The two programs used in this study were Microsoft Excel and SigmaPlot 12.0. Microsoft Excel is a spreadsheet-based application that features functionality such as statistical calculations, graphing tools, pivot tables and macro programming (i.e. a saved sequence of commands that can be recalled). SigmaPlot 12.0 is a software package designed specifically for scientific graphing and data analysis. Data analysis features in SigmaPlot include basic statistics to advanced mathematical calculations, and its graphing capabilities include an interactive Graph Wizard, a wide range of 2-D and 3-D options for representation and modeling of data, and publication quality outputs. SigmaPlot is compatible with Microsoft Office, which allows Microsoft Excel spreadsheets to be easily imported. Although both programs contain graphing functionality, Excel is primarily designed as a spreadsheet and data manipulation tool, whereas SigmaPlot was primarily designed as a graphing and data analysis tool.

Research Design

To answer the research questions posed by this study, a mixed-model teaching experiment methodology was used (Figure 1). Teaching experiments emerged in mathematics education in 1970 as an effective method of understanding the learning and reasoning exhibited by students (Steffe & Thompson, 2000; Von Glasersfeld, 1991). In a teaching experiment, the researcher also acts as the teacher with an objective of understanding the concepts and operations held by students. For my research, the teaching experiment included a pre-intervention qualitative questionnaire, two teaching interventions focused on simple and complex graph creation in Excel and SigmaPlot, a graphing task session with quantitative evaluation of student steps for each task, and a subsequent qualitative interview. The quantitative method of assessing student capability with each program through evaluation of steps taken for Excel and SigmaPlot was intended to supplement information collected in the qualitative stage of analysis (including final graph products and interview responses). The interventions in this study were necessary to provide students with a baseline experience in graphing with the less commonly encountered software package (SigmaPlot), as none of the students had previously used this program. Collection and analysis of both qualitative and quantitative data provided a comprehensive understanding of student preference and proficiencies for the software programs Excel and SigmaPlot.

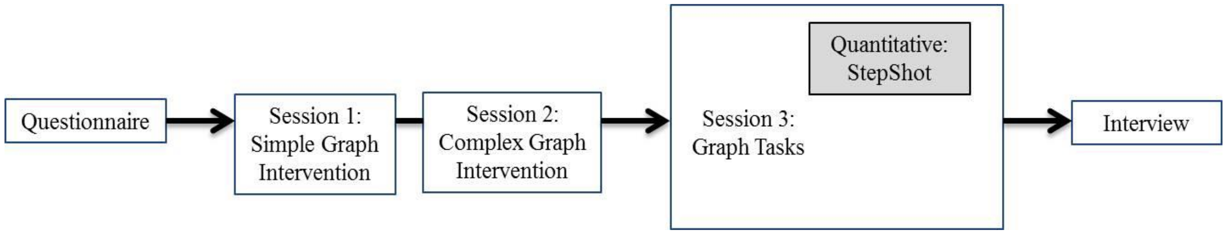

Figure 1. Teaching experiment model.

Setting and Participants

This research was conducted on the Point Loma Nazarene University campus in October 2013. It was conducted in accordance with Point Loma Nazarene University's Institutional Review Board's guidelines. Each participant was asked to provide their own laptop computer; several additional laptops were provided for those who did not have a compatible PC (required by SigmaPlot 12). Software trial versions were installed for SigmaPlot 12 and StepShot, the program used to record each students steps. Individuals were pre-selected using a questionnaire that evaluated their background experience with both programs, and their level of exposure to computers (Appendix A**)**. To accommodate student schedules and encourage participation for the three week duration of this study, each session occurred in the early evening (e.g. start time between 4 -5 p.m.), and student were provided with two dates and times for each session. Sessions were between 75 and 90 minutes in duration, and varied due to the amount of time required for students to complete the practice exercises and set of graphing tasks in the final session. All students were provided monetary compensation for their time at each session.

Sophomore and junior undergraduate biology students with self-identified moderate to high comfort levels for learning new software were recruited in compliance with IRB regulations for adult participants. Students who were comfortable with computers was desirable because they were less likely to be discouraged during the interventions, and more apt to participate in all three sessions. Sophomore and junior students were preferred because they have different levels of experience and familiarity with graphing, and with the program Excel. Second year students were expected to have less experience with Excel and graphs, whereas third year students were expected to be more confident with Excel and graphs. Ten biology program students were selected to participate in one of two teaching experiments that accommodated their schedule. The students were presented the same guided exercises in each session, resulting in two "Session 1: Basic Graphs" periods in the first week, two "Session 2: Complex Graphs" periods in the second week, and two "Session 3: Final Graph Tasks" periods in the third subsequent week. The participants included six Sophomore students and four Junior students, each of which participated in all three sessions.

Tutorial and Graphing Task Module Organization

Data used in this teaching experiment were derived from a 35-year data series of physical measurements obtained from the Rocky Mountain Biological Laboratory (www.rmbl.org). The datasets were formatted and condensed prior to conducting the tutorials in order to enable the students to focus on graphing instead of data manipulation. For graphing tasks in Session 3, data from the Rocky Mountain Biological Laboratory was combined with data generated based on research previously conducted on environmental change in the Colorado Rocky Mountain and long term morphological and abundance fluctuations in the yellow-bellied marmot (Ozgul, Childs, Oli, Armitage, Blumstein, Olson, Tuljapurka & Coulson, 2010).

After the instructor (E. Ferguson) demonstrated both simple and complex graphs using both programs, students were allotted time to try the graphing exercise on their own using similar data (Appendices B – D). In session 1, exercises were conducted in Excel and SigmaPlot 12 alternatively, whereas in session 2 all exercises were conducted first in Excel and subsequently in SigmaPlot 12 as the group preferred to do all activities in one program at a time. Session 1 included exercises for creating histograms and simple line graphs (one line graph using a singular dataset, and a second using two datasets of the same unit of measurement on the y-axis) (Appendix B, Tables B1 & B2). Session 2 included more complex graphing exercises such as box plots, scatter plots and line graphs with two units of measurement (and thus two y-axes)(Appendix C, Tables C1 & C2). During session 1 and session 2, students were allowed to request help or ask questions regarding how to generate graphs in both programs. In session 3, no guidance was provided on to how to generate graphs and all students were required to work independently using the program of their choice (Appendix D, Table D1-D3). Six graphing tasks were presented to students in session 3, each of which explicitly defined the type of graph that was to be completed (Appendix D, Table D1). An earlier pilot study demonstrated that when provided with the option for selecting the type of graph to use, students often selected basic, more familiar graphs. Controlling the type of graph to use in response to each task allowed students to focus on the program they prefer for the specific graphing technique and ensured that they attempted to create some of the more complex graphs (e.g. boxplots). Task 6 for Excel required some basic formatting of the data as did Task 3 for SigmaPlot; this was intentionally done to provide a "fair" scenario for use in either program.

Data Collection

Three types of data were collected during this study: qualitative interview responses, screenshots to use in quantitative coding analysis during the final session, and the final graph products produced in response to each of six tasks in session 3. To identify student program preference, at the end of the third session each student was asked which program they preferred to use for simple and complex graphing tasks, as well as weaknesses and strengths of each program (Appendix E). To obtain information about the process by which students progressed through the graph tasks, a screenshot recording program called StepShot (www.stepshot.net) was used for the duration of the third session. This software recorded screenshots or "steps" which provided important information about how proficient students were in creating each type of graph. Finally, data was collected in the form of graph products which were saved in each program upon completion of the third session.

Data Analysis

The data analysis for this study consisted of summarizing software preferences, qualitatively evaluating interview responses, coding screenshot "steps" used to quantitatively evaluate proficiency, and assessing final graph products. Program preference was calculated as a percentage of the group, as well as by undergraduate grade level. To qualitatively assess student preference, one-on-one interviews were conducted to further evaluate student likes and dislikes relating to each of the programs (Appendix E). Student responses to interview questions were recorded and compared to their Session 3 program choice. Interviews also provided the opportunity for other relevant factors to emerge during transcription and analysis.

Coding of StepShot data. Screenshots collected using StepShot were analyzed to determine students' software proficiency as they generated six different graphs during the third session (see Appendix D for the six graphing tasks). At the end of session 3, the StepShot file was saved and screenshots were exported to a Microsoft (MS) Word document; each step (or screenshot taken when the student clicked the mouse button) relating to the creation of a graph was assigned a code. Relevant steps included menu selection, graph type selection and settings selections, data selection, and completion steps (e.g. 'Finish' button). If a student re-did a step or series of steps related to the graph production, these were also included and identified as either 'reasonable' if they had made an error and were correcting it, 'correct' if they were starting over and taking the correct steps again, or 'unknown' if the objective behind the steps was unclear. Any other type of step unrelated to the creation of a graph, such as saving files, clicking on different windows, scrolling up and down in the spreadsheet etc., was excluded from coding. Extensive hesitations, wherein a student simply clicked the mouse button multiple times, were also excluded from coding. Students were not required to label axes in graphs nor provide a title for the graph as the objective of the task was to assess if the fundamental data manipulation was conducted correctly. Thus formatting steps were not included in the assessment. MS Word documents created from StepShot were generally very large (300-350 pages) and would repeatedly crash during coding, so were saved frequently. Sophomore 5's StepShot files were unable to be exported to MS Word due to the large file size, resulting in coding for only 9 (four Sophomore and four Junior) StepShot sessions. Table 1 provides a brief description of the codes used.

Table 1

Codes used to identify the steps taken by students in generating each graph.

Code	Description
C	Correct: This step is clearly correct in the process of the graphing task; since graphs may be created more than one way, it includes any correct step in the process; if a correct step was made after an error it was still considered correct.
R	Reasonable: This is a logical step that may not have been ideal, but is understandable given the graphing task; it is a step that is not efficient for creating the graph, but is understandable based on the nature of the software.
E	Error: This is an incorrect step. In this case there is no logical reason for this step; an error is also assigned if the user selected the wrong data to graph.
U	Unknown: This step is not identifiable as correct or incorrect, but is possibly due to the student performing a task or operation that is unrelated to the graph (e.g. pauses).

Results

Technology Issues Influencing Results

Several technological issues relating to computer performance and operating system compatibility occurred in the first two sessions, and were primarily corrected for the third session. However, major software compatibility problems were encountered during all three sessions related to operating requirements for SigmaPlot. The trial version of SigmaPlot is required to be run as an administrator, and due to limitations of recording as an administrator with StepShot, problems relating to the opening of SigmaPlot as well as final graph display issues were encountered. After session 1, I used an alternate Window's based recording program called "Problem Step Recorder" but could not continue to use it as provided insufficient data for analysis. Although I attempted to resolve the StepShot problem after session 2, students still encountered the SigmaPlot graph display, resulting in students either prematurely resigning their attempt at a graph using SigmaPlot or producing a corrupt final graph product. Despite errors in the final graph product of some students, StepShot tracked all steps students took to generate a graph in SigmaPlot, thereby providing information about their proficiency. Software problems are important to consider during interpretation of results.

General Findings

Subsequent to the final session, all students provided responses to interview questions about (a) their comfort level with graphing software, (b) their software preference for simple graphs, and (c) their software preference for complex graphs (Table 2). Students self-identified their comfort level with graphing software use on a scale of 1 to 10. Most students self-identified their level of comfort with using software as a 6 or 7. Amongst Junior biology students, 100% preferred Excel for simple graph tasks whereas 50% preferred Excel for complex graphs. Of the six sophomore students, 66% preferred Excel for simple graph tasks and 33% preferred SigmaPlot. For complex graphs, the reverse was observed with 66% selecting SigmaPlot and 33% selecting Excel.

Table 2

Students' self-identified comfort level with graphing software and software preference for simple and complex graphs.

Participant #	Grade Level & ID	Comfort Level Rating for Graphing Software (1-10)	Simple Graph Task Program Selection	Complex Graph Task Program Selection
1	Junior 1	8	Excel	Excel/SigmaPlot
2	Junior 2	7	Excel	SigmaPlot
3	Junior 3	7	Excel	Excel
5	Junior 4	6	Excel	Excel
4	Sophomore 1	5	Excel	Excel
6	Sophomore 2	10	Excel	Excel
7	Sophomore 3	6	SigmaPlot	SigmaPlot
8	Sophomore 4	7	Excel	SigmaPlot
9	Sophomore 5	6-7	SigmaPlot	SigmaPlot
10	Sophomore 6	6	Excel	SigmaPlot

Interviews consisted of a series of questions including student's course of study, software preference, weaknesses and strengths of each program, and if they would use the same software in their professional career. Students indicated that their preference for Excel was based on their familiarity with the program, its graphic layout and ease of data selection. Their familiarity with the program is the most commonly identified reason for selecting Excel. Yet, students also commonly indicated that SigmaPlot offered a more seamless and straightforward interface, and with additional experience they might select SigmaPlot. With respect to weaknesses for each software, students indicated that Excel is difficult to "figure out" or remember the numerous steps required to create complex graphs, while SigmaPlot requires the data to be formatted in a separate program (e.g. Excel) prior to being imported.

Several interesting responses were uncovered during the interviews that provide conflicting information regarding students' perception of software. In this example, Sophomore 2 is a pre-med Biology student who identified his comfort level as a 10 with graphing software and indicated that he strongly preferred Excel for simple and complex graphing tasks:

Instructor: ….so if you have pre-med data, like [data from] clinical trials, I'm assuming you would stick with the same program choice?

Sophomore 2: I would choose SigmaPlot

Instructor: Oh you would?

Sophomore 2: Well, I would try the SigmaPlot because I know that like our teachers talk about it a lot for our medical journals.

Instructor: Oh, ok. What specifically do they say?

Sophomore 2: Well, they just talk about how they use mass amounts of data to throw in, instead of like small ones like we are manipulating. So if it was me I would try Excel, but I know that I have to get myself more familiar with Sigma cause we're gonna start doing like journals and stuff. So I would choose Excel just for me.

Sophomore 2 clearly prefers Excel but has indicated in his response he would select SigmaPlot over Excel due to its use in professional settings for larger datasets and in medical publications. This provides evidence that students are sometimes aware of software needs within the scientific community, and recognize the need for learning these programs for future use in their careers.

In another example, Junior 1, a Bio-Chemistry major, selected Excel when asked about using graphing software in a professional career, but then revealed a preference for a more straightforward program:

Junior 1: I would probably stick with Excel. I actually just had a…it was for vertebrate physiology, but I had a lab and we had to do a statistical analysis, um, things I haven't done, and someone else showed me another program, that's just on the computers here that was super easy, so I actually just used that, instead of Excel.

Instructor: Oh cool, what's the name of it?

Junior 1: It's like IBM…I don't know, I forget what it's called. But it's on the computers in the lab. We had to do an ANOVA test? Do you know [what that is]? I've never heard of it before, but on that one, you just put [the data] in, and it just does it all for you. And with Excel, it's like, it would take hours.

Instructor: Did you try it in Excel and this other IBM program?

Junior 1: I had no idea where to start with Excel, and that, so I was just like trying to get help from my friend on this other one, and he just showed me it was real quick and easy.

In this example, Junior 1 points out that Excel can be complicated and time consuming when attempting to use some of its data analysis applications. Although this does not specifically relate to graphing data, it does indicate that this participant is aware that alternate software programs can sometimes perform better for certain applications. These responses indicate that biology students, while comfortable with Excel, experience limitations and complexities when using some of Excel's applications.

Although several students alluded to the benefits of SigmaPlot, Sophomore 3, a Bio-Chemistry Major, strongly indicated his preference for this software:

Instructor**:** What do you think are the strengths and weaknesses of SigmaPlot?

Sophomore 3: I'd say SigmaPlot, as far as strengths, uh, just all around [an] awesome program. That was my first time using it; I, I'm stoked on it, it's a great program. Very easy to use. Everything, as far as making graphs, it's all exactly the same, you just chose out of that little box and stuff, and having everything in one place is really convenient, so I would say that's a major strength. And a weakness? I would say formatting [the data], so other than that, I didn't really see a whole lot [of negatives].

Sophomore 3 clearly preferred SigmaPlot despite his previous undergraduate history using Excel.

The interviews provided a unique opportunity to obtain information regarding the students' perception of each software program. The verbal software preference indicated by each student can be compared to their interview responses to gauge their underlying opinions of each program. All students indicating a preference for Excel with either basic or complex graphing tasks (80% of the students) also provided inconsistent responses about their selection. They primarily indicated that given more time with SigmaPlot, they might prefer this software, particularly for more complex graphing tasks. These inconsistencies were represented in student responses about the strengths and weaknesses of each program, or in their responses to which program they would use in their own field. Only the responses of Sophomore 3 (who preferred SigmaPlot) solely indicated a strong preference for one program over the other; no student responded with consistent responses relating to their preference for Excel. Interview questions provided information regarding the student underlying perspectives about the use of these software programs.

Screenshot Coding

The results of the coding provided information about the students' program selection to perform (or attempt to perform) each graphing task, in addition to student proficiency with each of the graphing tasks. During session 3, five out of nine recorded sessions revealed that students attempted at least one graph in both programs, whereas the remaining four students only used Excel for the graphing tasks. Codes for each graph were tallied, and the number of correct steps and total steps were used to calculate the total percentage of correct steps taken. Percentage of correct steps is used as a proxy for evaluating the level of proficiency for each graphing task by software program; higher percentage scores indicate a higher level of proficiency while lower scores indicate low proficiency. The minimum, maximum and average percentages correct were calculated for each task for both Excel and SigmaPlot (Table 3). Across all tasks in Excel, students collectively had an average of 59.7% correct steps, whereas in SigmaPlot had an average of 67.9% correct steps. For Excel, students performed best on line graph Task 1 (69.6% correct steps) and Task 2 (65.6% correct steps), as well as the scatterplot using two datasets Task 5 (67.0% correct steps). Students did not perform as well in Excel with the single dataset scatterplot Task 4 (56.8% correct steps) and the histogram Task 6 (59%

correct steps). Students performed lowest on the box plots in Task 3 (40% correct steps). For SigmaPlot, the sample size is smaller than Excel because only five of the nine students selected SigmaPlot to attempt graphs, one of which only attempted to use SigmaPlot for one graphing task. For SigmaPlot, students performed the best on line graph Task 1 (78.7% correct steps) and the single dataset scatterplot Task 4 (86.7 % correct steps), although it should be noted that only one and two students contributed to these percentages respectively. Students performed relatively well for line graph Task 2 and the box plot Task 3 (63.5% correct steps for each task). The histogram Task 6 resulted in the lowest performance scores (47.1% correct steps).

Table 3

StepShot results by graph task.

	Excel			**SigmaPlot**		
	Minimum % Correct	**Average % Correct**	**Maximum % Correct**	**Minimum % Correct**	**Average % Correct**	**Maximum % Correct**
Task 1	33.3%	69.6%	100.0%	68.4%*	78.7%*	84.2%*
Task 2	33.3%	65.6%	96.7%	46.5%	63.5%	75.0%
Task 3	13.5%	40.5%	64.1%	25.7%	63.5%	89.7%
Task 4	20.6%	56.8%	95.5%	83.3%*	86.7%*	90.0%*
Task 5	6.8%	67.0%	95.7%	N/A	N/A	N/A
Task 6	0.0%	59.0%	96.9%	23.8%	47.1%	78.9%

sample size ≤2

Final Graph Products

Overall Results: The final graphs completed by students for each of the graphing tasks contributed to the results in several ways. The use of a recording program allowed for the final graph products to be identified even if it was not saved in the students' files. The evaluation of the final graphs for each task is based on its accuracy and completion and is represented as a percentage in Table 4. Many students attempted several graphing tasks in both Excel and SigmaPlot, so sample size and percentages are broken out by program as well as by task. Table 4 indicates whether a graph is “Correct,” “Mostly Correct” (only a few formatting errors), “Partially Correct” (approximately half of the elements of the graph are correct), “Incorrect” or if there is “No Graph” (wherein the final product was deleted and/or not recorded by StepShot). As previously mentioned, problems with the integration of the trial version of SigmaPlot with the recording program resulted in graph display errors, which often caused the students to revise the graph even when it should have been displayed correctly. For these instances, if a student made a series of steps that should have resulted in a correct graph, but encountered graph display errors, then these graphs were considered a “correct” final product (with a note that the graph “would have been correct” during review).

Table 4

Summary of state of "correctness" of final graph products organized by task, software and type of graph.

Software & Task	Type of Graph	Sample Size	Percentage Correct	Percentage Mostly Correct	Percentage Partially Correct	Percentage Incorrect	Percentage No Graph
Excel - Task 1	Line graph, single y-axis	9	100%				
SigmaPlot - Task 1	Line graph, single y-axis	3	100%				
Excel - Task 2	Line graph, two y-axes	9	44%		44%		11%
SigmaPlot - Task 2	Line graph, two y-axes	3	33%	67%			
Excel - Task 3	Boxplot	9	11%	33%	33%	22%	
SigmaPlot - Task 3	Boxplot	4		25%		25%	50%
Excel - Task 4	Single dataset scatterplot	9	44%	11%	11%	33%	
SigmaPlot - Task 4	Single dataset scatterplot	2*	100%*				
Excel - Task 5	Two datasets scatterplot	9	56%	22%		22%	
SigmaPlot - Task 5	Two datasets scatterplot	1*	100%*				
Excel - Task 6	Frequency distribution	7	71%			29%	
SigmaPlot - Task 6	Frequency distribution	3	33%			67%	

***** *indicates a low sample size***

The final completed graphs contributed information regarding the students' proficiency in creating simple and complex graphs. Overall, students performed well (100% Correct) for Task 1 in both Excel and SigmaPlot. They also performed well (100% correct) for Task 4 and 5 in SigmaPlot, but the sample size contributing to the performance on these two tasks for SigmaPlot was very low. For graphing Task 2, students performed slightly better using SigmaPlot than Excel (see Table 4). Students had difficulty with Task 3 in both programs and collectively performed worst on this task as compared to all other tasks. Of the two scatterplot tasks and only comparing for Excel (SigmaPlot users for this task were in the minority), students performed better with Task 5, which contained two datasets, than with Task 4, which only contained one dataset. Students performed better with Task 6 in Excel versus SigmaPlot. Although students were provided with instructions stating what type of graph to use and what datasets to use, students still encountered many problems with these graphing tasks. Tasks that students struggled with the most are identifiable in Table 4 as those that contained percentages in "Partially Correct," "Incorrect," or "No Graph" categories.

In the section below, I provide a few examples of individual student's challenges and successes with the graphing tasks.

Example 1: Inconsistencies Across Proficiency Level, Program Use and Program Selection. Although some students preferred Excel and also indicated a high level of comfort with graphing software in general, their final graph product in conjunction with their percentage correct scores per task indicated low levels of proficiency in Excel. For example, when asked in Task 4 to construct a singular scatterplot of YB (yellow bellied) Marmot Count and Maximum temperature, Junior 3 created two separate graphs. Figure 2 contrasts an example of a correct graph created by another student (Figure 2A) to the incorrect paired graphs Junior 3 created (Figure 2B & 2C).

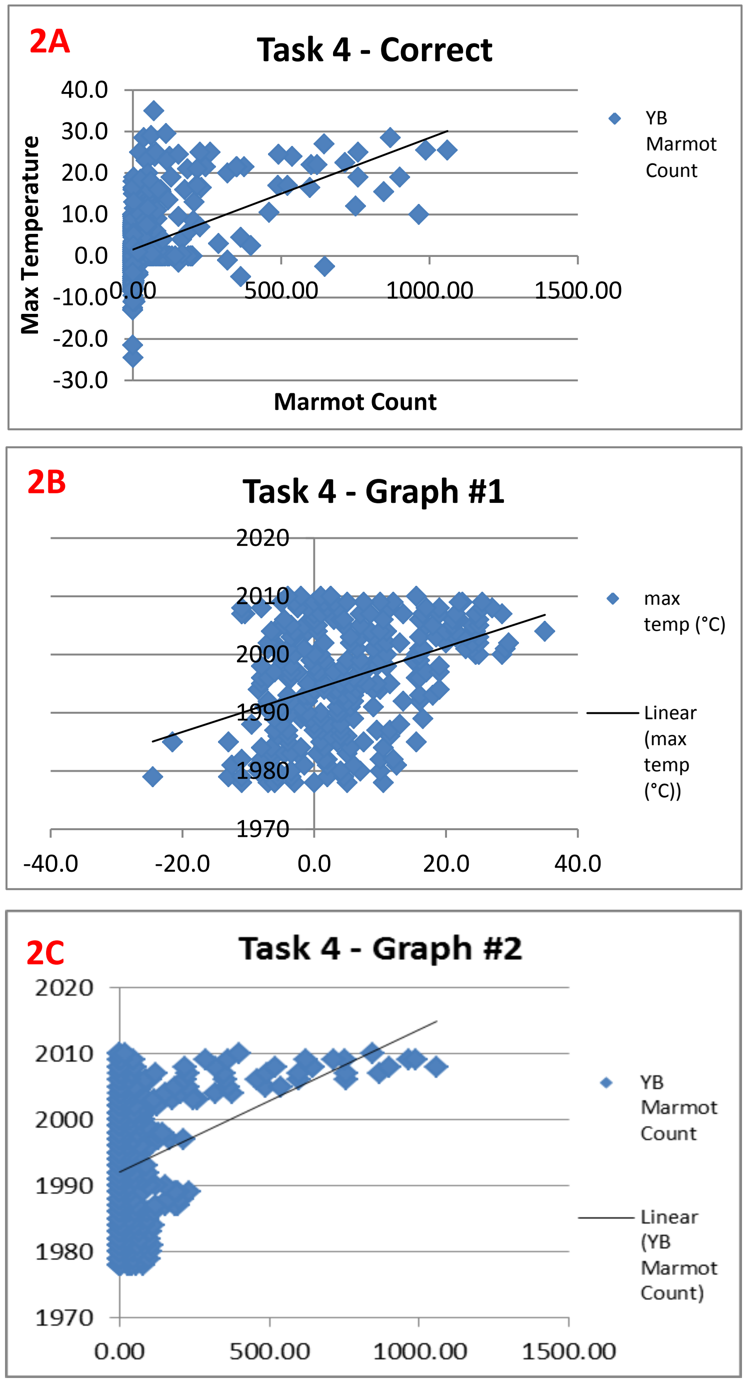

Figure 2. Comparison of correct graph (2A) versus Junior 3's pair of incorrect graphs (2B and 2C) for Task 4.

Junior 3 indicated a comfort level of 7 for using software programs for graphing purposes, and preferred Excel for both simple and complex graphing tasks. Table 5 provides a summary of this student's proficiency overall by task and program; Junior 3 seemed to struggle with graphing tasks 3, 4 and 5 based on the number of erroneous and repeated steps taken to create the graphs. In my notes during the analysis of Junior 3's graphing task, I wrote:

> *"There was a lot of back and forth between Excel & SigmaPlot –which resulted in an incorrect graph in excel; Junior 3 was very confused as to how to make this graph and repeatedly selected the wrong data to graph; Junior 3 returned to try this graph again after completing other graph tasks, and repeatedly made the same mistakes and was getting errors from choosing incompatible data"*

Interestingly, despite the indicated preference, Junior 3 attempted to use SigmaPlot for two of the six graphing tasks, which resulted in the correct creation of the frequency distribution graph, a task that many struggled with. My notes for her graphing proficiency with Task 6 in SigmaPlot indicated she knew exactly what steps to take for creating the graph in this program, and accomplished it with a minimal number of steps (Table 5). Even more interesting is that prior to graphing Task 6 in SigmaPlot, she also unsuccessfully attempted to create the frequency distribution in Excel (Figure 3).

Table 5

Screenshot coding and final graph results for Junior 3.

Software	Task	# of Correct Steps	# of Reasonable Steps	# of Erroneous Steps	# of Unidentified Steps	Total # of Steps	Percentage Correct Steps	Final Graph Results
SigmaPlot	Task 1	26	4	3	5	38	68.4%	Correct
Excel	Task 1	26	0	0	0	26	100.0%	Correct
Excel	Task 2	18	13	7	16	54	33.3%	No Graph
Excel	Task 3	75	66	57	125	323	23.2%	Incorrect
SigmaPlot	Task 3	18	27	6	19	70	25.7%	No Graph
Excel	Task 4	40	11	12	11	74	54.1%	Incorrect
Excel	Task 5	66	28	8	9	111	59.5%	Mostly Correct
Excel	Task 6	6	0	3	9	18	33.3%	Incorrect
SigmaPlot	Task 6	15	2	0	2	19	78.9%	Correct

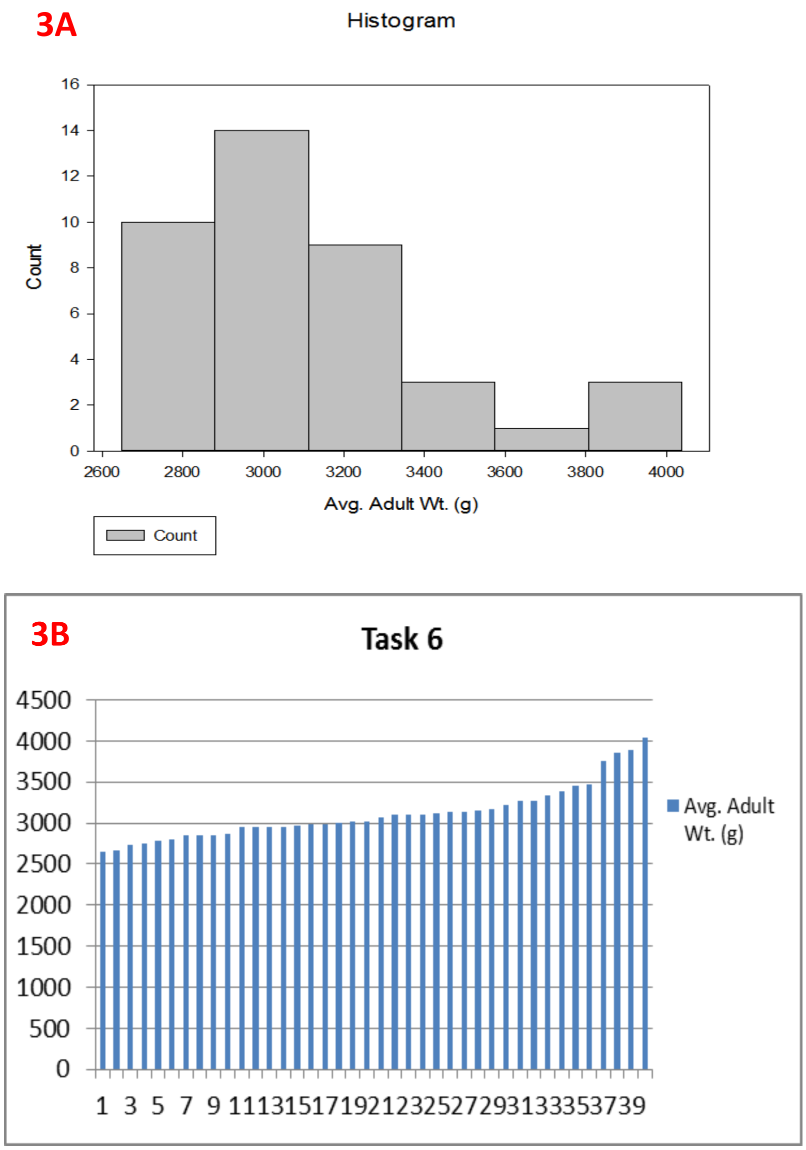

Figure 3. Junior 3's correct (3A) and incorrect (3B) final graph products for Task 6.

Example 2: Time on the X-axis. The results of this study identified several difficulties that students had with graphing tasks including difficulties with scatterplots, difficulties graphing data with two different y axes, and difficulties with boxplot and frequency distribution graphs. Several of these graphs were found to be difficult using either program, although the majority of these results were obtained from Excel-based graphing as that is the program most students selected (and used based on programs with administrator issues mentioned in the beginning). Despite clear instructions on which data to plot and the type of graph to use, student unexpectedly attempted to plot each dataset against time. Figure 4 exemplifies the problem Sophomore 5 and Sophomore 6 had with creating a scatterplot in Task 4. In this example, the students each decided to generate the incorrect graph shown in 4B as opposed to the correct scatterplot graph in 4A. Sophomore 6 had varying degrees of proficiency as indicated in Table 6 while Sophomore 5's proficiency scores via StepShot were unavailable due to an inability to extract the data. Interestingly, Sophomore 6's graphs for Task 5, which required the graphing of two datasets in a scatterplot, yielded a correct response (Figure 5).

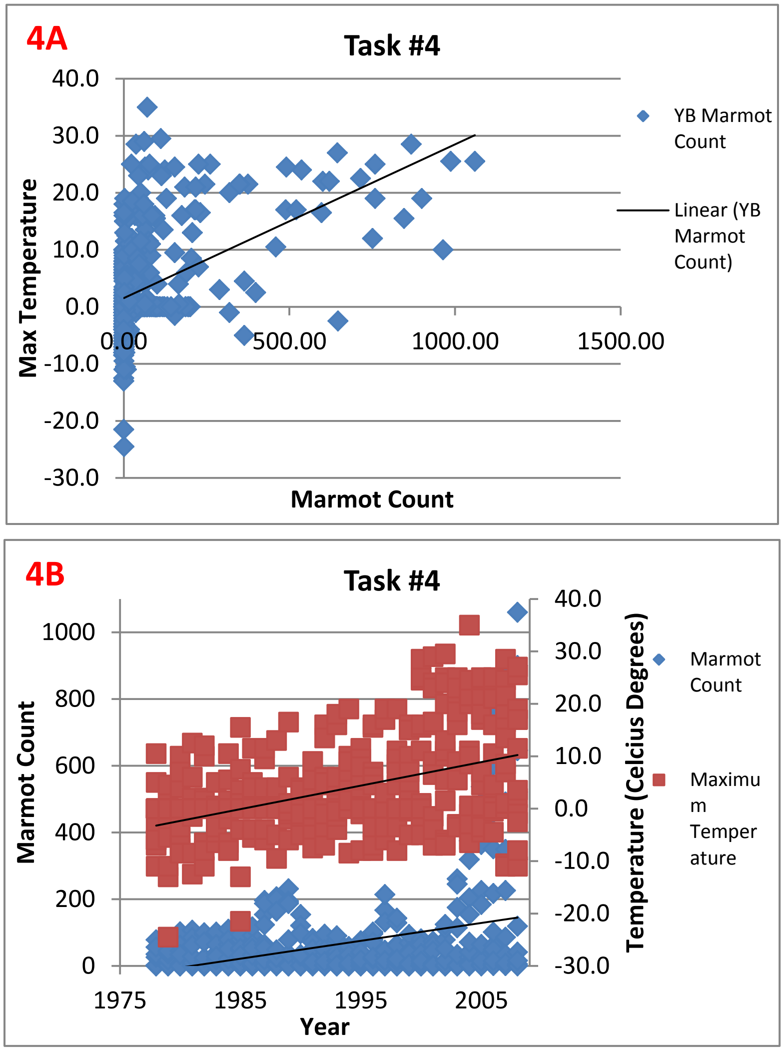

Figure 4. Comparison of correct graph (4A) compared with the incorrect graph (4B) created by Sophomores 5 and 6 for Task 4.

Table 6

Screenshot coding and final graph results for Sophomore 6.

Software	Task	# of Correct Steps	# of Reasonable Steps	# of Erroneous Steps	# of Unidentified Steps	Total # of Steps	Percentage Correct Steps	Final Graph Results
Excel	Task 1	27	3	3	6	39	69.2%	Correct
Excel	Task 2	29	2	3	9	43	67.4%	Correct
Excel	Task 3	56	11	6	19	92	60.9%	Partially Correct
Excel	Task 4	26	15	30	55	126	20.6%	Incorrect
Excel	Task 5	30	1	6	4	41	73.2%	Correct
Excel	Task 6	35	1	2	8	46	76.1%	Correct

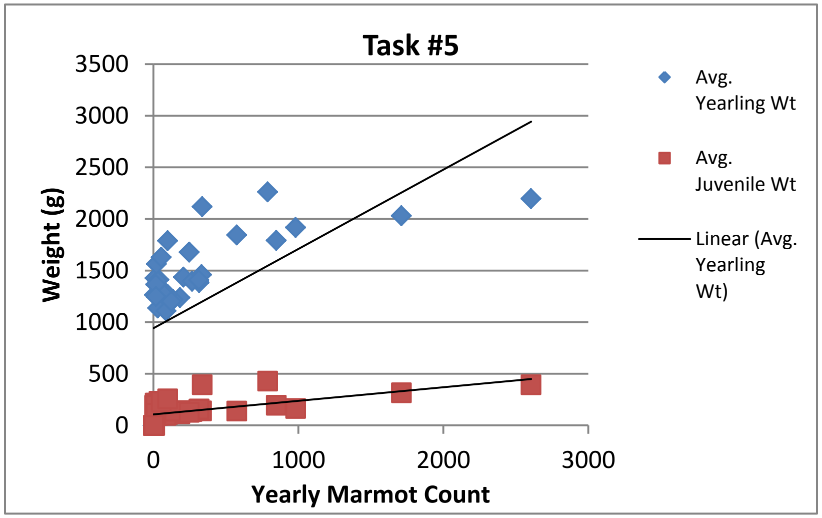

Figure 5. Sophomore 6's correct scatterplot in the two dataset task (Task 5).

Example 3: Two y-axes. Another problem students seemed to exhibit is graphing two datasets that contain different values of measurement (i.e. different y axes). Students regularly neglected to add the second y-axis in Excel in order to represent a different unit of measurement including Junior 1, Junior 2, Junior 3 (no final graph actually produced for this student in either program), Sophomore 2 and Sophomore 4 (Figure 6). Each student had varying levels of proficiency. Figure 6A depicts an accurate graph for Task 2 while Figure 6B represents the common graph for students listed above.

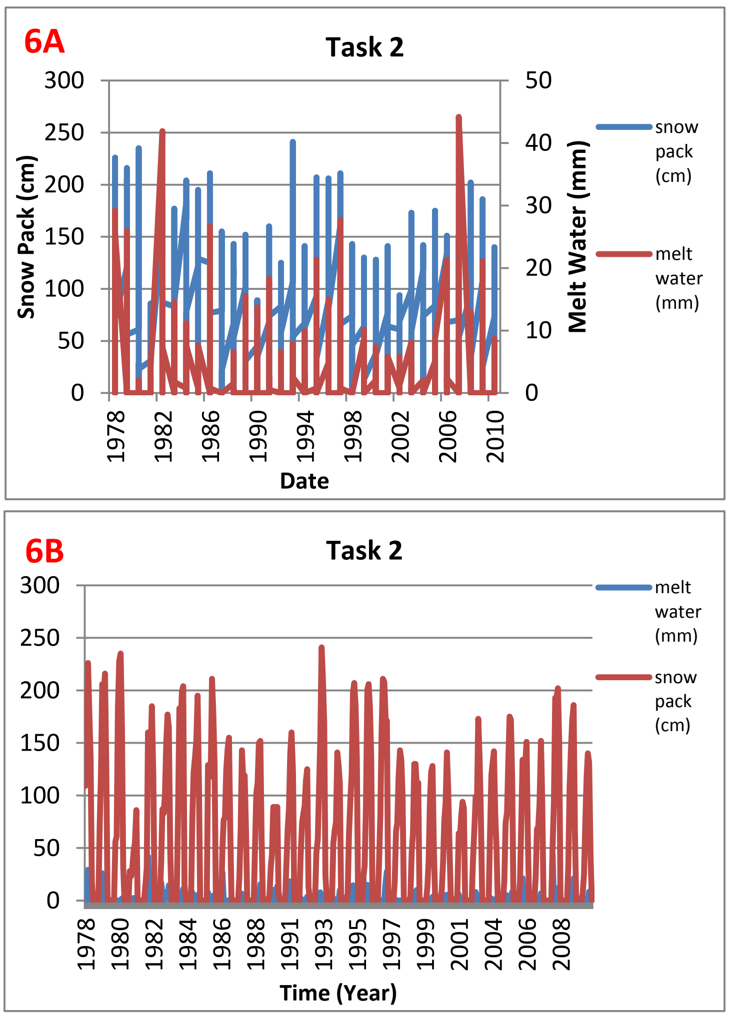

Figure 6. Comparison of a correct graph (6A) versus an incorrect graph (6B) created by five students for Task 2.

Example 4: Frequency Distributions. An additional final graphing problem existed for students attempting Task 6, the frequency distribution task. This particular graphing issue was mentioned in Junior 3's ability to create an accurate graph for Task 6 using SigmaPlot but not Excel. Several other students (using either software program) also ran into issues with this graphing task. In Figure 7, a correct graph is presented (Figure 7A), in comparison to an incorrect graph created in Excel by Sophomore 1 and an incorrect graph created in SigmaPlot by Sophomore 4 (Figure 7B & Figure 7C, respectively). Table 7 provides a summary of these two students' proficiency with graphing tasks; Sophomore 1 had the lowest scores of all 9 students and a high number of total steps, and Sophomore 4 did well with some Excel tasks.

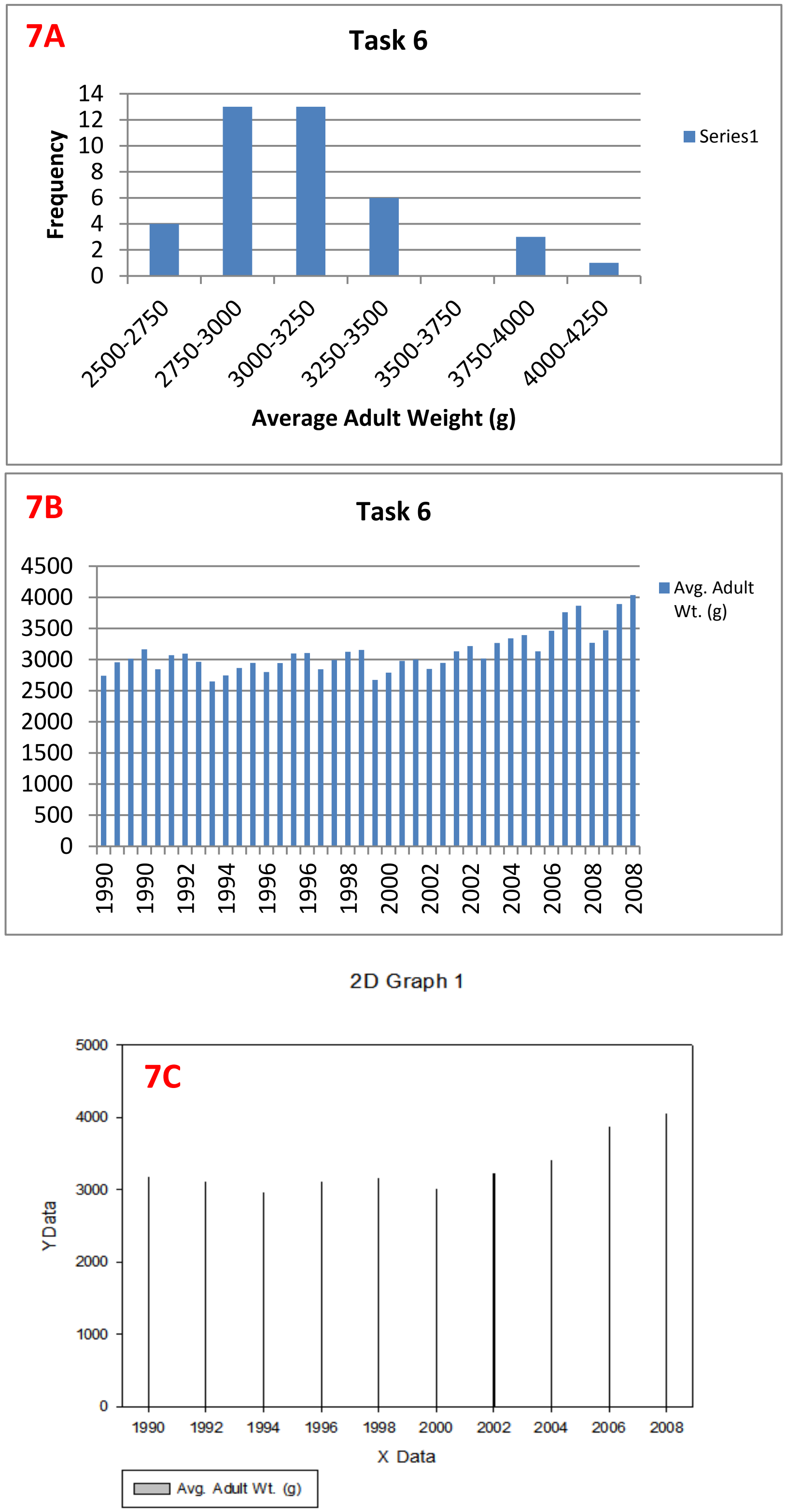

Figure 7. Comparison of correct graph (7A) as compared to an incorrect graph created in Excel (7B) and SigmaPlot (7C) for Task 6.

Table 7

Screenshot coding and final graph results for Sophomores 1 and 4 who generated the incorrect graphs in Figure 7.

Software	Task	# of Correct Steps	# of Reasonable Steps	# of Erroneous Steps	# of Unidentified Steps	Total # of Steps	Percentage Correct Steps	Final Graph Results
SOPHOMORE 1								
Excel	Task 1	42	17	15	25	99	42.4%	Correct
Excel	Task 2	53	16	22	67	158	33.5%	Correct
SigmaPlot	Task 2	20	17	5	1	43	46.5%	Mostly Correct
Excel	Task 3	18	30	35	50	133	13.5%	Partially Correct
Excel	Task 4	26	5	5	19	55	47.3%	Correct
Excel	Task 5	3	10	4	27	44	6.8%	Incorrect
Excel	Task 6	0	17	9	11	37	0.0%	Incorrect
SOPHOMORE 4								
Excel	Task 1	27	0	0	0	27	100.0%	Correct
Excel	Task 2	34	0	0	13	47	72.3%	Partially Correct
Excel	Task 3	33	8	42	46	129	25.6%	Incorrect
SigmaPlot	Task 3	260	0	5	25	290	89.7%	No Graph
Excel	Task 4	30	0	5	16	51	58.8%	Correct
Excel	Task 5	40	7	2	1	50	80.0%	Mostly Correct
SigmaPlot	Task 6	22	24	5	6	57	38.6%	Incorrect

However, despite the incorrect graph product for Task 6 (Figure 7C), I observed the following when analyzing Sophomore 4's production of this graph using SigmaPlot:

> *"Sophomore 4 had the right idea but did not do a histogram, instead did a bar chart so couldn't get the graph right. Sophomore 4 played around with this for a while after finishing but then gave up and saved it"*

Sophomore 4 made many of the correct steps for creating the graph for Task 6 but did not select the correct type of graph which prevented her from creating a correct final graph product. Sophomore 4 also tried to use SigmaPlot with Task 3 (BoxPlot) which she could not complete correctly in Excel. She made all the correct steps in SigmaPlot as well, but forgot that the data had to be formatted prior to graphing. Sophomore 4 thus acted consistently with her preference for SigmaPlot for more complex graphs.

Discussion

This research was intended to (a) assess the level of proficiency gained by students after being guided through graphing exercises in Excel and SigmaPlot, (b) determine the participants' preference of software programs, and (c) evaluate each student's level of proficiency in relation to their program selection. The results were partially inconclusive, but revealed some unexpected discrepancies and insightful indications of the students' true abilities with graphing data. The combined methods used for attaining information relating to the students' perceptions, capabilities and underlying viewpoints about these software programs was necessary for presenting a comprehensive response to each of my thesis questions. It was through the collective assessment of student program preference, recorded steps taken during graph creation in each program, and exploratory interview questions that my questions were addressed.

Through the course of this research, it became clear that my initial research question needed to be slightly refined. The training sessions were intended to provide the students with an opportunity to spend an equal amount of time performing graphing tasks with both Excel and SigmaPlot. However, due to the explicit difference in all students' experience with these two programs—all had experience with Excel and none had experience with SigmaPlot—the study results were affected by an unevenly weighted level of experience. All students started the study with a certain level of proficiency with Excel which enabled them to build upon the ability they already possessed, not develop a proficiency level from the exercises. I also expected Junior students to have a greater proficiency level with Excel than Sophomore students due to the extra year of instruction wherein they were likely to have spent more time with this program. Thus, in this discussion, the content tends to reflect how proficient students already were in Excel versus how proficient they become in SigmaPlot given the guided exercises, instead of how proficient they become with both programs due to the two intervention sessions.

Research Question A: Student Proficiency

Student's proficiency level was determined based on an analysis of the StepShot data, where both the percentage of correct steps taken by students and the state of their final graph products was examined. Although the "total number of steps taken" was used as a metric for evaluating the proficiency of the students for each graphing task, the total steps required for graphs in Excel versus SigmaPlot were inherently incomparable due to differences in the program functionality and infrastructure. SigmaPlot requires fewer steps for graph creation due to the underlying algorithms that allow for columns to be selected in comparison to cells of data in Excel. Therefore, the number of correct steps students performed in each program was converted to percentages to provide comparable results, and was one of the primary factors in evaluating proficiency level.

As indicated in Table 3, the minimum, average, and maximum percentage correct steps were calculated for each graphing task amongst all students. Percentage of correct steps was calculated by comparing the correct steps to the total number of steps (consisting of "correct", "incorrect", "reasonable", or "unknown" steps). A lower average percentage of correct steps indicated a greater number of incorrect, reasonable, or unknown steps in their repertoire of decisions regarding graph production. Although the average percentage of correct steps taken for graphs created in Excel and SigmaPlot were similar, students creating graphs in SigmaPlot performed slightly better on two of the six tasks. The minimum and maximum percentage of correct steps by students also provides information relating to proficiency because a lower minimum percentage of correct steps indicates students have a great amount of difficulty remembering the steps for creating a graph, whereas a high percentage indicates a clear understanding of the steps required within a program for completing a graph. Considering students' previous experience with Excel, it would stand to reason that they would have a lower minimum percentage of correct steps when attempting SigmaPlot graphs versus Excel graphs. However, for 5 out of 6 graphing tasks (the remaining task provided no SigmaPlot results for comparison as there was an error with the recording of the student who created this graph correctly using SigmaPlot), students attempting graphs in Excel had a lower minimum percentage of correct steps versus SigmaPlot. Additionally, they produced a maximum percentage of correct steps using SigmaPlot that was relatively comparable to Excel across all graphing tasks. These results suggest students were more efficient with SigmaPlot overall, which was not anticipated given their training and familiarity with Excel throughout their school careers (Figure 8). Although the lower sample size of students attempting graphs in SigmaPlot impacts the interpretation of these results, for those students who attempted graphs in SigmaPlot, this comparison suggests efficiency (and thus proficiency) may be slightly higher for students graphing in SigmaPlot versus Excel.

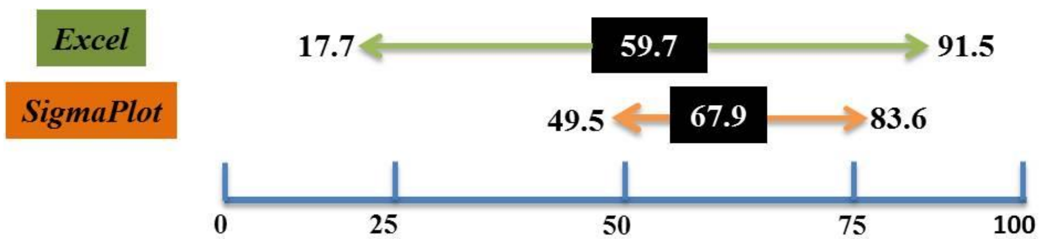

Figure 8. Range of percentage of correct steps from screenshot coding demonstrating a greater range of variability in students' performance using Excel versus SigmaPlot.

In reviewing the results from StepShot for all graphing tasks attempted using Excel, several details start to become apparent relating to students' proficiency level with this new program. As mentioned, the minimum and maximum percentage of correct steps for Excel showed a wide range of variability in the abilities of students using this program. In looking at the group as a whole, if students had an overall higher level of proficiency in Excel, I would expect to see less variability amongst students (e.g. minimum and maximum percentage of correct steps values would be expected to be closer to the average percentage of correct steps values for each task). This is expected because I assumed that students knew the steps required to create a graph in Excel, and thus should have a larger number of correct steps. Additionally, given the limited experience students had with SigmaPlot, I expected that the average percentage of correct steps in Excel would be better than the percentage of correct steps using SigmaPlot. I also expected to see a slightly higher average percentage of correct steps values for the more basic graphs such as the line graphs in Task 1 and Task 2. However, the results show that students do not have an equal distribution of proficiency for the creation of simple and complex graphs in Excel, and that their overall proficiency level is less than expected given their experience with Excel prior to this teaching experiment.

The final graph products provide another clue relating to the proficiency level of students for Excel and SigmaPlot. As indicated in Table 4, the final graphs produced by all students were evaluated to determine if they were correct, mostly correct, partially correct, incorrect or missing. These results enhanced the findings in Table 3 regarding the percentage of correct steps taken to generate a graph. I expected that despite the percentage of correct steps taken in Excel, students would ultimately produce a correct, completed graph; this was not typically the case. To further assess proficiency, I compared several of the graphing task results from Table 4 with the results in Table 3. For example, Task 1 resulted in 100% correct graph product for students using either SigmaPlot or Excel, but this is accompanied by differences in the average percentage of correct steps as students exhibited 69.6% average correct steps in Excel and 78.7% correct steps in SigmaPlot. The lower percentages of correct steps in Excel indicates that students had a higher quantity of incorrect, reasonable, or unknown steps in comparison with those students producing the graph in SigmaPlot. Table 4 further suggests that many of the graphs attempted by students, despite the trial and error relating to generating a graph, did not result in correct graphs. This data reveals that students' proficiency with Excel is less than expected considering their preference and experience with Excel-based graphing. Research in graphing capabilities indicates that students have difficulty developing proficient graph construction skills (Glazer, 2011; Singer et al., 2012). Fluency, or accomplishing a task quickly and with expertise, is also noted as lacking in undergraduate biology students despite its importance in the workplace (Airey & Linder, 2008; Ebenezer et al., 2011). My research indicates a lack of fluency among most students for graphing in Excel, however, there's also evidence of a competitive proficiency level in SigmaPlot even given a short experience with the program.

Research Question B: Student Program Preference

Regarding program preference, Excel was selected as the software program choice amongst many of the students, particularly with simple graphing tasks; however interview responses ultimately indicated conflicting viewpoints. One clearly evident trend was the stronger preference for Excel amongst Junior students as compared to Sophomore students. Junior students overwhelmingly indicated Excel as their preferred graphing software program of choice for both simple and complex graphing tasks. Sophomore students were almost equally split in preferences between Excel and SigmaPlot for simple graphing tasks, but predominantly preferred SigmaPlot for complex graphing tasks. This preference may be a result of Junior level students being privy to an additional year of course instruction wherein Excel is the only graphing software program used. Additionally, many students provided responses to interview questions that conflicted and were inconsistent with their program choice. As mentioned in the results, these inconsistencies were represented in the majority of the participants (80%).

Research Question C: Student's level of proficiency in relation to their program selection.

I also found inconsistency in software program choice and proficiency, which is best exemplified in Junior 3. This student indicated a strong preference for Excel, but still decided to use SigmaPlot (in addition to Excel) for three of the six graphing tasks. Her percentage of correct steps for all tasks aside from Task 1 in Excel and SigmaPlot and Task 6 in SigmaPlot were very low. In comparing Tasks 1 and 6 in each program, she completed correct graphs in both programs for Task 1, but only completed the correct graph in SigmaPlot for Task 6 (Excel for Task 6 was incorrect). This suggests that her proficiency with SigmaPlot was slightly greater than for Excel, despite her clear choice of Excel for program preference. Considering familiarity with Excel is one of the reigning lines of justification for the selection of this program, students' preference for Excel is not unexpected. Perhaps what is unexpected is the Junior and Sophomore students' level of proficiency of graphing within Excel as indicated in the previous section. I expected that students would have a higher level of proficiency with the software program they were more familiar with, but this was not the case.

When reviewing the participant responses about these software programs we can glean several overall impressions of both SigmaPlot and Excel. SigmaPlot was perceived as offering a more seamless and intuitive interface for graphing purposes, but was predominantly unfamiliar, and gave the impression that there was more formatting of data required prior to graphing. Excel provided students with a familiar platform containing attractive layout features and the ability to manually manipulate and select data, but was found to be burdensome with respect to the steps required for graphing that data. Student familiarity with Excel should have enabled them, particularly the more experienced students, to be relatively proficient with the graphing in Excel. But as this study shows, there is clearly less of an aptitude amongst most students for this program. Despite the endeavors of all students to complete each of the graphing tasks in Excel in an efficient manner, the results indicate (based on the percentage of correct steps and state of final graph products) that they collectively struggled to attain this objective. Their interview responses can be used as an indication of a common problem; one of Excel's biggest weaknesses is "figuring out" the steps to create the graph, as many students informed me. This suggests that Excel is heavily burdensome when used for tasks other than data manipulation. However, we must keep in mind that Excel is first and foremost a spreadsheet-centric program designed for storing, organizing and manipulating data. Its capabilities for these purposes have revolutionized many finance and scientific industries, and ancillary capabilities such as graphing and statistical analysis are naturally valuable features for any user. The problem arises from the origin – Excel is predominantly designed for data storage and though subsequent version modifications and developments have increased the capabilities of this software program, the steps required to obtain analytical and graph results remain complicated and cumbersome. These challenges are not newly identified as various researchers have indicated the deficiencies of Excel for graphing data (Abramovich et al., 2010; Baker, 2004; Lim, 2004).

Other Interesting Findings

During the course of this study, several interesting results emerged that were not directly related to my research questions. The final graph products provided information about graphing errors encountered by this group of students. Students experienced several problems regarding the appropriate physical display of data within two dimensional graphs as well as how datasets can be compared to one another. For instance in Example 1, Junior 3 was asked to create a scatterplot using two datasets and instead graphed each dataset separately, each with respect to time (year) for Task 4. Sophomore 5 and 6 (Example 2) also chose to graph two datasets with respect to time (year) for Task 4, but this time within one graph. In a subsequent scatterplot graph, Sophomore 6 accurately graphed a second scatterplot using both datasets for Task 5. The propensity to graph data along a timescale was also exhibited in Task 6 for Junior 3 using Excel, for Sophomore 1 using Excel and for Sophomore 4 using SigmaPlot. These graphical errors represent a tendency for students to graph data with respect to time, despite the nature of the requested graph. These results align with previous research (Bowen et al., 1999), and reveal a fundamental issue that students have with graphing data. This penchant to graph time on the x-axis could be due to the frequency of time oriented graphs occurring in news reports and other media.

Several complex graph tasks prompted unexpected errors in students' products. The Task 6 examples in my results demonstrated students’ difficulty in conceptualizing a frequency distribution. For this task, students incorrectly graphed all data points along the x-axis, which they did not seem to perceive as problematic because these incorrect graphs were kept as their final products. Another problem students encountered was in graphing data with two different types of measurement using the same y-axis in Task 2. Junior 1, Junior 2, Junior 3, Sophomore 2 and Sophomore 4 all made this same error. It is possible that units of measure in the same graph are difficult to perceive conceptually. Errors were exhibited from both Junior and Sophomore students, so this error may not be related to experience level. These students all varied with respect to proficiency levels, which could indicate that their conceptions are based on an overarching perspective on how data is represented graphically.

Conclusion

Implications for Undergraduate Biology Courses

The results from this study suggest that students are in need of more extensive training with Excel throughout their undergraduate biology careers to enable them to develop fluency with this skill. Furthermore, students may benefit from being exposed to different graphing programs, such as SigmaPlot, so that they may focus their efforts on graph interpretation rather than construction. The observed results for SigmaPlot in this study suggest that use of this software program for graphing exercises may be a beneficial addition to undergraduate biology courses. Responses from the student with the highest proficiency levels (Sophomore 2) and conviction in his preference for Excel revealed that SigmaPlot is still deemed a necessary component for growth in a career following graduation. This suggests that students may be aware of more powerful graphing programs that are available and ultimately necessary to learn as they progress towards their career. Junior 1 also provided an insightful perspective on the overall complexity of performing some analytical tasks in Excel. His description of an alternate program used in place of Excel for conducting an ANOVA statistical analysis provides a related example of how Excel can be complicated and burdensome. His comment specifically indicates that Excel would take too much time to learn the steps versus an alternate program; this suggests that students may be bogged down by the complexities of obtaining a result in Excel, such that it prevents them from focusing on the analysis results. Although this example does not refer to graphing, it does contribute to the common perspective that Excel can be difficult to navigate and execute for some tasks. Computer-based programs and technology that streamline the process of data manipulation and analysis should be incorporated into undergraduate biology programs such that students are provided with tools that enhance their learning of concepts and allow for better retention of graphing and analysis processes.

Limitations

Several limitations were encountered during this exploratory research that should be considered when interpreting these results. Initially and importantly, technological issues were encountered that impacted the use of SigmaPlot in this study. Operating system compatibility was and will remain a problem with SigmaPlot; at the time of this study, SigmaPlot was only available for Windows operating systems. This compatibility problem is important to consider if future use of the software in courses or research is considered. Additionally, administrator privileges required for the use of the trial version of SigmaPlot in conjunction with the StepShot recording program resulted in problems with SigmaPlot operation and display properties. This ultimately resulted in some students opting out of the use of SigmaPlot for their session 3 graphing tasks (as observed through initial StepShot step recordings). Recordings made during sessions 1 and 2 allowed for some problems to be counteracted, but did not result in a complete resolution to these problems. Careful consideration of these issues should be made prior to conducting future research with StepShot or the trial version of SigmaPlot (the licensed version is recommended as it resulted in no errors).

Sample size was another problem encountered in this study and should be noted when interpreting results. Ten students attended all three sessions, and of these ten students, six were Sophomores and four were Juniors in various undergraduate biology programs at Point Loma Nazarene University. Although a larger sample size of students, even distribution of grade levels and sampling from various universities and colleges is typically desired, this was an exploratory study and thus limited by the nature of the data collected. These results are considered qualitative and are intended to provide information to support future, extensive research. In addition to the smaller sample size, only four students selected SigmaPlot for use in their session 3 graphing tasks. This is directly relates to the problems described above, and should be noted when interpreting results from this study.

Future Research

Future research is required to provide a more comprehensive assessment of student proficiency using SigmaPlot, Excel, and various other graphing programs. As previously mentioned, the sample size of students using SigmaPlot was smaller than desired, and technology issues complicated results further. This exploratory study is intended to provide the basis for the evident need for more research in the area of graphing program use and proficiency. Although SigmaPlot was used in this study, testing and research of alternate graphing software programs is recommended to obtain data that identifies the best performing program.

Though Excel is a widely used software program that has infiltrated the core of most undergraduate biology programs within the United States, this study provides preliminary evidence that it falls short of enabling students to graph data correctly and retain the knowledge required to reproduce graphs. If students exhaust their effort on repeatedly re-learning the steps required to reproduce a graph or analysis task within a program, they are less likely to refine or invest energy in interpreting the results of their product. This ultimately leads to a problem that is two-fold for the work force: students are unable to successfully retain the capabilities for tasks required of their post-baccalaureate job, and are less experienced with developing an accurate representation of their results (Johnstone, 2012; Labov et al., 2010; Woodin et al., 2010). By providing students with an alternative, less burdensome method for graphing data during the undergraduate stage of education, they are both able to focus their efforts on correct interpretation of biological data, and able to retain information about the correct type of graph to use as opposed to how to create a specific graph.

Use of more intuitive software programs such as SigmaPlot found in the scientific industry setting can support this objective and expose students to appropriate tools for use in their careers. Ultimately this goal relates to the need of science and technology industries, and the need for the United States to be competitive in an internationally growing market (National Science Board, 2012). Providing students with the best tools available to analyze scientific data not only enables them to excel in post-graduate work, but also provides a feasible, application-based experience for beginning students that might influence them to continue in science undergraduate programs of study. Technology has permeated many advanced scientific disciplines through the use of computer-based software. To better prepare students in undergraduate biology programs with the instruments required of them in their careers, incorporation of these instruments within an academic environment is essential.

References

Åberg-Bengtsson, L. (2006). "Then you can take half… almost": Elementary students learning bar graphs and pie charts in a computer-based context (7-12 years). *The Journal of Mathematical Behavior*, *25*(2), 116-135.

Åberg-Bengtsson, L., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching*, *43*(1), 43-62.

Abramovich, S., Nikitina, G. V., & Romanenko, V. N. (2010). Spreadsheets and the development of skills in the STEM disciplines. *Spreadsheets in Education (eJSiE)*, *3*(3). Retrieved from http://epublications.bond.edu.au/ejsie/vol3/iss3/5

Airey, J., & Linder, C. (2008). A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, *46*(1), 27-49.

American Association for the Advancement of Science (2011). Vision and Change: A Call to Action, Final report, Washington, DC.

Baker, J. (2004). The charts that Excel cannot do. *Spreadsheets in Education (eJSiE), 1*(3). Retrieved from http://epublications.bond.edu.au/ejsie/vol1/iss3/6

Baker, J., & Sugden, S. J. (2007). Spreadsheets in education–The first 25 years. *Spreadsheets in Education (eJSiE)*, *1*(1). Retrieved from http://epublications.bond.edu.au/ejsie/vol1/iss1/2

Barton, R. (1997). Computer-aided graphing: A comparative study. *Journal of Information Technology for Teacher Education*, *6*(1), 59-72.

Bowen, G. M., Roth, W. M., & McGinn, M. K. (1999). Interpretations of graphs by university biology students and practicing scientists: Toward a social practice view of scientific representations practices. *Journal of Research in Science Teaching*, *36*(9), 1020-1043.

Bransford, J. D., Brown, A. L., & Cocking, R.R. (2000). *How people learn*. Washington, DC: National Academy Press.

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*, *24*(4), 385-395.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32-42.

Carlson, M. P. (1999). The mathematical behavior of six successful mathematics graduate students: Influences leading to mathematical success. *Educational Studies in Mathematics*, *40*(3), 237-258.

Department of Education. (2000). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st century*. Washington, DC: ERIC Clearinghouse.

Derting, T. L., & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology: Positive relationships with long-term student achievement. *CBE-Life Sciences Education*, *9*(4), 462-472.

Dori, Y. J., & Sasson, I. (2008). Chemical understanding and graphing skills in an honors case-based computerized chemistry laboratory environment: The value of bidirectional visual and textual representations. *Journal of Research in Science Teaching*, *45*(2), 219-250.

Ebenezer, J., Kaya, O. N., & Ebenezer, D. L. (2011). Engaging students in environmental research projects: Perceptions of fluency with innovative technologies and levels of scientific inquiry abilities. *Journal of Research in Science Teaching*, *48*(1), 94-116.

Eglen, S. J. (2009). A quick guide to teaching R programming to computational biology students. *PLoS Computational Biology, 5*(8), 1-4.

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*, *32*(2) 124-158.

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education, 47*(2), 183-210.

Goldey, E. S., Abercrombie, C. L., Ivy, T. M., Kusher, D. I., Moeller, J. F., Rayner, D. A, Smith, C.F. & Spivey, N. W. (2012). Biological inquiry: A new course and assessment plan in response to the call to transform undergraduate biology. *CBE-Life Sciences Education*, *11*(4), 353-363.

Goldstein, J., & Flynn, D. F. (2011). Integrating active learning & quantitative skills into undergraduate introductory biology curricula. *The American Biology Teacher, 73*(8), 454-461.

Gross, L. J. (2000). Education for a biocomplex future. *Science*, *288*(5467), 807-807.

Gross, L. J. (2004). Points of view: the interface of mathematics and biology interdisciplinarity and the undergraduate biology curriculum: finding a balance. *Cell Biology Education*, *3*(2), 85-87.

Guthrie, J. T., Weber, S., & Kimmerly, N. (1993). Searching documents: Cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary Educational Psychology*, *18*(2), 186-221.

Hattikudur, S., Prather, R. W., Asquith, P., Alibali, M. W., Knuth, E. J., & Nathan, M. (2012). Constructing graphical representations: Middle schoolers' intuitions and developing knowledge about slope and y-intercept. *School Science and Mathematics*, *112*(4), 230-240.

Hohenwarter, M. (2007). Dynamic mathematics with GeoGebra. *Journal of Online Mathematics and Its Applications*, *7*. Retrieved from http://www.maa.org/external_archive/joma/Volume7/Hohenwarter/index.html

Hoskinson, A. M., Caballero, M. D., & Knight, J. K. (2013). How can we improve problem solving in undergraduate biology? Applying lessons from 30 years of physics education research. *CBE-Life Sciences Education*, *12*(2), 153-161. Hoy, R. (2004). Points of view: The interface of mathematics and biology new math for biology is the old new math. *Cell Biology Education*, *3*(2), 90-92.

Hudnutt, B. S. (2007). *Teaching functions with dynamic graphing tools: A study of lesson plans.* (Unpublished master's thesis). North Carolina State University, Raleigh, N.C.

Jackson, D. F., Edwards, B. J., & Berger, C. F. (1993). Teaching the design and interpretation of graphs through computer-aided graphical data analysis. *Journal of Research in Science Teaching*, *30*(5), 483-501.

Johnstone, T. (2012). *California's need for engineers and STEM education* (Unpublished masters dissertation). California State University, Sacramento, CA. Jones, B. F. (1990). *Dimensions of thinking and cognitive instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates. .

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, *13*(2), 205-226.

Labov, J. B., Singer, S. R., George, M. D., Schweingruber, H. A., & Hilton, M. L. (2009). Effective practices in undergraduate STEM education part 1: Examining the evidence. *CBE-Life Sciences Education*, *8*(3), 157-161.

Labov, J. B., Reid, A. H., & Yamamoto, K. R. (2010). Integrated biology and undergraduate science education: A new biology education for the twenty-first century?. *CBE-Life Sciences Education*, *9*(1), 10-16.

Lave, J. (1993). Situated learning in communities of practice. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 63–82). Washington, DC: American Psychological Association.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, U.K.: Cambridge University Press.

Lim, K. (2004). A survey of first-year university students' ability to use spreadsheets. *Spreadsheets in Education*, *1*(2), 52-66.

Lindwall, O., & Ivarsson, J. (2004). What makes the subject matter matter? Contrasting probeware with Graphs & Tracks. In J. Ivarsson (Ed.), *Renderings & reasoning: Studying artifacts in human knowing* (pp. 115-143). G¨oteborg, Sweden: Acta Universitatis Gothoburgensis.

Linn, M. C., Layman, J. W., & Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Contemporary Educational Psychology*, *12*(3), 244-253.

Martin, P. (2006). Achieving success in industrial training. Proceedings of ICOTS-7: *International Association for Statistical Education.* Salvador, Bahia, Brazil.

Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*(4), 369-383.

National Research Council. (1996). *From analysis to action: Undergraduate education in science, mathematics, engineering, and technology.* Washington, DC: National Academy Press.

National Research Council. (2003). *BIO 2010: Transforming undergraduate education for future research biologists*. Washington, DC: National Academy Press.

National Science Board. (2012). *Science and Engineering Indicators*. Washington, DC: National Science Foundation.

Nemirovsky, R., Tierney, C., & Wright, T. (1998). Body motion and graphing. *Cognition and Instruction*, *16*(2), 119-172.

Ozgul, A., Childs, D. Z., Oli, M. K., Armitage, K. B., Blumstein, D. T., Olson, L. E., Tuljapurka, S. & Coulson, T. (2010). Coupled dynamics of body mass and population growth in response to environmental change. *Nature*, *466*(7305), 482-485.

Peterlin, P. (2010). Data analysis and graphing in an introductory physics laboratory: spreadsheet versus statistics suite. *European Journal of Physics, 31*(4), doi:10.1088/0143-0807/31/4/021.

Plass, J. L., Milne, C., Homer, B. D., Schwartz, R. N., Hayward, E. O., Jordan, T., Verkuilen, J., Florrie, N., Wang, Y., & Barrientos, J. (2012). Investigating the effectiveness of computer simulations for chemistry learning. *Journal of Research in Science Teaching, 49*(3), 394-419.

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

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_feb.pdf

Roth, W-M., & Bowen, G.M., (2001). Professionals read graphs: A semiotic analysis. *Journal for Research in Mathematics Education*, *32*, 159-194.

Serra, H., & Godoy, W. A. C. (2011). Using ecological modeling to enhance instruction in population dynamics and to stimulate scientific thinking. *Creative Education*, *2*(2), 83-90.

Seymour, E., & Hewitt, N. M. (1994). *Talking about leaving: Factors contributing to high attrition rates among science, mathematics & engineering undergraduate majors: Final report to the Alfred P. Sloan Foundation on an ethnographic inquiry at seven institutions*. Boulder, CO: Bureau of Sociological Research, University of Colorado.

Shubert, C., Ceraj, I., & Riley, J. (2009). Bringing research tools into the classroom. *Journal of Computers in Mathematics and Science Teaching*, *28*(4), 405-421.

Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (Eds.). (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.

Smetana, L. K., & Bell, R. L. (2012). Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education*, *34*(9), 1337-1370.

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.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 267–306). Mahwah, NJ: Lawrence Erlbaum Associates.

Summerlin, L., & Gardner, M. (1973). A study of tutorial-type computer assisted instruction in high school chemistry. *Journal of Research in Science Teaching*, *10*(1), 75-82.

Thornton, R. K. (1987). Tools for scientific thinking-microcomputer-based laboratories for physics teaching. *Physics Education*, *22*(4), 230.

Von Glasersfeld, E. (1991). *Radical constructivism in mathematics education*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Vygotsky, L. S. (1979). Consciousness as a problem in the psychology of behavior. *Journal of Russian and East European Psychology*, *17*(4), 3-35.

Weller, H. G. (1996). Assessing the impact of computer-based learning in science. *Journal of Research on Computing in Education*, *28*(4), 461-85.

Wood, W. B. (2009). Innovations in teaching undergraduate biology and why we need them. *Annual Review of Cell and Developmental*, *25*, 93-112.

Woodin, T., Carter, V. C., & Fletcher, L. (2010). Vision and change in biology undergraduate education, a call for action—initial responses. *CBE-Life Sciences Education*, *9*(2), 71-73.

Appendix A: Sophomore and junior undergraduate biology student subject questionnaire.

Please answer the following on a scale of 1 to 5, with 5 being "Strongly agree" and 1 being "Disagree." Answer N/A if you are unsure or if the question does not apply to you.

1. I am comfortable navigating computers.
2. I enjoy learning new software programs.
3. I am able to follow along with guided exercises for software programs I am unfamiliar with.
4. I am familiar with basic data graphing methods (i.e. line graphs, histograms)
5. I have taken at least one course that included graphing biological data.
6. I have used Excel for biological graphing exercises.
7. I have used SigmaPlot for biological graphing exercises.
8. I can provide a laptop computer for my use in this research study, have Microsoft Excel available on my laptop, and am willing to install trial versions of two programs.

Appendix B. Datasets used for Session 1: Simple Graphing Exercises in Excel and SigmaPlot

Table B1
Segment of data for use in session one with Excel simple graph method tutorial; data are selections from the Rocky Mountain Biological Laboratory (www.rmbl.org) long term data series. Participants will use data for generating line graphs and histograms/frequency distributions in Microsoft Excel.

dd/mm/yyyy	year	month	day	min temp (°C)	max temp (°C)		year	mean max temp (°C)	std. error
1/1/2006	2006	1	1	-7.5	-0.5		2006	9.4	0.54
1/2/2006	2006	1	2	-13.0	-1.5		2007	10.3	0.56
1/3/2006	2006	1	3	-8.5	0.5		2008	8.1	0.60
1/4/2006	2006	1	4	-13.5	-4.0		2009	8.6	0.55
1/5/2006	2006	1	5	-12.0	-4.5		2010	6.7	0.64
1/6/2006	2006	1	6	-15.0	3.0				
1/7/2006	2006	1	7	-12.0	3.5				
1/8/2006	2006	1	8	-9.5	-2.5				
1/9/2006	2006	1	9	-17.0	-7.0				
1/10/2006	2006	1	10	-20.0	-7.0				
1/11/2006	2006	1	11	-16.5	-0.5				
1/12/2006	2006	1	12	-14.5	-6.0				
1/13/2006	2006	1	13	-19.0	-5.0				
1/14/2006	2006	1	14	-13.0	1.0				
1/15/2006	2006	1	15	-10.0	-3.5				
1/16/2006	2006	1	16	-14.5	-7.0				
1/17/2006	2006	1	17	-14.0	-2.0				
1/18/2006	2006	1	18	-7.0	-1.0				
1/19/2006	2006	1	19	-8.5	-4.5				
1/20/2006	2006	1	20	-19.0	-8.5				
1/21/2006	2006	1	21	-22.5	-10.0				
1/22/2006	2006	1	22	-21.5	-8.5				
1/23/2006	2006	1	23	-24.0	-7.5				

Table B2

Segment of data for use in session one with SigmaPlot 12.0 simple graph method tutorial; data are selections from the Rocky Mountain Biological Laboratory (www.rmbl.org) long term data series. Participants will use data for generating line graphs and histograms/frequency distributions in SigmaPlot.

dd/mm/yyyy	year	month	day	total snow (cm)	snow pack (cm)	year	mean snow pack (cm)	std. error
1/1/1995	1995	1	1	410	93	1995	82	4.41
1/2/1995	1995	1	2	410	91	1996	76	4.12
1/3/1995	1995	1	3	411	91	1997	78	4.28
1/4/1995	1995	1	4	415	94	1998	53	3.00
1/5/1995	1995	1	5	433	109	1999	46	2.84
1/6/1995	1995	1	6	443	112			
1/7/1995	1995	1	7	467	129			
1/8/1995	1995	1	8	482	132			
1/9/1995	1995	1	9	484	127			
1/10/1995	1995	1	10	484	122			
1/11/1995	1995	1	11	500	134			
1/12/1995	1995	1	12	515	138			
1/13/1995	1995	1	13	522	138			
1/14/1995	1995	1	14	523	132			
1/15/1995	1995	1	15	523	127			
1/16/1995	1995	1	16	535	130			
1/17/1995	1995	1	17	537	127			
1/18/1995	1995	1	18	538	126			
1/19/1995	1995	1	19	539	124			
1/20/1995	1995	1	20	539	122			
1/21/1995	1995	1	21	539	121			
1/22/1995	1995	1	22	539	120			
1/23/1995	1995	1	23	539	120			

Appendix C. Datasets used for Session 2: Complex Graphing Exercises in Excel and SigmaPlot

Table C1
Segment of data for use in session two with Excel complex graph method tutorial; data are selections from the Rocky Mountain Biological Laboratory (www.rmbl.org) long term data series. Participants will use data for generating complex line graphs, scatter plots and box and whisker plots in Microsoft Excel. Data for the years 2000 and 2009 box and whisker plot activity purposefully left blank to demonstrate data preparation needs.

dd/mm/yyyy	year	month	day	min temp (°C)	max temp (°C)	melt water (mm)	total snow (cm)	snow pack (cm)
11/1/1980	1980	11	1	-9.0	7.0	0.00	121	28
11/2/1980	1980	11	2	-11.0	5.5	0.00	121	27
11/3/1980	1980	11	3	-8.5	5.0	0.00	121	25
11/4/1980	1980	11	4	-6.0	4.0	0.00	121	24
11/5/1980	1980	11	5	-8.0	8.0	0.00	121	20
11/6/1980	1980	11	6	-8.0	6.5	0.00	121	17
11/7/1980	1980	11	7	-4.5	7.0	0.00	121	14
11/8/1980	1980	11	8	-4.5	5.0	0.00	121	10
11/9/1980	1980	11	9	-4.5	7.0	0.00	121	8
11/10/1980	1980	11	10	-7.5	8.0	0.00	121	5
11/11/1980	1980	11	11	-8.5	7.0	0.00	121	4
11/12/1980	1980	11	12	-4.0	5.0	0.00	121	3
11/13/1980	1980	11	13	-5.0	-1.0	16.26	144	20
11/14/1980	1980	11	14	-9.0	-3.0	2.29	149	21
11/15/1980	1980	11	15	-21.5	-4.0	0.00	149	20

	1980	1990
Max	8.0	6.5
3rd Quartile	4.5	2
Median	1.0	-2.0
1st Quartile	-1.5	-8.5
Min	-8.0	-20.5

Table 2C

Segment of data for use in session two with SigmaPlot 12.0 complex graph method tutorial; data are selections from the Rocky Mountain Biological Laboratory (www.rmbl.org) long term data series. Participants will use data for generating complex line graphs, scatter plots and box and whisker plots in SigmaPlot.

dd/mm/yyyy	year	month	day	min temp (°C)	max temp (°C)	melt water (in)	melt water (mm)	snow pack (cm)	Rainfall (mm)
6/9/2007	2007	6	9	0.0	21.0	0.00	0.00	0	0.00
6/10/2007	2007	6	10	1.0	23.0	0.00	0.00	0	0.00
6/11/2007	2007	6	11	3.0	21.0	0.00	0.00	0	0.00
6/12/2007	2007	6	12	6.0	13.0	0.00	0.00	0	17.78
6/13/2007	2007	6	13	2.0	19.5	0.00	0.00	0	0.00
6/14/2007	2007	6	14	1.5	23.0	0.00	0.00	0	0.00
6/15/2007	2007	6	15	3.5	25.0	0.00	0.00	0	0.00
6/16/2007	2007	6	16	4.0	25.0	0.00	0.00	0	0.00
6/17/2007	2007	6	17	4.5	24.0	0.00	0.00	0	0.00
6/18/2007	2007	6	18	4.5	21.5	0.00	0.00	0	0.00
6/19/2007	2007	6	19	1.5	24.5	0.00	0.00	0	0.00
6/20/2007	2007	6	20	4.0	25.5	0.00	0.00	0	0.00
6/21/2007	2007	6	21	3.5	25.0	0.00	0.00	0	0.00
6/22/2007	2007	6	22	4.5	22.5	0.00	0.00	0	2.29
6/23/2007	2007	6	23	3.0	26.0	0.00	0.00	0	0.00
6/24/2007	2007	6	24	4.0	26.5	0.00	0.00	0	0.00
6/25/2007	2007	6	25	3.5	25.5	0.00	0.00	0	0.00
6/26/2007	2007	6	26	3.5	25.5	0.00	0.00	0	0.00
6/27/2007	2007	6	27	5.0	26.5	0.00	0.00	0	0.00

Appendix D. Tasks and datasets for Session 3: Final Graph Tasks

Table D1
List of tasks for participants to address using graph methods in the program of their choice, Excel or SigmaPlot. Tasks specifically delineate the graphs that should be used to address each task to eliminate variability in user selection of graph type.

Task #	Use the following tasks to address the research question: Has the Yellow-Bellied Marmot maintained consistent population characteristics in its indigenous Rocky Mountain habitat over a 30 year observation period? Has the Rocky Mountain habitat been environmentally consistent over 30 years?
1	**Task**: How has environmental conditions changed over the 30 year period? *Graph the minimum and maximum temperatures from 1978 to 2010 using a line graph with time on the x axis and the same y axis.*
2	**Task**: Is there a difference in the amount of snow pack over time and how does this relate to melt water in inches? *Graph both datasets using a line graph with time on the x axis and two different Y axes*.
3	**Task**: Has the population of Yellow-bellied Marmots remained consistent over 30 year period, an how does their populations vary during the year? *Create box and whisker plots for YB Marmot Count for each year, with year on the X axis and count on the Y axis.* (If you decide to use Excel, there are data for every few years, so you can just graph what's available)
4	**Task:** How does the Yellow-bellied Marmot population vary with maximum temperature from 1978 to 2008. *Graph YB Marmot Count and Max Temp using a scatter plot. Add a trendline.*
5	**Task**: Do the same trends in adult weight and abundance occur with juveniles and yearlings? *Graph both the Avg. Yearling Wt. (g) and Avg. Juvenile Wt. (g) datasets with YB Marmot Count on the x axis using a scatter plot. Add trendlines for both datasets.*
6	**Task:** What is the distribution of seasonal Yellow-Bellied Marmot average adult weight in two year intervals from 1990 to 2008*? Create a histogram/frequency distribution of Avg. Adult weights (g). Break up the weights into 5-6 different ranges/number of bins.*

Table D2

Subset of the Rocky Mountain Biology Laboratory to be used for session 3 graphing exercises in response to tasks list. Spreadsheet is condensed to include only one day per year of physical data, and information on yellow-bellied marmot count was generated based on research in this region (http://rmbl.info/rockymountainbiolab/Ozgul_etal_2010_Nature_all.pdf).

year	month	day	min temp (°C)	max temp (°C)	new snow (cm)	melt water (in)	melt water (mm)	snow pack (cm)	YB Marmot Count
1978	1	1	-27.0	-11.0	0	0.00	0.00	109	1.00
1978	2	1	-20.0	-7.0	0	0.00	0.00	133	0.00
1978	3	1	-9.0	-6.0	45	1.15	29.21	226	0.00
1978	4	1	-6.0	-3.0	10	0.23	5.84	179	2.00
1978	5	1	-8.5	0.0	3	0.03	0.76	137	2.00
1978	6	1	-4.5	10.5	0	0.00	0.00	18	26.00
1978	7	1			0	0.00	0.00	0	35.00
1978	8	1			0	0.00	0.00	0	78.00
1978	9	1			0	0.00	0.00	0	56.00
1978	10	1			0	0.00	0.00	0	43.00
1978	11	1	-5.0	5.0	0	0.00	0.00	0	5.00
1978	12	1	-7.5	-6.0	44	1.01	25.65	71	3.00
1979	1	1	-31.0	-24.5	0	0.00	0.00	121	1.00
1979	2	1	-15.5	-13.0	32	1.02	25.91	206	0.00
1979	3	1	-12.5	-3.5	0	0.00	0.00	196	0.00
1979	4	1	-19.0	-5.5	5	0.10	2.54	216	2.00
1979	5	1	-3.5	2.0	8	0.15	3.81	130	17.00
1979	6	1			0	0.00	0.00	0	36.00
1979	7	1			0	0.00	0.00	0	85.00

Table D3

Subset of the Rocky Mountain Biology Laboratory to be used for session 3 graphing exercises in response to tasks list. All data relating to yellow-bellied marmot count were generated based on research using the physical data from this region (http://rmbl.info/rockymountainbiolab/Ozgul_etal_2010_Nature_all.pdf).

Year	Season	YB Marmot Count	Avg. Juvenile Wt. (g)	Avg. Yearling Wt. (g)	Avg. Adult Wt. (g)
1990	Winter	1	N/A	N/A	2741.3
1990	Spring	184	112.34	1237.64	2956.24
1990	Summer	316	159.67	1384.34	3014.51
1990	Fall	39	234.11	1411.97	3164.78
1992	Winter	1	N/A	N/A	2845.32
1992	Spring	49	110.34	1334.61	3070.21
1992	Summer	208	143.64	1437.11	3096.11
1992	Fall	13	217.64	1427.33	2963.21
1994	Winter	1	N/A	N/A	2649.37
1994	Spring	30	96.37	1137.33	2744.34
1994	Summer	136	128.33	1219.34	2864.37
1994	Fall	8	194.67	1264.33	2947.33
1996	Winter	1	N/A	N/A	2799.31
1996	Spring	99	106.24	1269.16	2944.75
1996	Summer	268	134.65	1396.34	3097.65
1996	Fall	29	201.34	1400.94	3104.25
1998	Winter	1	N/A	N/A	2843.63
1998	Spring	77	101.34	1264.38	2994.1
1998	Summer	333	141.67	1459.69	3124.61
1998	Fall	21	196.37	1564.03	3154.25
2000	Winter	0	N/A	N/A	2674.19
2000	Spring	88	96.88	1108.72	2789.91
2000	Summer	124	138.64	1200.69	2978.12

Appendix E. Interview questions for participants

1. Which program do you prefer to use for basic graph methods, Excel or SigmaPlot, and why?
2. Which program do you prefer to use for advanced graphing methods, Excel or SigmaPlot, and why?
3. What do you think are the strengths and advantages of Excel? What are its deficits? What do you think are the strengths and advantages of Sigmaplot? What are its deficits?
4. You were given one example of biological and environmental data but this might not represent the type of data you would encounter in your specific biological discipline. If you were graphing data specific to the biological subject you would like to pursue in your career, would you anticipate using one program over the other? If so which program would you prefer to use and why?