Spring 5-6-2012

Bar Graphs & Baselines: Student Perceptions of Distortions in Real World Graphs

Robert K. Janes III
University of Connecticut - Storrs, robert.janes@uconn.edu

Follow this and additional works at: https://opencommons.uconn.edu/srhonors_theses

Part of the Science and Mathematics Education Commons

Recommended Citation
https://opencommons.uconn.edu/srhonors_theses/258
Bar Graphs & Baselines:

Student Perceptions of Distortions in Real World Graphs

Robert K Janes III

University of Connecticut
# Table of Contents

Abstract ........................................................................................................................................... 5  

Introduction ..................................................................................................................................... 6  

Literature Review ............................................................................................................................ 9  
  - Definition and Levels of Graph Comprehension ................................................................. 9  
  - Empirically Based Factors Influencing Graph Comprehension ....................................... 11  
    - Dimensionality ................................................................................................................ 12  
    - Coloration ....................................................................................................................... 14  
    - Proximity .......................................................................................................................... 14  
    - Visual Chunks .................................................................................................................. 15  
    - Familiarity ......................................................................................................................... 16  
  - Theories of Graphical Distortions ..................................................................................... 17  
  - Graphical Distortions in Context ...................................................................................... 18  
  - Summary .............................................................................................................................. 20  

Method .......................................................................................................................................... 21  
  - Note about Bar Graphs and Histograms ........................................................................... 21  
  - Survey Instrument ............................................................................................................. 21  
    - Selected Response Comprehension Items in Forms A, B, and C. ........................ 23  
    - Constructed Response Prompts in Forms A, B, and C. ........................................ 26  
    - Data About Validity in Form D. ............................................................................... 27
BAR GRAPHS & BASELINES

Demographic Items in Form E................................................................. 28
Lack of Reliability Information ............................................................... 28
Participants ................................................................................................. 28
Participant Demographics ........................................................................ 29
Local Demographics ................................................................................ 29
Local and State Curriculum ...................................................................... 30
Data Collection .......................................................................................... 31
Data Analysis .............................................................................................. 32
  Coding of Constructed Response Items on Forms A, B, and C ............... 33
  Tests for Construct Validity ................................................................. 35
Results and Discussion ............................................................................... 36
  Varying Scores across Three Levels of Distortion ............................... 36
  Some Heavy Distortions may be too Obvious ...................................... 37
  Difference between Scores on Qualitative and Quantitative Items ....... 38
  Scores across Four Class Levels .......................................................... 40
    Relationship between Course Level and Graph Comprehension ........ 40
    Disparity between Qualitative and Quantitative Scores by Course Level 41
  Trends in Constructed Responses ......................................................... 43
    Participants want a Reason for Trends .............................................. 44
    Reporting the Non-Zero Baseline ...................................................... 45
Abstract

It is important for every educated member of our society to be able to read, comprehend, and interpret graphs. To that end, the National Council of Teachers of Mathematics (NCTM) and the Common Core State Standards have endorsed a kindergarten through eighth-grade mathematics curriculum that is rich in data analysis and graphical literacy skills. These skills are important, as graphs in the public sphere may include certain features that can cause misperceptions of the data. Such features may be intentional or unintentional and can include non-zero baselines, representing data with extra dimensions, stretching and shrinking graphics, not displaying outliers, and more. It is unclear whether the recommended mathematics curriculum prepares students to accurately comprehend these kinds of graphs.

This study investigates how effectively the kindergarten through eighth-grade public mathematics curriculum prepares students to read graphs that contain distortions. It focuses on student perceptions and interpretation of data presented in bar graphs. A survey instrument was created to measure student perceptions that included comprehension questions referring to bar graphs with zero baselines and others with non-zero baselines. The instrument was used to assess 181 ninth-grade students in a school in a New England suburb. The results from this survey suggest that many students are susceptible to graphical misperceptions while reading a graph with a non-zero baseline. It suggests that while the current curriculum may include strategies to solve mathematically rigorous graph comprehension problems, it may not equip students with the ability to make simple qualitative conclusions about real world graphs. It is the hope of the researcher that this study may inform future curricula on a local and state level.

Keywords: Math, Education, Graph, Comprehension
Introduction

*Graphical comprehension*, i.e. the process of reading, understanding, and interpreting information from a graph, is of utmost importance in modern society. Graphs can provide the reader with a quick summary of a set of data while simultaneously displaying individual data points. As a result, they are often used to make informed decisions about a variety of topics, including medicine, finance, nutrition, politics, and safety (Galesic & Garcia-Retamero, 2010).

As a result, leaders in mathematics education have pushed for a curriculum that is rich in data analysis techniques with frequent construction and interpretation of graphs. The Common Core State Standards highlight data collection and statistical analysis in kindergarten through eighth-grade (National Governors Association Center for Best Practices, Council of Chief State School Officers [Common Core], 2010). Students are expected to be able to draw and interpret picture graphs and bar graphs in second grade (2.MD.10), and quantitative statistical analysis begins in sixth grade (6.SP.5). Similarly, The National Council for Teachers of Mathematics (NCTM) now recommends a rigorous curriculum including data analysis and graph comprehension for students in kindergarten through eighth grade (Curcio, 2010). Note that Common Core and NCTM place a reduced emphasis on data analysis and graph comprehension once students enter high school in ninth grade.

It is important to first understand the history of graphs in order to gain perspective. Most commentators credit William Playfair with the first use of modern graphs in his works *Commercial and Political Atlas* (1786) and *Statistical Breviary* (1801). In the mid-20\(^{\text{th}}\) century, data graphics (i.e. graphs) became an increasingly popular method for conveying quantitative data. With the dramatic increase in popularity of data graphics came an increase in the number of graphs with features that could cause the viewer to misinterpret the data being represented. These
features are called *graphical distortions*. These distortions ranged from accidental to intentional and can include non-zero baselines, representing data with extra dimensions, stretching and shrinking graphics, not displaying outliers, and more. Distortions can appear in a variety of types of graphs including bar graphs, picture graphs, line graphs, and pie charts, and many more. As a result, many authors began to comment on these graphical distortions by classifying the different types and suggesting ways to remedy each (Huff, 1954; Jones, 1995; Tufte, 1983; Turkey & Wilk, as cited in Tufte, 1983).

Despite the work of these and other commentators, graphical distortions are still commonplace today. While most have a minimal effect, some distortions can have widespread consequences. For instance, the Space Shuttle *Challenger* disaster has been blamed in part on a NASA team who made the decision to launch the shuttle based on data displayed in a faulty graph. NASA engineers needed to determine if a certain component would fail under any adverse temperatures. The component was tested, and a graph was created to display the data. The graph organized the data by testing date rather than temperature, the most critical factor. Commentators claim that the choice to organize the data by test date created a graphical distortion that caused misperceptions among NASA engineers (Tufte, 1997). In addition, critics have noted that the CO$_2$ graphs from the film *An Inconvenient Truth* are distorted to make the issue of global warming seem more salient (Guggenheim, 2006). The baseline of these graphs does not start at zero, and critics claim that the change in the baseline causes temperature fluctuations to seem larger than they actually are.

Empirical studies have also played an important role. Researchers have investigated connections between a participant’s level of graph comprehension and specific features of the graph. These features include the familiarity the participant has with the graph (Xi, 2010), the
coloration of the graph (Stewart, Cipolla & Best, 2009), the type of graph being used (Shah, Mayer, & Hegarty, 1999), the number of dimensions represented in the graph (Shah & Carpenter, 1995; Stewart et. al., 2009), and the proximity of graphical elements (Shah et al., 1999; Trickett & Trafton, 2004).

Although research has shown that there are many factors that influence graph comprehension, there is little research that examines graphical misperceptions. Graphical misperceptions occur when a viewer’s level of graphical comprehension is lowered in the presence of graphical distortions. There have been some studies that explore graphical misperceptions (Galesic & Garcia-Retamero, 2010; Monteiro & Ainley, 2003), but none have focused on high school students. Given the prevalence of graphical distortions in today’s society (Galesic & Garcia-Retamero, 2010), it is important for students graduating from high school to be able to recognize graphs that misrepresent the data and maintain a high level of graph comprehension despite these potentially misleading features. Yet, it is unclear whether the recommended curriculum prepares students to accurately comprehend distorted graphs.

Therefore, this study will investigate whether students are able to maintain a high level of graph comprehension while reading bar graphs with non-zero baselines. Bar graphs are one of the most basic types of graphs, and non-zero baselines are one of the most frequently occurring graphical distortions. The focus will be on high school students in ninth grade, since these students have recently completed eighth grade, and have recently completed a mathematics curriculum modeled after NCTM recommendations. These students will just be entering high school, a critical point in mathematics education. If students are unable to detect distortions at this level, educators still have ample time to address these concepts. A review of the literature is in order.
Literature Review

In this section, I explicitly define what is meant by graphical comprehension, and discuss a framework to specify levels of comprehension. Although there has been some debate as to the titles of these levels, there is a general consensus among researchers that there are three unique levels with distinct characteristics.

This review will also discuss previous research that focuses on factors influencing graph comprehension. Discussion will start with empirical studies that have reported connections between specific features of a graph and a participant’s level of graph comprehension. Unfortunately, there are very few general theories of graphical distortions that are backed by empirical studies. As a result, this review will also discuss non-empirical theories. Although these theories rely on the notions of commentators and not on data driven research, they are heavily cited and accepted by most researchers as true. Lastly, there will be an examination of the few studies that do investigate connections between distortion in graphs and the perceptions of the readers. Taken together, this review will present the current understanding of the effect distortions have on graph comprehension.

Definition and Levels of Graph Comprehension

In general, there are three types of comprehension behaviors: translation, interpretation, and extrapolation (Jolliff, 1991). Translation involves a change in form of a communication, such as by explaining numerical data using words or describing a picture verbally. Interpretation requires the ability to rearrange information in order to draw inferences. Extrapolation is an extension of interpretation, and includes conclusions about consequences and effects.

In a study examining the effects of prior knowledge on graph comprehension, Curcio (1981a; 1981b; 1987) postulated that there are three specific levels of graph comprehension.
These are similar to the three levels proposed by Jolliff. They are labeled as follows: *reading the data*, *reading between the data*, and *reading beyond the data*. NCTM has recognized this three level system and currently uses it to classify comprehension questions intended for student assessment (Curcio, 2010). Various researchers have used similar systems, and although different labels have been used to describe the three levels, there is a general consensus as to the definitions of these levels (Bertin, 1983; Carswell, 1992; McKnight, 1990; Wainer, 1992). For the purposes of this review, this study will use the labels provided by Curcio and recognized by NCTM. The three levels are defined as follows:

*Reading the data* is the most elementary level of graph comprehension; it requires the reader to extract information from the data (Friel, Curcio, & Bright, 2001). Questions of this type generally focus on observing single facts and relationships that are directly stated in the data (McKnight, 1990). The reader simply reports facts that are apparent by inspecting the graph, graph title, or axis labels. There is no interpretation or calculation at this level, and reading that requires this kind of comprehension is considered to be a low level cognitive task (e.g., How tall is Maria?; Curcio, 2010).

*Reading between the data* is an intermediate level of graph comprehension at which the reader is able to find relationships in the data through interpretation and integration (Friel et al., 2001). This level requires the ability to compare more than one quantity (e.g., bigger than, less than, heavier than) and use other mathematical concepts (e.g., addition, subtraction, multiplication, division). The reader uses at least one step of logical reasoning to determine the answer to a question (e.g., How much taller is Maria than Jose?; Curcio, 2010).

The third and most advanced level of graph comprehension is *reading beyond the data*. This behavior can involve extending, predicting, or inferring from the data to draw conclusions.
(Friel, Curcio, & Bright, 2001). These conclusions can then be used to create, support, or reject a proposition (i.e., to provide evidence; McKnight, 1990). Reading beyond the data often requires the reader to tap into an existing schema: background knowledge or knowledge already in memory. Whereas reading between the data necessitates using only the information in the graph, reading beyond the data involves information that must be inferred. Inferences may include extrapolating, finding a probability, describing a long-term trend, or creating a general rule for the data (e.g., Will Jose ever be taller than Maria? Why or why not?; Curcio, 2010).

Multiple studies have adopted this three level system of graph comprehension (Galesic & Retamero, 2010; Monteiro & Ainley, 2003). One such study by Stewart, Cipolla, and Best (2009) explored the differences in levels of comprehension required to understand two-dimensional and three-dimensional graphs. Participants answered comprehension questions based on various two and three dimensional graphs. Curcio’s three level system was used by Stewart et al. to classify the difficulty of their test questions and create an objective method of scoring their participants. In the discussion, researchers agreed that the three level system of graph comprehension was an appropriate way to classify comprehension questions (p. 38).

In general, graph comprehension is a reader’s ability to derive meaning from a graph, which can be categorized by Curcio’s three level system according to cognitive difficulty (Friel et al., 2001). This system will be used to classify the level of comprehension required to answer questions about graphs for the remainder of the study.

**Empirically Based Factors Influencing Graph Comprehension**

Research has indicated that there are multiple variables that affect the level of comprehension of a particular graph (Katz, Xi, Kim, & Cheng, 2004; Shah & Carpenter, 1995; Stewart et al., 2009). A subset of these variables concerns the construction of the graph. Altering
these variables may cause information to become more or less salient, and to require differing amounts of mental computation by the readers (Shah et al., 1999). This in turn can impact levels of graph comprehension. While many variables have been examined, the following are of importance to this study: dimensionality, coloration, proximity, number of elements, familiarity, and type. Each of these will be discussed here.

**Dimensionality.** The number of physical dimensions used in a graph to describe data has an effect on the average level of graph comprehension. Research indicates that three dimensional graphs interfere with graph comprehension. Stewart and colleagues (2009) found that participants had higher accuracy and decreased response time when answering comprehension questions about two-dimensional graphs, and lower accuracy and increased response time when answering comprehension questions about three dimensional graphs. In addition, the researchers found that while there was no statistically significant difference in accuracy of responses to easy questions (i.e., reading the data) with two-dimensional and three-dimensional graphs, there was a statistically significant difference in the accuracy of responses to difficult questions, i.e., *reading beyond the data*. The biggest disparity was present in difficult questions referring to bar graphs; participants were able to answer 78.1% of the comprehension questions related to two-dimensional graphs, but only 60.4% of those for three-dimensional graphs. Similar data were collected pertaining to response time: participants took longer to answer comprehension questions related to graphs with more dimensions.
Figure 1. Two of the graphs used by Shah and Carpenter (1995) in their experiment on the number of graphical relationships that can be understood at once.

Shah and Carpenter (1995) found that graph readers may have a limit as to how many graphical dimensions they can retain and process at one time. Participants in their study were students from Carnegie Mellon University. Researchers first presented participants with a multiple line graph, and asked them to interpret the information verbally. Graph A in Figure 1 is one of these graphs. Notice that the graph depicts three continuous variables. Noise level and achievement were represented on the x-axis and y-axis, respectively. Room temperature was the third variable, also known as the z-variable. It was displayed at two different values by two different lines on the graph; the two values are indicated in the key. This essentially transformed the third variable from continuous to discrete. Participants frequently drew conclusions relating the y-axis variable to the x-axis variables, but infrequently connected the y-axis to the z-variable. The researchers speculated that these errors were caused in part by a limitation on the number of relationships that can be comprehended and simultaneously retained.
Shah and Carpenter (1995) also presented participants with two multiple line graphs from different perspectives. The graphs were comparable, except the data from the y-axis and z-axis were swapped. The participants were then asked if the two graphs represented the same data set. Figure 1 contains a pair of graphs that do depict the same data set. On average, participants answered correctly 62% of the time. Note that if the participants answered randomly, they would be correct approximately 50% of the time. The conclusion was that participants had difficulty mentally manipulating multiple relationships, and thus were unable to compare two graphs of differing perspectives. In general, these studies show that graph readers have difficulty drawing conclusions from graphs when they must interpret multiple dimensions of data simultaneously.

**Coloration.** The presence of color in a graph may have an effect on graph comprehension, although theorists’ and researchers’ findings are conflicting. Theorists have suggested that color is an important part of graph comprehension (Tufte, 1983, 1997) and that the addition of color in a careful and meaningful way may lead to higher graph reading accuracy (Kosslyn, 1994). However, one study determined that there was no statistically significant difference in graph comprehension between graphs in color and graphs in grayscale for 22 participants with little statistical background (Stewart et al. 2009). The researchers in this study speculated that the coloration of a graph may affect a graph reader’s initial perception of a graph but not the overall graph comprehension. The authors commented that more research should be done on the presence of color in graphs to reach a more definite conclusion.

**Proximity.** The proximity, or physical distance between elements of a graph, can cause differing levels of graph comprehension by making parts of the data more salient than others. Simply moving one graphical element farther away from another decreases the probability that the reader will make a comparison between those two elements (Shah, Mayer, & Hegarty, 1999).
Additionally, proximity can affect the accuracy of comparisons. If the reader is forced to make thoughtful comparisons between two elements, the accuracy of those comparisons decreases as the physical gap between the elements increases. A recent study has shown that even experienced graph readers make mistakes when using large spatial transformations to extract information (Trickett & Trafton, 2004). Note that although saliency and accuracy of comparisons are related, they are two different constructs. In both cases, the proximity of elements in a graph has an effect on the reader’s ability to read between the data.

A quantitative study by Shah, Mayer, and Hegarty (1999) showed that visual chunking has an effect on the types of relationships participants found in the data. They explained that multiple elements can be visually chunked together if they are in close physical proximity or if they are connected with a line or other physical element of the graph. The researchers took various graphs found in a middle school history textbook and asked participants to make conclusions based on the graph. The researchers then reconstructed the graphs by changing how the data were visually grouped on the graph. They found that participants were more likely to discuss the relationships between individual data points when they were placed in close proximity or connected with a line. In addition, they found that participants were able to make more accurate statements about trends in data when the data points were grouped together through proximity or being physically connected, i.e. the data points were touching or connected with a straight line in the case of line graphs. Thus, the proximity and connectedness of data points also has an influence on participants’ ability to read beyond the data.

**Visual Chunks.** A study by Katz, Xi, Kim, and Cheng (2004) revealed that graphs with fewer visual chunks facilitated higher levels of graph comprehension. Their research continued the work done by Shah and colleagues (1999) on visual chunks, but focused on the number as
opposed to the proximity of the chunks. The number of visual chunks in a graph is defined as the number of unique graphical elements that are adjacent to each other (in the case of bar graphs) or connected with a line (in the case of line graphs). A change in the slope of a line from positive to negative is said to delineate two separate visual chunks. Researchers asked participants to describe line and bar graphs with a varying number of visual chunks. Participants’ responses were scored by their fluency, organization, and content. The results indicated graphs with fewer visual chunks led to more coherent, cohesive, and higher level descriptions. In addition, participants took less time to scan the few-chunks graphs before speaking (Katz et al., 2004). Thus, a participant’s level of graph comprehension is affected negatively by the number of visual chunks present within a graph.

**Familiarity.** Research has shown that graph readers tend to read data more accurately when they are more familiar with the type of graph in question. Xi (2004) gave participants a verbal exam consisting of graph description tasks followed by a survey designed to measure their familiarity with various types of graphs. Participants consisted of college graduates and undergraduates. The data collected showed that participants’ responses contained higher level content and better organization when describing graph types they felt familiar with. Thus, participants were able to locate, retrieve, and integrate information from graphs with more proficiency when they were comfortable with that particular graph type. Interestingly, the participants’ level of fluency was unaffected by graph familiarity. Fluency refers to the ease with which the participants spoke. This could indicate that communication skills remain unhindered even when the participant has trouble processing information (Xi, 2004). Thus, familiarity with the type of graph in question can lead to increased levels of graph comprehension.
As has been shown, there are many graphical variables that can impact the graph comprehension of the reader. Taken together, this set of graph variables (dimensionality, coloration, proximity, number of elements, familiarity, and type) can be used to create a graph that should present the optimal situation for high graph comprehension. This sort of graph should have a limited number of dimensions, have a low number of highly connected visual chunks, and be familiar to the viewer. On the other hand, the research also provides a way to create a graph intended to elicit the lowest level of graph comprehension possible. This kind of graph would have a high number of dimensions within the same graph, have a large number of disjoint elements, and utilize a graph type which was foreign to the viewer. These are the types of graphs that will be explored further in this review: graphs that purposefully aim to lower graph comprehension. The elements used to achieve this are known as *graphical distortions*. Graphical distortions are any graphical element which noticeably decreases the graph comprehension of the viewer.

**Theories of Graphical Distortions**

A number of theorists have discussed the presence and of graphical distortions, usually in an attempt to warn their audience about the dangers associated with them. These theorists attempt to classify, and occasionally quantify, the various types of distortions that are used to purposefully lower levels of graph comprehension and encourage graphical misperceptions. It should be noted that most of these theorists do not rely on empirical data to arrive at their conclusions; rather, they use their judgment and expertise to make conjectures.

The most relevant of these conjectures is the idea that graphical trends can be made more salient by changing the scale of a graph (Huff, 1954; Jones, 1995; Tufte, 1983). Authors have noted that graph readers often associate the bottom of a graph with a value of zero. For example,
the lowest value on a bar graph occurs at the x-axis which normally represents a y-value of zero. This value is called the baseline of the graph. Graph designers can set the baseline to a different value in an effort to “stretch” or “shrink” the graphs. Suppose a set of data had a range of 1990 to 2000, and suppose this data set fluctuated randomly within this range. If this data set was displayed with a baseline of zero, the changes in the data would only appear on the upper 5% of the graph. As a result, the fluctuations would appear to be insignificant. But, if the baseline was changed to 1990, the data points would appear to be farther apart, and the fluctuations would become more obvious (Huff, 1954; Jones, 1995; Tufte, 1983). Thus, theorists believe that the ability to detect trends, or read beyond the data, can be influenced by the baseline that a particular graph uses.

Commentators such as Huff, Jones, and Tufte claim that there are other types of graphical distortions including dimensionality, coloration, type, and proximity. In addition, they believe these kinds of graphical distortions are widespread, and that being able to identify them is crucial to accurate graph comprehension. Given the examples from Space Shuttle Challenger and An Incontinent Truth mentioned in the introduction, there is reason to believe that their claims are at least partially true. However, it is important to recall that these commentators have supported their conjectures with few research studies.

**Graphical Distortions in Context**

Until this point in the review, theories of graphical distortions have not been supported directly with empirical data. There is a low amount of research available that has studied graphical distortions in the context of real-world situations. This is because it is difficult to measure the amount of distortion within one graph. While more controlled studies can manipulate one graphical element at a time, graphs as described by Huff, Jones, and Tufte are
difficult to quantify. However, some data have been collected on the impact of generally distorted graphs on levels of graph comprehension.

Monteiro and Ainley (2003) explored the use of critical sense when reading real world graphs. The researchers presented 10 undergraduate student teachers (of various concentrations) with two graphs taken from news magazines. Although both graphs contained two data sources, one combined the two sources into one graphic, while the other used two separate graphics. In an interview, participants were asked to respond to the question, “If you could talk to the person who produced this graph, are there any questions you would like to ask?” The researchers noted that the responses differed drastically between the two graphs.

Participants were more likely to make comparisons between data sets that were presented with two separate graphs, rather than one merged graph. The researchers also noted that participants were able to ask more critical questions when data was displayed in two separate graphics. They conjectured that participants thought more critically about this graph because the spatial transformation required to compare the two graphics demanded a higher level of graph comprehension. This agrees with the previous work by Shah et al. (1999) in that larger spatial transformations require a higher degree of comprehension. However it is interesting to note that in this study, the need for increased levels of graph comprehension led to more critical and insightful comments. It seems that in some contexts, increasing the level of graph comprehension required to interpret a graph may fail to hinder the reader. Instead, it can provoke a higher level of understanding.

Additional studies have created measurement systems for graph comprehension, and in some instances these instruments include questions related to graphical distortion (Diezmann & Lowrie, 2009; Galesic & Retamero, 2010). However, little research has been done to compare
the level of comprehension participants can achieve on accurate graphs to the level of comprehension on deceptive graphs. Even less empirical research has been done to specifically determine what kinds of graphical distortions are most troublesome to viewers. The research that explores graphical distortions in context is sparse.

Summary

Graphical distortions are ever present in today’s society, and the ability to detect such distortions is of utmost importance for the general public. Given the focus on graph comprehension by key groups in mathematics education, we would expect students entering high school to be able to discern between accurate and biased graphs. Yet, there has been little research to date that explores this topic.

A large number of studies have been able to identify many variables that effect graph comprehension. These studies tend to focus on the theoretical implications, rather than the real world applications of their results. On the other hand, there have been many books and articles written about the dangers of biased graphs, with little empirical evidence to support them. Thus, this study will draw from both past theoretical research and from the ideas of various commentators in an attempt to decide if high school freshmen are able to distinguish between accurate and biased graphs. In other words: do students maintain a high level of graphical comprehension to answer comprehension questions about bar graphs with zero-baselines and non-zero-baselines with the same level of accuracy? This study will investigate.
Method

Note about Bar Graphs and Histograms

*Bar graphs* are diagrams in which the numerical values of variables are represented by the height or length of lines or rectangles of equal width. *Histograms* are sometimes considered a specialized type of bar graph where the dependent variable is numerical. For the remainder of this study, the term bar graph will be used to refer to both bar graphs and histograms.

Survey Instrument

The survey instrument was designed to measure participants’ levels of graphical comprehension while viewing graphs with and without distortions. The paper survey was created by the researcher and comprised Form A, B or C, along with Forms D and E. Each of these forms had a specific purpose and format.

Forms A, B, and C are different versions of the same form. They each contain four bar graphs that display information about the number of books in a fictional library over a span of six years. These graphs contained varying levels of distortions. After each graph is a set of reading beyond the data level comprehension questions. The purpose of these questions is to determine the participants’ level of graphical comprehension while viewing graphs with varying levels of distortions. Participants were randomly given either given Form A, B or C.

Form D contains a bar graph and six comprehension questions that vary in level. This form was adapted from an instrument used in Curcio’s initial graph comprehension study (1981b, p. 134). It is intended to be used as a point of comparison between an established survey instrument and the one created for this study. Form E contains demographic questions, for determining background information about the sample population. Every participant was given Form D and E.
The librarian at the Greenville High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

1. Notice that the number of books in the Greenville Library is lower in the 2007 school year than in the other school years. Circle the most reasonable explanation for this drop.
   a. The library lost a small amount of books after a book sale.
   b. The library lost more than half of its books after a flood.
   c. The library lost all of its books after a fire.
   d. The librarian forgot to record the number of books.

2. How confident are you that your answer is correct?

**Figure 2.** Comprehension item 1 and confidence item 2. Each item appears in Form A, B, and C.

**Figure 3.** A sample representative graph. This graph corresponds to items 1 and 2 and appears in either Form A, B or C.

**Figure 4.** The moderately distorted version of the graph in Figure 3.

**Figure 5.** The heavily distorted version of the graph in Figure 3.
Selected Response Comprehension Items in Forms A, B, and C. The purpose of Forms A, B, and C were to determine the level of graph comprehension of the participant while viewing graphs with and without distortions. Each of the odd numbered questions required students to use a bar graph to answer multiple choice graph comprehension questions. There
were four graphs and six questions in total, and all of the questions were at the reading beyond the data level. Each graph was created using a unique data set.

**Qualitative and Quantitative Items.** Two different types of questions were used to assess each participant’s level of graph comprehension. *Qualitative items* ask participants to make qualitative inferences about the data represented by the graph. On the survey, participants were asked to provide a reasonable real-world explanation for the increase or decrease in the data. Figure 2 contains one of the two qualitative questions (Questions 1 and 3 on Form A, B, or C, were both qualitative questions).

*Quantitative items* ask participants to make numerical predictions based on the data represented by the graph. In these questions, participants were asked to make extrapolations based on the linear trend of the data. Responses are always numerical. Figure 6 contains a quantitative question that requires participants to use the trend of the data determine when the number of books from a particular year will double. There were four quantitative questions (Items 5, 7, 9, and 11 on Form A, B, or C, were all quantitative questions.)

**Participant Confidence.** Each of the odd numbered comprehension items was followed by an even numbered item that asked for the participants’ confidence in answering the previous comprehension question. Participants responded on a five-point Likert scale, with responses ranging from “A: I guessed” to “E: I am sure” (see Figure 2). The purpose of these questions was to determine if participants felt less comfortable while viewing heavily distorted graphs. Unfortunately, the results from these confidence questions will not be discussed in this report.

**Three Levels of Graphical Distortions.** Three versions of each of the bar graphs were used during the survey: *representative, moderately distorted, and heavily distorted.* These three
versions were designed to determine whether participants correctly comprehended the data despite the graphical distortions, or if they were susceptible to graphical misperceptions.

Representative graphs used a baseline of zero. Figure 3 is an example of a representative graph. The second version used a baseline that ranged between 35% and 75% of the highest data point in each data set. All ranges in this section were constructed for the purposes of this study. These versions will be called moderately distorted graphs. Figure 4 is a moderately distorted graph. It has the baseline at 2400 book and the highest data point at 3250 books, for a baseline that is 74% of the highest data point. The third version used a baseline that ranged between 75% and 95% of the highest data point. These versions will be called the heavily distorted graphs. Figure 5 is a heavily distorted graph with a baseline that is 86% of the highest data point. Thus, the representative graphs contain no graphical distortions, the moderately distorted graphs contain moderate graphical distortions and the heavily distorted graphs contain the largest graphical distortions. Note that the terms representative, moderately distorted, and heavily distorted are specific to this study, and have not been adopted.

The percentages for moderately and heavily distorted graphs are large because the amount of visual distortion depends on range of the data as well as the baseline. Data that has a small range will require a higher baseline to achieve the same level of distortion that data with a larger range would. Thus, the percentages are large to accommodate for similar visual distortions in different data sets.

Distribution of Graphs within Forms A, B, and C. Recall that Forms A, B, and C each contained four graphs. These four graphs were created using unique data sets, meaning all four graphs depicted a different set of data. Three different versions of each of the graphs were made from each data set. For example, Figure 3 is the representative version of graph 1, Figure 4 is the
moderately distorted version of graph 1, and Figure 5 is the heavily distorted version of graph 1. There are three versions of graphs 2, 3, and 4, for a total of 12 graphs.

The twelve different graphs were distributed into Forms A, B and C. The distribution was done such that each form contained one of each graph: graph 1, graph 2, graph 3, and graph 4. In addition, each form contained at least one of every level of graphical distortion. For example, Form A contained the representative version of graph 1, the moderately distorted version of graph 2, the heavily distorted version of graph 3, and the representative version of graph 4. See Appendix C for a complete listing of the graphs present in each form.

**Scoring of Selected Response Items in Forms A, B, and C.** Each of the responses to the comprehension questions was scored correct or incorrect. A complete scoring guide is available in Appendix C.

An alternate scoring guide was also used to score responses. This alternate scoring guide scored each response correct, moderately incorrect, or heavily incorrect. The purpose of this guide was to establish multiple levels of understanding. This scoring guide will not be used in this study, and is included in Appendix C for use in future research.

**Constructed Response Prompts in Forms A, B, and C.** Forms A, B, and C also contained two constructed response questions which were more qualitative in nature. Participants were asked to respond to the following prompts: “If you could talk to the librarians who made the previous four graphs, what questions would you ask them? Why would you ask those questions?” and, “Did you find any of the previous four graphs confusing? Explain why or why not.” These were items 13 and 14, respectively.

The purpose of these prompts was to determine if a particular type of response correlated with performance on the selected response questions. For example, perhaps participants who
commented on the non-zero baseline performed better on the comprehension questions. Maybe participants that commented about confusions they had performed worse. In addition, if a large number of participants indicated that the survey was confusing, then the results would be less valid.

**Data About Validity in Form D.** Form D contained a bar graph that had been adapted from an instrument used in a major study (Curcio, 1981b, p. 134). The first two items ask viewers to read the data, next two items ask viewers to read between the data, and the last two items require participants to read beyond the data. Thus, the original objective of the instrument was to determine the level of graphical comprehension a participant can achieve: reading the data, reading between the data, or reading beyond the data.

The purpose of Curcio’s instrument in this survey is to determine construct validity. The results from the reading beyond the data items in Form D indicate each participant’s ability to answer high level graphical comprehension questions. From this information, the population can be divided into two groups: participants that have a high level of graph comprehension, and participants that do not have a lower level of graph comprehension. Then, the results from the items referring to representative graphs in Forms A, B, and C for both of the two groups can be compared.

The selection of reading beyond the data items from Form D and items referring to representative graphs in Forms A, B and C is purposeful. This section ensured that all of the comprehension questions used in the comparison was similar. Each of the selected items require a high level of graph comprehension and refer to graphs do not contain distortions. If Forms A, B and C measure the same construct as Form D, there should be a significant difference in scores
on Forms A, B, and C between the high level group and the low level group (groups are determined by Form D).

**Demographic Items in Form E.** Form E asks participants to report demographic information including gender, grade level, and title of current mathematics class. Participants are also asked what kinds of academic grades they typically receive in all of their classes and their math classes specifically. Lastly, participants were asked if there were any school years that they did not attend a school in the United States, in their current state, or their current school district. The results of these eight questions can be used to determine if there is a correlation between a participant’s level of graph comprehension and the factors described previously (gender, academic grades, years not in the current school district, etc.).

**Lack of Reliability Information.** Unfortunately, it is difficult to apply standard tests of reliability to the survey used in this study. Most measures of internal consistency do not apply. For example, a Cronbach's alpha is inappropriate because of the low number of questions, the high level of variability between questions, and the number of forms used.

Furthermore, the use of multiple forms with different questions is not useful for inter-method or parallel-forms reliability. Lastly, inter-rater reliability within the constructed response coding scheme was not tested. Recommendations for testing and improving reliability will be discussed later.

**Participants**

The researcher chose to survey ninth-grade students because of the sudden shift between the kindergarten through eighth-grade mathematics curriculum and the ninth-grade through twelfth-grade mathematics curriculum. NCTM, the Common Core State Standards, and the local curriculum all stipulate that students should have sufficient graphical comprehension skills at the
completion of eighth grade (Curcio, 2010; Common Core, 2010). Yet, if students are still susceptible to graphical misperceptions in ninth grade, there is still time to correct these issues.

**Participant Demographics.** The researcher surveyed 181 students at a public high school in an affluent New England suburb, yielding 159 usable surveys. Five regular mathematics teachers volunteered their freshman level classes as part of the study, for a total of 10 classes. Two of the classes were an honors geometry course, two of the classes were a middle level algebra course, five of the classes were a middle to low level algebra course, and one class was a remedial algebra course. More specifically, 38 students were in an honors geometry course, 44 were in a middle level algebra course, 67 were in a low level algebra course, and 10 were in a remedial algebra course.

Of the 159 students surveyed, 47.8% were male and 52.2% female. There were 3.7% of students who reported that they had attended a school outside of the United States for at least one academic year. In addition, 11.3% of students reported that they had attended a school outside of the state for at least one academic year. Lastly, 35.2% of students reported that they attended a school outside of the local district for at least one academic year.

**Local Demographics.** The following demographic information pertains to the entire local high school population, and is meant to serve as an indicator of educational need (Connecticut Dept. of Ed., 2010). Note that the group of survey participants is a subset of this larger population:

- 7% of students are eligible for free/reduced-price meals
- < 1% of students are not fluent in English
- 10% of students are identified as having a disability
- 7% of students are identified as gifted and talented
The following demographic information is meant to serve as an indicator of mathematics performance of the local high school as compared to other public high schools in the state:

- 94 percentile on the state standardized mathematics test
- 95 percentile on the mathematics portion of the SAT I

The following demographic information is meant to serve as an indicator of post-high school preparation of the local high school as compared to other public high schools in the state:

- 57% of juniors and seniors are enrolled in a course or courses for college credit
- 99% graduation rate
- 98% pursue higher education

As is evident, this is a very high-performing school. CAPT scores, SAT scores, college credit enrollment, and graduation rates are all exceptionally high. Note that this district was not chosen purposefully, but due to convenience caused by a pre-existing relationship with the district. Nevertheless, the high level of achievement within the school suggests that the students have undergone very rigorous curricula in all of their classes. If students in this district are unable to maintain a high level of comprehension while viewing graphs with distortions, then it is unlikely that a lower-achieving district would be able to maintain a high level of comprehension either.

**Local and State Curriculum.** The local district has stated that it follows a curriculum based off of recommendations from the Connecticut Department of Education. 64.8% of students that were surveyed had completed their entire kindergarten through eighth-grade education within the local district, and 88.7% had completed it within the state of Connecticut. It is reasonable to assume that the majority of students received instruction that was based off of the state curriculum.
It is important to note that students involved in this study should have received some instruction on graph comprehension. Whether this instruction actually occurred is unclear, but this study will assume that students have received some of the graph comprehension instruction required by the state. The state curriculum is currently in transition from the previous Connecticut Mathematics Curriculum Framework and Connecticut Mathematics Curriculum Standards to the Common Core State Standards. The Connecticut Mathematics Curriculum Framework and Connecticut Mathematics Curriculum Standards both emphasize the domains of Geometry and Measurement, and Working with Data in their elementary and middle school curriculums (Connecticut Dept. of Education, 2005; 2007). The Common Core State Standards also recognizes the domains of Measurement & Data and Statistics & Probability (Common Core, 2010). Although the state curriculum is in transition, it is clear that both the current and future standards support the instruction of measurement, data analysis, and statistical understanding.

Data Collection

The researcher initially introduced the study to 181 students two weeks before the date of the survey. The introduction took place during the first five minutes of their regularly scheduled mathematics class. At that time, students received an informational letter as well as a notification of refusal form.

On the date of the survey, the researcher reintroduced the study, and read an information sheet that instructed students not to write their names or any other identifying marks on the surveys. This was done to protect anonymity. The students also informed that they did not need to answer any question, and that they could stop the survey at any time. The survey was then distributed, completed, and collected as it was finished. Note that the three forms (Form A, Form
B, and Form C) were distributed at random. The survey took students approximately 15 minutes to complete. All surveys were completed in the students’ regularly scheduled mathematics classes under the supervision of the researcher and the regular classroom teacher.

All paper surveys were kept in a secure location until the data were entered into an electronic format.

**Data Analysis**

A total of 181 surveys were collected. Sixty one surveys contained Form A, sixty contained Form B, and sixty contained Form C. Fourteen surveys were removed because the participant was not in ninth-grade. Six surveys were removed due to nonresponse, and an additional two surveys were removed because the participant circled every possible response. This yielded 159 usable surveys. 55 surveys contained Form A, 54 contained Form B, and 50 contained Form C.

Surveys were initially inspected for names or other identifying marks; all marks were blacked-out. Then, each survey was coded with a number, 1 through 181. Responses to demographic and comprehension questions were inputted into a spreadsheet manually and scored electronically. The unusable surveys were removed as they were inputted. Constructed responses were transcribed electronically. All further analysis was also completed electronically.
**Coding of Constructed Response Items on Forms A, B, and C.** The responses to the constructed response items were reviewed and coded. Note that some responses may have received more than one code, while others may have received none. Recall the prompt for item 13, “If you could talk to the librarians who made the previous four graphs, what questions would you ask them? Why would you ask those questions?” Item 13 received 135 responses, while item 14 received 143 responses. The possible codes and sample responses for item 13 are as follows:

Table 1

*Codes for Constructed Response Item 13: Questions for the Authors of the Graphs*

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reason for Trends or Changes in Number of Books</td>
<td>The response asked for or suggested a reason for the change in the number of books.</td>
<td>What happened to the books? Why did the books drop, increase, or stay the same? Were there natural disasters? How busy is the library?</td>
</tr>
<tr>
<td>Accuracy of Count</td>
<td>The response questioned the accuracy of the graph or the data collection methods.</td>
<td>How did the authors count? Is their count accurate? Did they forget, misplace, or lose books?</td>
</tr>
<tr>
<td>Motivation to Count and Create Graphs</td>
<td>The response questioned the reason the authors created the graph.</td>
<td>Why did they count the books?</td>
</tr>
<tr>
<td>Non-Zero Baseline or Scale</td>
<td>The response comments on the baseline, y-axis or scale of the bar graphs.</td>
<td>Why didn’t the graph start at zero? The baseline was wrong.</td>
</tr>
<tr>
<td>Choice of Book Stacks to Represent Bars</td>
<td>The response comments on the bars of the bar graphs.</td>
<td>Why did you use books instead of rectangles? The stacks of books were confusing.</td>
</tr>
<tr>
<td>General Confusion</td>
<td>The response indicates confusion, but does not cite a reason.</td>
<td>I was confused.</td>
</tr>
<tr>
<td>No Questions</td>
<td>The response indicates that the participant has no questions. Does not include nonresponse.</td>
<td>I have no questions</td>
</tr>
</tbody>
</table>
Recall the prompt for item 14, “Did you find any of the previous four graphs confusing? Explain why or why not.” The possible codes and sample responses for item 14 are as follows:

Table 2

_Codes for Constructed Response Item 14: Was the Graph Confusing?_

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reason for Trends or Changes in Number of Books</td>
<td>The response asked for or suggested a reason for the change in the number of books.</td>
<td>What happened to the books? Why did the books drop, increase, or stay the same? Were there natural disasters? How busy is the library?</td>
</tr>
<tr>
<td>Non-Zero Baseline or Scale</td>
<td>The response comments on the baseline, y-axis or scale of the bar graphs.</td>
<td>Why didn’t the graph start at zero? The baseline was wrong.</td>
</tr>
<tr>
<td>Choice of Book Stacks to Represent Bars</td>
<td>The response comments on the bars of the bar graphs.</td>
<td>Why did you use books instead of rectangles? The stacks of books were confusing.</td>
</tr>
<tr>
<td>Wording of Questions</td>
<td>The response indicated that the participant was confused about the diction or syntax of the selected response items.</td>
<td>I was confused about what number 5 was asking.</td>
</tr>
<tr>
<td>Participants Inability to Complete Mathematical Calculations</td>
<td>The response indicated that the participant was unable to complete mathematical calculations required to respond to any of the items.</td>
<td>I didn’t know how to count the number of books in number 9.</td>
</tr>
<tr>
<td>No Confusion</td>
<td>The response indicates that the participant has no confusions. Does not include nonresponse.</td>
<td>I am not confused.</td>
</tr>
</tbody>
</table>

Nonresponse was also coded accordingly for both items.
Tests for Construct Validity. It is necessary to check that the selected response questions in Forms A, B, and C accurately measure the subject’s level of graph comprehension while viewing reading beyond the data questions. Recall that the results of the reading beyond the data items from Form D were compared to items referring to representative graphs in Forms A, B and C.

Each participant received two numerical scores for the purposes of determining construct validity. The first was a proportion of correct responses for items corresponding to reading beyond the data graphs in Form D. The second was a proportion of correct responses for items corresponding to accurate graphs in Forms A, B, and C. For this test, all responses were scored correct or incorrect. All scores were represented as proportions because participants answered a different number of questions that related to representative graphs on Forms A, B and C.

Then, participants were divided into two groups: those with perfect scores and those with imperfect scores on reading beyond the data questions in Form D. It was assumed that students with perfect scores had a high level of graph comprehension, while those with imperfect scores did not. A t-test was performed to compare the scores on Form A, B, and C between the group with a high level of graph comprehension and the group with a lower level of graph comprehension. The results of this test would help determine the construct validity of Forms A, B, and C for graphical comprehension.
Results and Discussion

Once the survey responses had been put into aggregate form, a few trends became visible. The most obvious trend was that students performed better while viewing representative graphs and poorer while viewing distorted graphs. Upon further analysis, more subtle patterns appeared in the data. For example, students scored lower on qualitative items than quantitative items. These findings will be now be presented and discussed in detail.

Table 3
Proportions of Correct Responses for Each Item in Forms A, B, and C

<table>
<thead>
<tr>
<th>Type of Graph Viewed</th>
<th>Item Number</th>
<th>Qualitative Items</th>
<th>Quantitative Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Representative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>55</td>
<td>50</td>
<td>54</td>
</tr>
<tr>
<td>Correct Responses</td>
<td>.800</td>
<td>.691</td>
<td>.745</td>
</tr>
<tr>
<td>Moderately Distorted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>54</td>
<td>55</td>
<td>50</td>
</tr>
<tr>
<td>Correct Responses</td>
<td>.273</td>
<td>.127</td>
<td>.582</td>
</tr>
<tr>
<td>Heavily Distorted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>50</td>
<td>54</td>
<td>55</td>
</tr>
<tr>
<td>Correct Responses</td>
<td>.400</td>
<td>.255</td>
<td>.418</td>
</tr>
</tbody>
</table>

Note. Qualitative items require participants to provide an explanation for the trends in the data. Quantitative items require participants to find a numerical extrapolation based on the trends in the data.

Varying Scores across Three Levels of Distortion

The most salient trend in the data is the prevalence of correct scores among participants viewing representative graphs. Table 3 reveals that the percentage of correct responses falls
between 65.5% and 89.1% for items referring to representative graphs. The researcher expected students to score at about this range, due to the assumption that most students had completed a curriculum that included graphical comprehension skills. This result indicates that most students are able to read and interpret the accurate graphs used in the survey.

Another noticeable trend is that the scores for items referring to a moderately or heavily distorted graph are low. Table 3 indicates that the percentage of correct responses drops to between 12.7% and 65.5% for items referring to moderately or heavily distorted graphs, with an average score of 42.3%. Moreover, the average score for each item was lower when it referred to a distorted graph than when it referred to a representative graph. This suggests that participants misperceive the data when it is represented by a distorted graph.

Taken together, these results can start to answer the main research question. It seems that students are generally not able to maintain a high level of graph comprehension while viewing bar graphs that contain distortions. But, we cannot stop here. Further analysis and discussion will reveal complexities, insights, and a few surprises about the comprehension of distorted graphs.

**Some Heavy Distortions may be too Obvious.**

The researcher expected average scores to decrease as the level of distortion increased. The conjecture was that heavier distortions cause heavier misperceptions. This expectation holds true for the quantitative items. The average score for each item was lower when it referred to a heavily distorted graph than when it referred to a moderately distorted graph (see Table 3). The same cannot be said for the qualitative items.

The scores on the qualitative items were higher when they referred to a heavily distorted graph than when they referred to a moderately distorted graph. Students who viewed moderately distorted graphs scored 27.3% and 12.7% correct on qualitative items 1 and 2, respectively,
whereas students who viewed heavily distorted graphs scored 40% and 25% correct (see Table 3). One reason for this unexpected difference may have been the stark contrast caused by the absence of bars. The heavily distorted, qualitative graphs had data points that fell below the baseline. As a result, there were some bars did not appear in the graph area at all. Participants may have noticed the absence of a bar and begun to suspect that the graph was odd in some way. They may have inspected the graph more closely and discovered the non-zero baseline. It seems that while moderate graphical distortions tend to create misperceptions, grossly large distortions can “give away” the distortion. This may ultimately help some readers realize and correct their own misperceptions.

**Difference between Scores on Qualitative and Quantitative Items.**

A particularly interesting result of this survey was the difference in scores between qualitative items and quantitative items. Recall that qualitative items require participants to provide an explanation for the trends in the data. Quantitative items require participants to find a numerical extrapolation based on the trends in the data.

There was no significant difference between the scores of these qualitative and quantitative items while viewing the representative graphs. However, there was a difference while viewing moderately and heavily distorted graphs. With a distortion present, participants were able to answer 26.0% of the qualitative items correctly \((n = 213)\) and 50.3% of the quantitative items correctly \((n = 418)\). Thus, the population scored higher on the quantitative items than the qualitative items.

This difference was unexpected, and the reasons for the result are a bit unclear. The quantitative items required linear extrapolation while the qualitative items required basic forms of proportional reasoning. Extrapolation is considered to be a more advanced skill than basic
forms of proportional reasoning because it is introduced later in the kindergarten through eighth-grade curriculum (Common Core, 2010; Connecticut Dept. of Ed., 2005; 2007). The researcher hypothesized that items requiring more advanced mathematical skills would be more difficult, but the results support the opposite conclusion. Perhaps the wording of the quantitative items prompted students to use comprehension strategies that they have learned in their mathematics classes. Mathematically precise words such as “trend,” “increase,” “decrease”, “double”, and “half” may have caused students to use the method of common differences to establish a linear trend. Consequently, students were able to predict the proper year without needing to use information related to the baseline shown for the graph.

On the other hand, the qualitative items contained no mathematic-specific vocabulary. The questions simply required students to make real-world interpretations of the graph; advanced strategies were not necessary. Students only needed to attend to the visual trends in the bars of the graph without referencing the scale. This was not enough for most students to realize the non-zero baseline. As a result, students frequently responded to these questions incorrectly.

One plausible explanation then is that the majority of students were able to use advanced graph comprehension strategies effectively. When these strategies were not used, or when they did not apply, students tended to misperceive the data represented by the graph. Perhaps students were more familiar with types of graphs that require advanced comprehension strategies (e.g., scatter plots, histograms) and less familiar with graphs that require basic interpretation skills (e.g., bar graphs, pie charts). It may be the case that students receive sufficient instruction in advanced quantities comprehension skills, but not enough in basic qualitative skills. While students could have received this sort of basic qualitative instruction at some point in their school
career, it may not have been reinforced recently. While these reasons may account for the trend in the data, more data is needed to support these conjectures.

Table 4

*Proportions of Correct Responses Across Course Levels*

<table>
<thead>
<tr>
<th>Course Description</th>
<th>n</th>
<th>Qualitative</th>
<th>Quantitative</th>
<th>Form D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reading</td>
<td>Between</td>
<td>Beyond</td>
</tr>
<tr>
<td>Honors Geometry</td>
<td>38*</td>
<td>0.487</td>
<td>0.796</td>
<td>1.000</td>
</tr>
<tr>
<td>High Level Algebra</td>
<td>44</td>
<td>0.420</td>
<td>0.676</td>
<td>0.966</td>
</tr>
<tr>
<td>Middle Level Algebra</td>
<td>67*</td>
<td>0.448</td>
<td>0.530</td>
<td>0.933</td>
</tr>
<tr>
<td>Remedial Algebra</td>
<td>10*</td>
<td>0.333</td>
<td>0.250</td>
<td>0.833</td>
</tr>
</tbody>
</table>

*Scores were omitted if the participant failed to respond to all questions within the same form.

Reading is an abbreviation of reading beyond the data, Between is an abbreviation of reading between the data, and Beyond is an abbreviation of reading beyond the data.*

**Scores across Four Class Levels**

In order to examine the trends between students with different mathematical abilities, the scores were analyzed by the mathematics courses the participants were currently taking. The course a participant was in can serve as a good indicator of their mathematical ability, since students were placed in these courses according to their grades in past courses.

**Relationship between Course Level and Graph Comprehension.** The most prominent and expected trend is the decrease in scores from students in higher level mathematics courses to students in lower level mathematics courses (see Table 4). This decrease occurred in every item and in every applicable form with one exception. There was a slight increase in scores from High Level Algebra students and Middle Level Algebra on qualitative questions. This increase is minimal, and can be considered an anomaly.
A second expected trend was the increase in scores between the three levels of questions in Form D. The average score in every class increased as students answered higher level comprehension questions. Curcio (1981b, 2010) suggests that the latter types of questions require a higher level of graph comprehension, and are more difficult for most participants to answer. These results indicate that students in higher level courses have a higher level of graph comprehension than those in lower level courses.

**Disparity between Qualitative and Quantitative Scores by Course Level.** It is critical to understand that the skills required to answer the qualitative and quantitative are not distinct. Students who are in higher level courses should have a higher level of mathematical achievement, and consequently, one would expect them to perform better on both qualitative and quantitative items. This relationship is supported by the evidence discussed in the previous section. The researchers expected that the difference in scores between qualitative and quantitative items would remain relatively constant from one class to another. Yet, the disparity between scores on the two item types varies heavily from higher level courses to lower level courses, and requires further investigation.

Average scores from students in the Geometry course were 31% percentage points higher on quantitative items than on qualitative items. Scores from High Level Algebra were 26% higher, scores from Middle Level Algebra were 8% higher, and scores from Remedial Algebra were 8% lower (See Table 4). Students in higher level mathematics courses seemed to show more of an increase between qualitative and quantitative items. Students in lower level courses show less of an increase, or even a decrease, in their scores.

This observation enhances the previous discussion that suggested that students have sufficient graph-comprehension strategies to answer quantitative questions, but are less skilled
at answering basic questions about graphs with distortions. The results indicate that students in upper level courses have a greater understanding of these strategies than students in lower level courses do. Perhaps students in upper level courses received a curriculum that was more rigorous in data analysis strategies than students in lower levels. Whatever the case may be, it is apparent the disparity between quantitative and qualitative abilities is greatest within the upper level courses.
Table 5  
*Prevalence of Specific Codes within Constructed Responses: Forms A, B, C*

<table>
<thead>
<tr>
<th>Code or Attribute</th>
<th>Number of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Item 13: Questions for the Librarians (Authors of the Graphs)</strong></td>
<td></td>
</tr>
<tr>
<td>Reason for Trends or Changes in Number of Books</td>
<td>77</td>
</tr>
<tr>
<td>Accuracy of Count</td>
<td>39</td>
</tr>
<tr>
<td>Motivation to Count and Create Graphs</td>
<td>11</td>
</tr>
<tr>
<td>Non-Zero Baseline or Scale</td>
<td>6</td>
</tr>
<tr>
<td>Choice of Book Stacks to Represent Bars</td>
<td>9</td>
</tr>
<tr>
<td>No Questions</td>
<td>11</td>
</tr>
<tr>
<td>Nonresponse</td>
<td>24</td>
</tr>
<tr>
<td><strong>Item 14: Confusion about Graphs</strong></td>
<td></td>
</tr>
<tr>
<td>General Confusion</td>
<td>4</td>
</tr>
<tr>
<td>Reason for Trends or Changes in Number of Books</td>
<td>31</td>
</tr>
<tr>
<td>Non-Zero Baseline or Scale</td>
<td>12</td>
</tr>
<tr>
<td>Choice of Book Stacks to Represent Bars</td>
<td>13</td>
</tr>
<tr>
<td>Wording of Questions</td>
<td>13</td>
</tr>
<tr>
<td>Participants Inability to Complete Mathematical Calculations</td>
<td>7</td>
</tr>
<tr>
<td>No Confusions</td>
<td>67</td>
</tr>
<tr>
<td>Nonresponse</td>
<td>16</td>
</tr>
</tbody>
</table>

**Trends in Constructed Responses**

The constructed responses provide some insight into the thoughts of the students as they answered each of the comprehension questions. From their responses, it is evident that the students are thinking critically about the reason for the trends they see in each bar graph. However, the responses do not indicate that many students noticed the non-zero baseline.
These responses will also determine if students were confused while they took the survey. Perhaps one or more of the graphs, prompts, or responses created unnecessary confusion and will render the results invalid. Thorough analysis and discussion is required.

**Participants want a Reason for Trends.** Recall that the constructed response questions asked participants for questions they would ask the authors of the graphs, and if they had any confusions about the graphs. When asked, 89 unique participants responded to these items with questions about the trends in the previous set of graphs. There were 77 participants who responded in this manner to item 13 and 31 who responded in item 14. Some participants mentioned trends in both responses. Below is a sample of typical student responses:

- “why don't they stay the same? why do they have to increase or decrease?”
- “Why did they drop so many books and then gain a lot back?”
- “why are the rate of the book being rased and dropping. is the library popular these days because of the people taking out the books?”
- “what happened to all the books? why didn't it decrease at times because I am interested.”

The frequency of the “trend” code is interesting, and it suggests a number of possibilities. First, students may have received formal instruction on what kinds of questions to ask the author when reading a graph. Students may have also been asked these kinds of questions by their teachers and peers when they learned to construct their own graphs. Still, the lack of contextual information in the selected response items may have motivated students to want to discover what the contextual information might be. Whatever the case may be, it is apparent that most students are interested in finding out more about the cause of the trends present in the graphs.
Reporting the Non-Zero Baseline. In total, only fifteen unique participants responded that some graphs contained non-zero baselines or an abnormal scale. Yet, it is presumptuous to assume that only 15 participants noticed the scale; many more could have noticed the baseline without noting it in their response.

In addition, there is no significant difference in scores between participants who reported the non-zero baseline and those who did not. An independent-sample t-test was conducted to compare scores from Forms A, B, and C between participants who noted the baseline and participants who did not. There was not a significant difference in the scores between participants who noted the baseline ($n = 15, M = 3.27, SD = 1.22$) and those who did not ($n = 144, M = 2.99, SD = 1.40$); $t(18) = 0.81, p = 0.427$. Thus, there is no apparent relationship between reporting the non-zero baseline and answering the comprehension questions correctly. Without further questioning, no conclusions can be made about the subset of participants who noticed the non-zero baseline.

Confusions about Graphs. Only 72 out of 159 unique participants noted any sort of confusion on the graphs. Recall that the objective of this study was to assess for graphical misperceptions; misperceptions are often synonymous with confusions. Some of the confusions that were reported were intentional byproducts of graphical misperception. Of those that reported confusion, only 17 out of 159 participants noted confusions that would have caused unintentional results (general confusion, wording of questions). This is an acceptably small number to conclude that most participants understood each of the items on Forms A, B, and C, and that there were minimal unintentional confusions.
Table 6

t-Test of Scores on Representative Graphs, between Participants with Perfect and Imperfect Scores on Form D, Reading Beyond the Data Items

<table>
<thead>
<tr>
<th>Participant Subgroup</th>
<th>Score on Representative Graphs from Forms A, B, C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
</tr>
<tr>
<td>Form D, Reading Beyond the Data</td>
<td></td>
</tr>
<tr>
<td>Participants with Perfect Score</td>
<td>115</td>
</tr>
<tr>
<td>Participants with Imperfect Score</td>
<td>44</td>
</tr>
</tbody>
</table>

*Note. t-Test of Difference = 0 (versus ≠ 0); Difference = (Perfect Score) - (Imperfect Score)

Estimate for difference: 0.200; $t = 2.69$; $p = 0.009$*; $DF = 61$

* $p < .001$.

Construct Validity

It is important to confirm that the survey items collectively were a valid measure of graphical comprehension. In order to measure the construct validity of Forms A, B, and C, the scores of the items corresponding to representative graphs in Forms A, B, and C were compared to the scores of the reading beyond the data items in Form D. Note that all scores in this comparison correspond to representative, high level graph comprehension questions. An independent-samples t-test was conducted to compare scores from Forms A, B, and C between participants who obtained a perfect score on Form D and participants who obtained an imperfect score on Form D. Although the sample is not normally distributed, the large sample size makes a t-test appropriate. The results indicate that there was a significant difference between participants who obtained a perfect score on Form D and those who did not. The specific results are included in Table 6. There is strong evidence to suggest that students who performed well on Curcio’s
reading beyond the data comprehension questions scored higher than those who did not on representative questions in Forms A, B, and C.

This result indicates a high level of construct validity for Forms A, B, and C. Forms A, B, and C were designed to assess a participant’s level of graphical comprehension. Form D was adapted from a study that was used to measure participant’s level of graph comprehension. The t-test indicates that students who scored highly on the adaptation of Curcio’s survey scored higher than their imperfect scoring peers on Forms A, B and C. If we assume that Curcio’s survey is a valid assessment of graphical comprehension, then the representative portion of this survey is also a valid assessment of graphical comprehension.
Conclusions and Implications

Expected Conclusions

As was expected, students generally scored lower when answering comprehension questions while viewing graphs that contained distortions. This result indicates that students are generally susceptible to misperceptions about non-zero baselines in bar graphs. Of course, the change in baseline must not cause data points to disappear completely, as it did with some of the heavily distorted graphs. This sort of drastic change seemed to potentially alert the reader to the graphical distortions.

Additionally, students in higher level courses tended to score higher than students in lower level courses. It is unclear whether this is due to differing student abilities before leveling occurs, or if it is a result of different instruction between the levels. Students are divided into leveled classes starting in sixth-grade. They continue through eighth grade, and into high school. It is not clear if three years of leveled classes has an effect on graph comprehension instruction. In any case, the disparity in scores is an unfortunate but expected result of a leveled mathematics curriculum.

Conclusions about the Rigor of Graphical Comprehension in Math Education

The difference between students’ scores on the qualitative and quantitative items on the survey may be a reflection of the manner in which graph comprehension skills are taught. Recall that students performed better on items involving quantitative predictions, and worse on items involving qualitative inferences. Furthermore, students in upper level courses scored drastically better on the quantitative items, while students in lower level courses scored similarly on both types of items.
Remember that the sample population was a subset of a school population that scored in the 94 percentile on the state standardized mathematics test and 95 percentile on the mathematics portion of the SAT I. In addition, 57% of juniors and seniors are enrolled in a course or courses for college credit, 99% of students graduate, and 98% pursue higher education (Connecticut Dept. of Ed., 2010). It can be assumed that the local mathematics curriculum prepares students for the kinds of problems they could see on a standardized test, in college, or while working in a STEM field. To that end, students in the local district probably learn a variety of skills, strategies, and techniques for a variety of topics, including graph comprehension.

Yet, the evidence suggests that students are more able to make qualitative predications and less able to make simple, qualitative judgments based on the data in a distorted graph. This difference may be caused by the way graph comprehension is taught. Students probably learn a very specific set of comprehension techniques that they then practice through repetition. Techniques such as extrapolation, interpolation, finding measures of central tendency, and finding measures of variation can be taught in a very straightforward, mechanized fashion. Based on the results from this study and the local demographic information, it seems that the mathematics curriculum is doing more than enough to teach students how to answer these kinds of questions.

The results from this study imply that greater emphasis should be placed on informal, qualitative analysis of graphs. Students may need to be taught about graphical distortions and given suggestions on how to avoid graphical misperceptions without resorting to quantitative techniques. This instruction may already exist in the lower grades, but it should continue throughout elementary school, and into high school. Without this type of instruction, we may be
preparing students to analyze and extrapolate complex trends in data without the common sense to see the apparent distortion in a bar graph.

**Suggestions for Further Research**

The researcher suggests that this study be repeated with slight alterations to increase the depth and expand the scope of the conclusions that can be made.

**Increase in Number of Items.** It is the suggestion of this researcher that this study be repeated with a revised survey form. Due to requests from the cooperating school, the final version of the survey was shorter than anticipated. As a result, the researcher was unable to measure the reliability or internal consistency of the survey instrument. This study should be repeated with a longer survey form. Such a survey would include all twelve graphs instead of a sampling of four. This survey would allow for an accurate measure of internal consistency, and include a measure of parallel-forms reliability.

**Refinement of Constructed Response and Addition of Interviews.** If survey is used again, the constructed response items should be refined. Recall that the constructed response items required students to report any questions for the authors and confusions they had. These two items were too broad, and did not target any specific understandings. Instead, participants commented on whatever understandings or misunderstandings were most obvious to them. As a result, the responses did not provide a valid indication of the students’ understandings, misconceptions, and confusions.

At least one question should be added to determine whether or not the student has noticed the non-zero baseline. Care must be taken in the phrasing of the question to ensure that the participant does not realize the distortions as a result of the question. Additional questions should also be added to target specific misunderstandings, for example, confusion about the book stacks,
questions about the trends, and misinterpretations of the questions. Perhaps participants could complete the comprehension items in one packet, and the aforementioned “misunderstanding” items in another. The first packet could be collected before the second is distributed to make sure participants do not change their answers on the comprehension items as a result of the misunderstanding items. Otherwise, participants’ scores may not be an accurate reflection of their perceptions of graphs.

Moreover, a repeat of this study should include follow up interviews for select participants. While the constructed response questions gave some insight, they did not provide the richness or depth required to make any conclusions. The refined constructed response questions could fail to produce the intended insight, but interview questions could be adjusted during the interview to ensure proper depth.

**Additional Sample Populations.** The study should be repeated within schools with different socioeconomic statuses. The sample used in this study was from an affluent suburban town. This is not a representative sample of students in Connecticut. Perhaps future research could focus on a school populations that are more representative of Connecticut’s demographics.

The study could also be repeated in a different state. Curricula differ from state to state, and it would be worthwhile to examine the effects of different curricula on graphical perceptions. Although there is no literature that compares the level of graphical comprehension instruction across states, there has been research into the effectiveness of schools at math and science instruction. White & Cottle (2011) developed and implemented *The Science and Engineering Readiness Index*, which ranks states according to scores on various math and science tests, as well as the level of teacher qualification within the state. Massachusetts, Minnesota, New Jersey, New Hampshire, and New York ranked the highest, while Mississippi, Alabama, Louisiana,
Nevada, Arizona, New Mexico, Oklahoma, Nebraska, and West Virginia ranked lowest. The researchers suggest that additional research occur within at least one high-ranking state and one low-ranking state. Differences or similarities in the results from these states could lead to suggestions about what each state needs to improve in terms of graphical comprehension instruction.

While further research should involve a variety of ages, this researcher suggests that an emphasis be placed on students in kindergarten through twelfth-grade. This age group receives the most instruction on graph comprehension, and any results or implications from further research would be best applied in these grades.

**Different Types of Distortions.** Additional studies should also investigate different types of distortions present in different kinds of graphs. There have been numerous studies completed that attempt to isolate the effects of specific graphical features on graphical comprehension, but there are very few studies that investigate the effect of graphical distortions holistically (Katz et al., 2004; Shah & Carpenter, 1995; Stewart et al., 2009). On the other hand, a number of commentators and theorists have derived conclusions about the effects of graphical distortions with little empirical evidence to support their claims (Huff, 1954; Jones, 1995; Tufte, 1983). More empirical data should be collected on the effects of graphical distortions on graph comprehension. It is the hope of this researcher that through studying the effects of graphical distortions on students, the national mathematics curriculum will place a greater emphasis on reading real-world graphs that may contain distortions.
References


### Appendix A: Table of Definitions

**Table 7**  
*Important Terms and Definitions*

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>The lowest value represented by the ordinate scale on a bar graph. It can also refer to the horizontal line on a graph corresponding to this value.</td>
</tr>
<tr>
<td>Bar Graph</td>
<td>A diagram in which the numerical values of variables are represented by the height or length of lines or rectangles of equal width.</td>
</tr>
<tr>
<td>Graphical Comprehension</td>
<td>The process of reading, understanding, and interpreting information from a graph.</td>
</tr>
<tr>
<td>Graphical Distortion</td>
<td>A visual feature of a graph that may cause the viewer to misinterpret the data being represented.</td>
</tr>
<tr>
<td>Graphical Misperception</td>
<td>The lowering of a viewer’s level of graphical comprehension in the presence of graphical distortions.</td>
</tr>
<tr>
<td>Heavily Distorted Graph</td>
<td>A bar graph with a baseline that is greater than 75% of the highest data point in the entire data set. This definition could be extended to additional types of graphs.</td>
</tr>
<tr>
<td>Moderately Distorted Graph</td>
<td>A bar graph with a baseline that ranges between 35% and 75% of the highest data point in the entire data set. This definition could be extended to additional types of graphs.</td>
</tr>
<tr>
<td>Reading the Data</td>
<td>The most elementary level of graphical comprehension. At this level, the reader extracts information about the data from a graph.</td>
</tr>
<tr>
<td>Reading Between the Data</td>
<td>An intermediate level of graphical comprehension. At this level, the reader finds relationships in the data through interpretation and integration.</td>
</tr>
<tr>
<td>Reading Beyond the Data</td>
<td>The most advanced level of graphical comprehension. At this level, the reader extends, predicts, or infers from the data to form conclusions. The reader may also integrate outside information to aid their ideas.</td>
</tr>
<tr>
<td>Representative Graph</td>
<td>A bar graph with no identifiable graphical skews. This definition could be extended to additional types of graphs.</td>
</tr>
<tr>
<td>Qualitative Items</td>
<td>Items that require participants to make inferences about the data represented by the graph. Specifically, participants were asked to provide a reasonable real-world explanation for the increase or decrease in the data.</td>
</tr>
<tr>
<td>Quantitative Items</td>
<td>Items that require participants to extrapolate in order to make predictions based on the data represented by the graph. Responses are always numerical.</td>
</tr>
</tbody>
</table>
Appendix B: Survey Instrument

FORM A

The librarian at the Greenville High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

1. Notice that the number of books in the Greenville Library is lower in the 2007 school year than in the other school years. Circle the most reasonable explanation for this drop.
   
a. The library lost a small amount of books after a book sale.
b. The library lost more than half of its books after a flood.
c. The library lost all of its books after a fire.
d. The librarian forgot to record the number of books.

2. How confident are you that your answer is correct?
   
The librarian at the Springfield High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

3. Notice that the number of books in the Springfield Library is higher in the 2007 school year than in the previous two years. Circle the most reasonable explanation for this increase.
   a. The library bought a small amount of books to fill a few bookshelves.
   b. The library doubled its number of books when added a new wing.
   c. The library was not open before the 2007 academic year, and it needed all new books.
   d. The librarian forgot to record the number of books in the 2005 and 2006 years.

4. How confident are you that your answer is correct?
The librarian at the Franklin High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

**Number of Books in the Franklin Library**

Notice that the number of books in the Franklin Library has increased every year. Assume that this trend continues even after the 2010 academic year.

5. In what academic year does the number of books from the 2005 year double?
   
   a. 2025  
   b. 2015  
   c. 2006

6. How confident are you that your answer is correct?
   
   a. I guessed  
   b. Not confident  
   c. Moderately confident  
   d. Very confident  
   e. I am sure

7. In what academic year does the number of books from the 2010 year double?
   
   a. 2035  
   b. 2025  
   c. 2016

8. How confident are you that your answer is correct?
   
   a. I guessed  
   b. Not confident  
   c. Moderately confident  
   d. Very confident  
   e. I am sure
The librarian at the Clinton High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

Notice that the number of books in the Clinton Library has decreased every year. Assume that this trend continues even after the 2010 academic year.

9. In what academic year will be half as many books as there are in the 2006 academic year?

10. How confident are you that your answer is correct?

11. In what academic year will there be zero books in the library?
    a. 2030  b. 2020  c. 2010

12. How confident are you that your answer is correct?
13. If you could talk to the librarians who made the previous four graphs, what questions would you ask them? Why would you ask those questions?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

14. Did you find any of the previous four graphs confusing? Explain why or why not.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
FORM B

The librarian at the Greenville High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

1. Notice that the number of books in the Greenville Library is lower in the 2007 school year than in the other school years. Circle the most reasonable explanation for this drop.

   a. The library lost a small amount of books after a book sale.
   b. The library lost more than half of its books after a flood.
   c. The library lost all of its books after a fire.
   d. The librarian forgot to record the number of books.

2. How confident are you that your answer is correct?

The librarian at the Springfield High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

![Number of Books in the Springfield Library](image)

3. Notice that the number of books in the Springfield Library is higher in the 2007 school year than in the previous two years. Circle the most reasonable explanation for this increase.

   a. The library bought a small amount of books to fill a few bookshelves.
   b. The library doubled its number of books when added a new wing.
   c. The library was not open before the 2007 academic year, and it needed all new books.
   d. The librarian forgot to record the number of books in the 2005 and 2006 years.
   e. I do not have enough information to answer this question.

4. How confident are you that your answer is correct?

The librarian at the Franklin High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

Notice that the number of books in the Franklin Library has increased every year. Assume that this trend continues even after the 2010 academic year.

5. In what academic year does the number of books from the 2005 year double?
   a. 2025  b. 2015  c. 2006

6. How confident are you that your answer is correct?

7. In what academic year does the number of books from the 2010 year double?
   a. 2035  b. 2025  c. 2016

8. How confident are you that your answer is correct?
The librarian at the Clinton High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

![Number of Books in the Clinton Library](image)

Notice that the number of books in the Clinton Library has decreased every year. Assume that this trend continues even after the 2010 academic year.

9. In what academic year will be half as many books as there are in the 2006 academic year?

10. How confident are you that your answer is correct?

11. In what academic year will there be zero books in the library?
    a. 2030  b. 2020  c. 2010

12. How confident are you that your answer is correct?
13. If you could talk to the librarians who made the previous four graphs, what questions would you ask them? Why would you ask those questions?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

14. Did you find any of the previous four graphs confusing? Explain why or why not.

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
FORM C

The librarian at the Greenville High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

![Bar Graph of Number of Books in the Greenville Library]

1. Notice that the number of books in the Greenville Library is lower in the 2007 school year than in the other school years. Circle the most reasonable explanation for this drop.

   a. The library lost a small amount of books after a book sale.
   b. The library lost more than half of its books after a flood.
   c. The library lost all of its books after a fire.
   d. The librarian forgot to record the number of books.
   e. I do not have enough information to answer this question.

2. How confident are you that your answer is correct?

The librarian at the Springfield High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

![Graph showing number of books in the Springfield Library over academic years 2005-2010.]

3. Notice that the number of books in the Springfield Library is higher in the 2007 school year than in the previous two years. Circle the most reasonable explanation for this increase.
   a. The library bought a small amount of books to fill a few bookshelves.
   b. The library doubled its number of books when added a new wing.
   c. The library was not open before the 2007 academic year, and it needed all new books.
   d. The librarian forgot to record the number of books in the 2005 and 2006 years.

4. How confident are you that your answer is correct?
The librarian at the Franklin High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

Notice that the number of books in the Franklin Library has increased every year. Assume that this trend continues even after the 2010 academic year.

5. In what academic year does the number of books from the 2005 year double?
   a. 2025   b. 2015   c. 2006

6. How confident are you that your answer is correct?

7. In what academic year does the number of books from the 2010 year double?
   a. 2035   b. 2025   c. 2016

8. How confident are you that your answers are correct?
The librarian at the Clinton High School Library counts the number of books in the library each year. The graph below displays the number of books for the past 5 years. Use the graph to answer the following questions.

Notice that the number of books in the Clinton Library has decreased every year. Assume that this trend continues even after the 2010 academic year.

9. In what academic year will be half as many books as there are in the 2006 academic year?

10. How confident are you that your answer is correct?

11. In what academic year will there be zero books in the library?
    a. 2030  b. 2020  c. 2010

12. How confident are you that your answers are correct?
13. If you could talk to the librarians who made the previous four graphs, what questions would you ask them? Why would you ask those questions?

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

14. Did you find any of the previous four graphs confusing? Explain why or why not.

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________
Use the graph above to answer the following questions.

1. How many children celebrate a birthday in February?
   a. 0  
   b. 1  
   c. 2  
   d. 4

4. How many children are in Mr. Kahn’s class?
   a. 10  
   b. 30  
   c. 44  
   d. 55

2. During which month are there eight children who celebrate a birthday?
   a. May  
   b. July  
   c. September  
   d. November

5. What is the probability that the birthday being celebrated in December occurs on December 25th?
   a. 1/31  
   b. 25/31  
   c. 30/31  
   d. 1

3. What does this graph tell you?
   a. There are more birthdays during June and November than during any other month of the year.
   b. There are more birthdays during May than during any other month of the year.
   c. There are the least number of birthdays during June and November then during any other month of the year.
   d. As the year progresses from January to December, the number of birthdays decreases.

6. Sally’s birthday is in February. According to the graph, which of the following statements is correct?
   a. Sally was probably not born on February 29th.
   b. Sally is not in Mr. Kahn’s class.
   c. Sally is in Mr. Kahn’s class.
   d. Sally is the only one celebrating a birthday in February.
FORM E

1. What gender do you identify with?
   a. Male
   b. Female

2. What year are you in high school?
   a. Freshman
   b. Sophomore
   c. Junior
   d. Senior

3. What math class are you currently in?
   a. Essentials for Algebra
   b. Algebra 1B – 1
   c. Algebra 1A
   d. Geometry A, L-1
   e. Other: __________

4. What kind of grades do you generally get in all of your classes?
   a. Mostly A’s
   b. Mostly B’s
   c. Mostly C’s
   d. Mostly D’s
   e. Mostly F’s

5. What kind of grades do you generally get your math classes?
   a. Mostly A’s
   b. Mostly B’s
   c. Mostly C’s
   d. Mostly D’s
   e. Mostly F’s

6. Were there any school years (or grades) that you did not attend in a United States School?

7. Were there any school years (or grades) that you did not attend in a Connecticut school?

8. Were there any school years (or grades) that you did not attend in a Glastonbury Public School?
### Appendix C: Scoring Guides

**Scoring Guide for Forms A, B, C**

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Qualitative Items</th>
<th>Quantitative Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Responses</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Responses</td>
<td>a</td>
<td>e</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Responses</td>
<td>e</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

**Note.** All non-correct responses are considered incorrect.

*Rep.* is an abbreviation of *representative graph*, *Mod.* is an abbreviation of *moderately distorted graph*, and *Heavy* is an abbreviation of *heavily distorted graph*.

**Scoring Guide for Form D**

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Reading the Data</th>
<th>Reading Between the Data</th>
<th>Reading Beyond the Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

| Correct Response | a | d | a | c | a | b |

**Note.** All non-correct responses are considered incorrect.
Alternate Scoring Guide

The following scoring guide creates a distinction between correct, moderately incorrect, or heavily incorrect responses in Forms A, B, and C. The distinction in responses was designed to correspond to perceptions that could have been caused by the three different versions of the graphs (representative, moderately distorted, and heavily distorted). The reason for this advanced scoring guide was to provide further insight into the types of responses students were giving. However, it was not relevant to the discussion, conclusions, or implications of this study and has been added in this appendix for use by future researchers.

For example, Figure 3 is a representative graph. In this graph, the correct response (choice a) was also the most plausible response based on a quick viewing. Figure 4 is a moderately distorted graph with the same comprehension question and responses as Figure 3. Like Figure 3, the correct response is also choice a. However, the graph in Figure 4 is distorted in such a way that the moderately incorrect response (choice b) would seem to be the most plausible to participants with graphical misperceptions. Figure 5 is a heavily distorted graph with the same comprehension question as the representative graph (Figure 3). Thus, participants who were susceptible to graphical misperceptions would have chosen a response that corresponded to the type of distortion present in the graph, instead of consistently choosing the correct answer. A complete scoring guide is available in Appendix C.
**Alternate Scoring Guide for Forms A, B, C**

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Qualitative Items</th>
<th>Quantitative Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 3</td>
<td>5 7 9 11</td>
</tr>
</tbody>
</table>

**Form A**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Responses</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>Moderately Incorrect Responses</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Heavily Incorrect Responses</td>
<td>c, d</td>
<td>c, d</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

**Form B**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Responses</td>
<td>a</td>
<td>e</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>Moderately Incorrect Responses</td>
<td>b</td>
<td>b, c, d</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Heavily Incorrect Responses</td>
<td>c, d</td>
<td>a</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

**Form C**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Responses</td>
<td>e</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>Moderately Incorrect Responses</td>
<td>b, c, d</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Heavily Incorrect Responses</td>
<td>a</td>
<td>c, d</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* *Rep.* is an abbreviation of representative graph, *Mod.* is an abbreviation of moderately distorted graph, and *Heavy* is an abbreviation of heavily distorted graph.
Also included in this appendix is an altered version of Table 3 that reflects the distinction between moderately incorrect and heavily incorrect responses.

*Alternate Version of Table 3*

<table>
<thead>
<tr>
<th>Type of Response</th>
<th>Item Number</th>
<th>Qualitative Items</th>
<th>Quantitative Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Participants Viewing Representative Graphs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>55</td>
<td>50</td>
<td>54</td>
</tr>
<tr>
<td>Correct</td>
<td>.800</td>
<td>.691</td>
<td>.745</td>
</tr>
<tr>
<td>Moderately Incorrect</td>
<td>.000</td>
<td>.091</td>
<td>.200</td>
</tr>
<tr>
<td>Heavily Incorrect</td>
<td>.200</td>
<td>.109</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>Participants Viewing Moderately Distorted Graphs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>54</td>
<td>55</td>
<td>50</td>
</tr>
<tr>
<td>Correct</td>
<td>.273</td>
<td>.127</td>
<td>.582</td>
</tr>
<tr>
<td>Moderately Incorrect</td>
<td>.436</td>
<td>.782</td>
<td>.291</td>
</tr>
<tr>
<td>Heavily Incorrect</td>
<td>.273</td>
<td>.091</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>Participants Viewing Heavily Distorted Graphs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>50</td>
<td>54</td>
<td>55</td>
</tr>
<tr>
<td>Correct</td>
<td>.400</td>
<td>.255</td>
<td>.418</td>
</tr>
<tr>
<td>Moderately Incorrect</td>
<td>.455</td>
<td>.691</td>
<td>.127</td>
</tr>
<tr>
<td>Heavily Incorrect</td>
<td>.055</td>
<td>.036</td>
<td>.436</td>
</tr>
</tbody>
</table>

*Note.* Qualitative items require participants to provide an explanation for the trends in the data. Quantitative items require participants to find a numerical extrapolation based on the trends in the data.