CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

EFFECTIVENESS OF INFOGRAPHIC DISPLAYS ON DATA INTERPRETATION OF GRE QUESTION-RELATED DATA

A graduate project submitted in partial fulfillment of the requirements

For the degree of Master of Arts in Psychology

Human Factors and Applied Psychology

by

Veronica Becerra

August 2015

The graduate project of Veronica Becerra is approved:

_____ _____

Jill Quilici, Ph.D. Date

_____ _____

Paula DiMarco, Ph. D. Date

_____ _____

Tyler Blake, Ph. D., Chair Date

California State University, Northridge

ACKNOWLEDGEMENT

Professor Blake…your guidance and "Glimmers of a Possible Future" by shaping "[Any of] that figuring result[ing] in an actual shape [?]" made this portion of *the dream* possible… like you said "…limiting your view to faux perceptions of self-competence is simply snatching defeat from the jaws of victory…" Thank you!
Professor Quilici… you transformed my life in 488 and I hope to "nudge" others in the way you have enlightened me…Thank you!
Professor DiMarco…you accepted this project and made it possible in record time and created the collaborative effort with the artists that furthered its survival…Thank you!
Professor Oh…you are the perfect combination of discipline and nurturing, I can fish…Thank you!

My Parents…you created me…
My Brother Daniel…you have continuously stretched my endurance…
My Sister Michele…there is NO WAY I could have done this (and so many other things) without you, you share your strength, your passion, your creativity, your curiosity and your tenderness with me and that drives my fire!
Lizette…my sister before my sister, I will "bake" and with your guidance know when I am ready…
Ali…you are the calm light that elucidates the logic…I am lucky to have you in my life…
Laura… my cheerleading captain, my sun…that was the best summer in my academic life…
Jen…my "ethereal baby sister", I don't "get" things without your "half"…
Dr Seiver…my O…life would not be without your "situational ethics"…I am blessed to have you in my life…
Dr Schillinger… your knowledge and caring combined have been an inspiration…
My JCWJ…may I make you proud and reflect your vision…
My LAW…Benji's father, my life, my compass…without you, lost I would be!
The rest of my cheerleaders…thank you for your support, your authenticity and for allowing me to be!

Table of Contents

List of Tables

List of Figures

Abstract

Effectiveness of Infographics Displays on Data Interpretation of GRE Question Related Data

By

Veronica Becerra

Master of Arts in Psychology,

Human Factors and Applied Psychology

Data interpretation is a skill that is increasingly important for society as more data become available for analysis in many of areas of daily life. Different methods and alternate presentation formats exist to address various types of data. Infographics are visual representations that integrate information and graphics in order to convey messages that assist in data interpretation. Literature on interpretation of infographics as currently used by industry is increasing, but much more is needed to optimize the design and application of infographics. Graduate Record Examinations standardized exam questions (derived from the data analysis section) were used to measure accuracy of interpretation. Twenty GRE questions were selected, ten were converted into infographic representations and ten were presented with their original graph display. Five questions in each display category (infographic representation and original graph) were either easy or medium level math difficulty. A total of 77 college students completed the questionnaire using a desktop with an onscreen calculator. The independent factors in this repeated-measures design were the type of graphic display and the level of math difficulty. Results showed that the participants benefited from the infographic display for the medium level difficulty and that males scored

significantly higher than females. Participants were not interviewed about their thought process when extracting the required infographic components needed to answer the GRE questions. Future studies can include collection of qualitative data that could further our understanding of infographic interpretation as well as addressing gender differences affecting the reading and interpreting of information presented graphically.

Introduction

Data producers and consumers in the information age are confronted with vast and growing amounts of data. Those who interface with these data face challenges in three specific areas: availability, accessibility and interpretation (National Academies Press [NAP], 1995). Programmers have developed algorithms to help the users of data in these task areas: sort, organize, access and display (Sagar, Hamid, Khalim 2012;). However, computational tools are not enough to overcome these challenges. Human expertise is necessary to achieve effective designs of data presentation, reliable sources of data, and comprehensive interpretations of data (Leonelli, 2014). Leonelli states that collaboration between a visualization expert and a subject expert is critical in order to create accurate representations of data. The NAP (1995) suggests that Human Factors research can provide direction in the evaluation of the effectiveness of different visualization representations relative to the complexity of the data they present. The goals for their suggested research are twofold: to determine the type of tasks that would significantly benefit from having the assistance of a visualization tool, and to evaluate the effectiveness of those visualization tools for improving user performance.

A review of Human Factors visualization research by Tory & Moller (2004) summarizes ways that visualization tools can assist cognition. These include visualization tools that function as a type of external memory, which in turn reduces demands on working memory (by grouping related information for easier search and recall) and assisting pre-attentive monitoring through visually recognizable visualizations (Tory & Moller, 2004). These visualizations are an important feature when interpreting data, particularly with multiple and complex data sets. According to Grolemund & Wickham (2014), data

interpretation based on multiple data sets requires complex cognitive processing that is analogous to understanding relationships among schemas and "sensemaking". The emerging field of Information Visualization (i.e., "InfoVis") seeks to integrate data, information and knowledge in order to facilitate effective mental representations (i.e., visualizations) of complex information.

Infographics are visual representations that integrate information derived from data and graphics to convey a message. These visualizations are frequently used to aid in data interpretation. Currently, design guidelines for infographics exist broadly under the umbrella of several disciplines that include semiotics and graphic design. The effectiveness of these guidelines is mostly anecdotal and has not been validated by applied research. Kosara & Mackinlay (2013) argue that research that would result in more concrete guidelines or best practices for data presentation and communication through visualization could provide more reliable guidance for infographic designers, and thus more effective information interpretations by users.

Related Work

Smiciklas (2012) formally defines infographics as "a visualization of data or ideas that tries to convey complex information to an audience in a manner that can be quickly consumed and easily understood" (p.3). He also states that "the process of developing and publishing infographics is called data visualization, information design, or information architecture" (p.3). Data visualization is a subset of information visualization that focuses on representing quantitative and categorical data (Friendly, 2009).

Infographics, along with other visualizations (e.g., maps, statistical graphs, scientific visualizations, way-finding tools, signage, information illustration and information visualization) are all components of the broad concept of *information design*.

Common bar graphs, line graphs and pie charts have existed as tools for data presentation since the 18th century. These formats can sometimes be found as elements that contribute to a larger image when used in infographic displays. Typically, in order to interpret an infographic message, it is necessary to read these graphs individually and to interpret the data they present.

Design for facilitating interpretation

According to the Handbook of Human Factors and Ergonomics (North, 2012), there are four methods that are used to create display design that support user decision-making and action: the aesthetics psychophysical, the attention-based, the problem-solving and the decision-making approaches. Commonly used parameters to represent data visually within the aesthetics psychophysical method include data-ink ratio and data density (Tufte, 1983, Salvendy, 2012). Note: for the purposes of this discussion, hard copy (e.g., paper) and soft copy (i.e., display screens) research will be integrated. Thus, terms used to describe aspects that are equivalent in terms of cognitive interpretation (e.g., pixel density for displays and ink density for paper) are used as appropriate to the media used in the cited research.

The data-ink ratio, as defined by Tufte (1983), is a measure of the proportion of ink used to visually represent data (only that which is needed) to the total ink used in the data visual representation; data-ink is, thus, the "non-erasable core of a graphic" (p.93). Data density is the number of data points on the graphic divided by the area of the graphic display.

Tufte concludes that higher density is the more effective option. These two parameters are alleged by Tufte to be an objective measure of effective data visualizations.

A study to measure the memorability of visualizations was conducted by Borkin et al. (2013), in an attempt to provide the groundwork for further studies relating to visualizations and comprehension. Borkin's team collected 5,693 online static data visualizations from the internet, categorizing them into four source categories. These visualizations were: infographics, scientific publications, news media and government/world organization. They proceeded to categorize each visualization as either single-panel or multiple-panel visualizations depending on the number of discrete graphic representations. Based on their comprehensive review of the static visualizations, the research team created specific taxonomy in order to describe, sort, and group the visual representations into defining characteristics. They identified a total of twelve visualization categories (i.e., types of graphs) and two characteristic categories (*properties* and *attributes*) that were then applied to each of the twelve visualization categories. Visual characteristic subtypes were assigned to the *properties* and *attributes* (see Table 1 and Figure 1).

Subtype Groups Assigned to Characteristic Categories: Properties & Attributes	
Properties	**Attributes**
1. Dimension [2D; 3D] 2. Multiplicity [Single; Multiple; Grouped; Multi-Panel; Small Multiples; Combination] 3. Pictorial [Pictorial; Pictorial Unit] 4. Time [Time Series]	1. Black & white: is the visual black & white 2. Color [1, 2-6, >=7]: 3. How many colors used in the visual display 4. Data-to-ink ratio [*good, medium, bad*]. - Good was defined as a subjective assessment that the ratio was low therefore more readable- thus a high ratio resulted in a *bad* rating. 5. Visual Density [low, medium, high] 6. Human Recognizable Objects [yes, no]: whether in a photograph or drawing 7. Human Depiction [yes, no]: whether in a photograph or drawing.

Table 1. Subtypes of Properties and Attributes Category

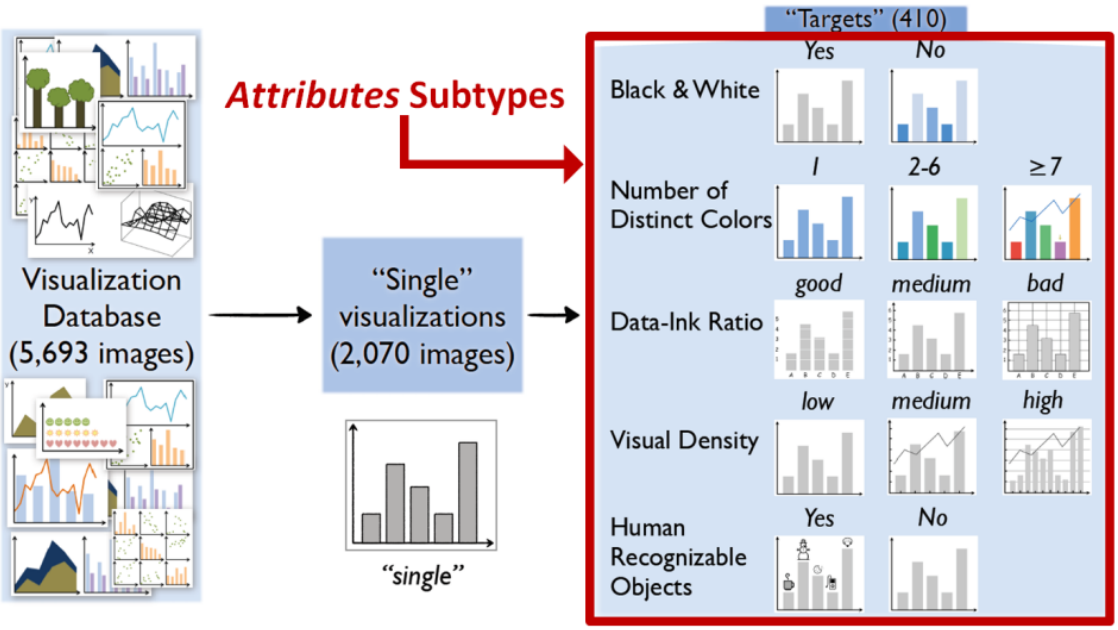


Figure 1. Examples of Attributes Subtypes (adapted from Borkin et al. 2014)

A total of 2,070 visualizations were categorized as single-pane from the original set of 5,693 visualizations. From the single-pane group, 410 visualizations were selected and assigned attributes in preparation for measuring memorability. Participants were presented with images that included the target visualizations among filler images using previous methodology developed by Isola et al (2011) that measured memorability. In this memory game, played on an online platform, participants would press a key upon viewing an image that had previously been displayed (Isola et al., 2011). The results from the Borkin et al. (2013) study provided evidence that visualizations with color, human recognizable objects, low data-ink ratio, high visual density, and pictorials (as opposed to common graphs) were more memorable (see Figure 2 selected from the Borkin et al. [2013] supplemental material).

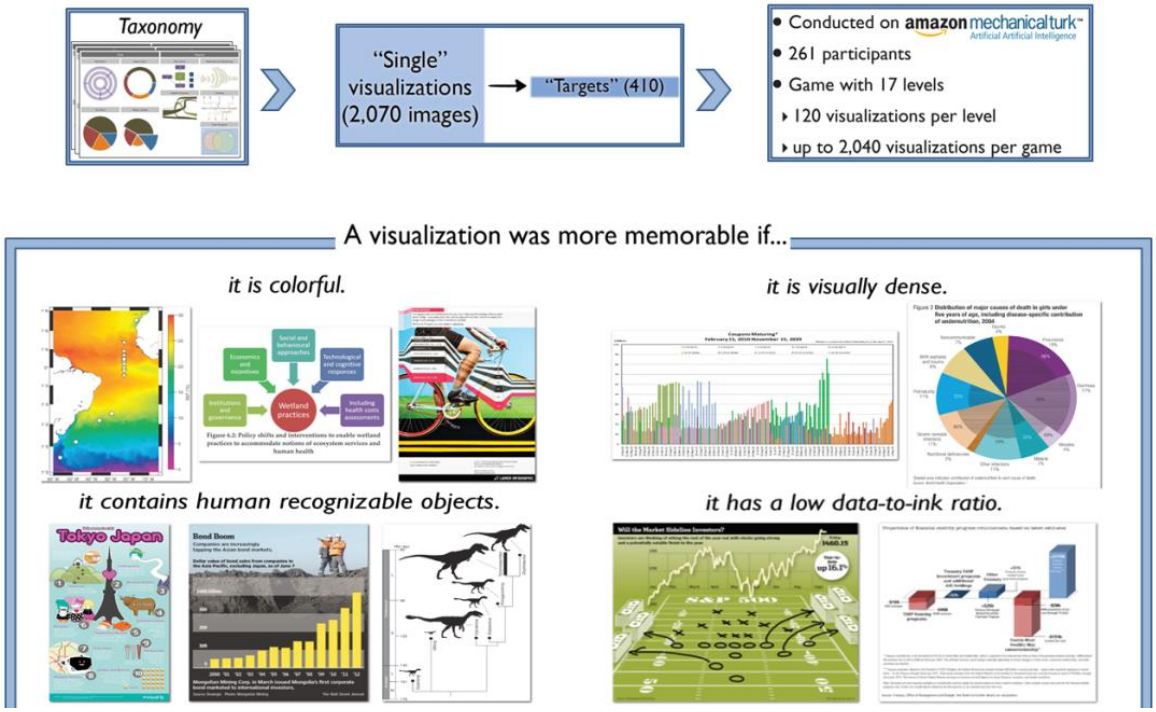


Figure 2. Study Progression and Results (adapted from Borkin et al. 2013)

Bateman et al. (2010) examined the effect of charts with visual embellishments (versus non-embellished charts) on comprehension and memorability. They selected fourteen charts from the book *Designer's Guide to Creating Charts and Diagrams* by Nigel Holmes that contained visual embellishments (such as images or cartoons) used as a visual attraction that would engage the reader and facilitate the message. The research team converted the selected charts into non-embellished generic charts.

The study consisted of two phases: the chart reading and description phase and the recall phase. Participants were not aware that the second phase consisted of recalling the graphs from the first phase. During the chart reading and description phase, each participant was presented with embellished and non-embellished charts and was asked to answer the same four questions about each graph (See Table 2). The second phase was divided into two conditions: immediate recall and long term recall. Participants either played a five minute card game immediately after the chart description phase (immediate recall) or returned one to two weeks later to complete the second phase (long term recall). After the second phase, participants were asked to choose their preferred visualization. Data collected included experimenter-coded response scores, eye-gaze data and a preference questionnaire (see Table 2).

Results showed that Holmes charts received better value message descriptions than the generic charts. Recall score differences for participants in the immediate condition were significant for the value message. Participants in the long-term recall condition had significant differences for all three question categories (See Table 2). Results also showed that participants preferred charts with visual embellishments for the following categories: most enjoyed, most attractive, easiest to remember, easiest to remember the details in the

embellished chart, fastest to describe and fastest to remember. Gaze data was collected to determine what percentage of the overall time the participant spent looking at the data on the embellished chart versus the plain charts. The percentage of time spent on the embellished charts was 40% and 22% for the plain charts. The longer time spent on observing the embellished charts did not result in more accurate chart description. The authors conclude that participants were able to recognize value messages in the Holmes charts due to the visual message conveyed by the embellishments. They also note the potential value offered by including visual imagery in charts. Finally, they highlight that issues related to chart design and context-specific use are important areas to explore.

Experimenter Coded Response Scores and Interview Questions	
4 Interview Questions	**Response Scores (0-3 points)**
1. **Subject**: • What is the chart about? 2. **Values**: • What are the displayed categories and values? 3. **Trend**: • What is the basic trend of the graph? 4. **Value Message**: • Is the author trying to communicate some message on the chart?	• **3 = all correct**: the participant provided all of the relevant information in a given category. • **2 = mostly correct**: the participant provided most of the relevant information in a given category, but omitted one or two important details. • **1 = mostly incorrect**: the participant provided one or two correct or relevant information in a given category, but omitted most relevant details. • **0 = all incorrect**: the participant provided no correct or relevant information in a given category. • **0 = "I don't know"**: the participant stated that they did not know any of the relevant information in a given category.
Categories for Preference Question	**Preference Measure**
• Most preferred • Most enjoyed • Most attractive • Easiest to describe • Easiest to remember • Easiest to remember details • Most accurate to describe • Most accurate to remember • Fastest to describe • Fastest to remember	• Participants chose their preferred chart format based on preference question categories.

Table 2. Interview Questions (adapted from Bateman et al., 2010)

The GRE data interpretation question

The quantitative section contains a number of data interpretation questions. This section is presented as a set of questions that require analyzing data presented using common graphical methods that are limited to bar graphs, segmented bar graphs, circle graphs, histograms, scatterplots, and time plots. The GRE quantitative exam preparation document (Educational Testing Service [ETS], 2010) describes the goal of the data interpretation section as being able to "make better decisions as a result of understanding the data within trends and making predictions with a level of certainty" (p.61).

The data interpretation questions are housed in the quantitative sections of the GRE. There are 2 quantitative sections that each last 35 minutes averaging 1 minute 45 seconds per question for 20 questions. Springmeyer, Blattner and Max (1992) observed a multidisciplinary group of scientists in order to determine characteristics of their processes involved in scientific data analysis. The observations resulted in two categories of this process they labeled as Investigation and Integration of Insight. The Investigation process is summarized as interacting with the data representations and applying math; Investigation and Integration of Insight are bridged by what they refer to as "maneuvering of data". "Maneuvering of data" refers to the process of going through and organizing, extracting, and synthesizing the data collected during the data investigation phase. The final process of Integration of Insight is expressing ideas derived from the Investigation process. The steps of this process parallel to the process of answering a GRE data interpretation question.

Selection of the GRE questions

GRE questions from 4 different test preparation books were examined and categorized by number or mathematical operations required. The participants' task is to

answer the questions selected based on minimal operations required. This was done to avoid effects on accuracy in responding to the GRE question based on mathematical skill level.

In a study where visualization effectiveness is measured using electroencephalography, Anderson et al., (2012; 2011), state that it is difficult to measure the cognitive requirements of visualizations by trying to objectively quantify the essentially subjective metrics used in the methodologies typically used for such studies. They also note that it is difficult to evaluate different visualization methods based on task complexity (Anderson et al., 2011). They conclude that the key to addressing these "objective evaluations" may be by using visualization techniques based on task complexity.

Ashcroft & Krause (2007) highlight that as the complexity of the math problem increases, so does the demand on working memory. They provide evidence that mathematical problem solving using a strategy approach demands more working memory resources than does a memory recall process. In other words, mathematical operations that required a strategy, such as counting, relied on working memory as opposed to automatic retrieval. This was measured through reaction time (RT) as well as accuracy of a recall while performing a mathematical task. Participants' response to "using a strategy" while performing the mathematical operations matched their increased reactions times. Finally, they also point out that cognitive demand also increases with the number of steps in the mathematical problem. This was determined from results of slower reaction times of participants answering 2-column addition problems that required a carry (e.g., 27+14) and algebraic calculations that required two successive brackets to be expanded. Errors from the algebraic calculations were found to be a result of working memory failure as opposed to inadequate algebraic knowledge (Ayres, 2001).

Working memory, memory encoding, memory retrieval, attention focusing and decision-making (Menon, 2010; Stillman, 1996) are some of the key cognitive processes involved in mathematical operations. The extent to which each process is involved can be estimated by the complexity of the problem. Smith and Stein (1998) identify task characteristics that form a four-part taxonomy of cognitive demand.

These cognitive demand categories include two "l*ower-level*: functions (i.e., memorization and procedures without connections) and two "*higher-level"* (procedures with connections and doing mathematics). There are two levels of task difficulty: (1) simple recall of previously learned fact or formula to address an unambiguous question where connections between concepts and numbers are not expected (lower-level memorization) and (2) tasks requiring unpredictable and ambiguous conceptual connections, non-algorithmic processing and problem solving. Tasks were analyzed with this model to develop a 4-level hierarchy of GRE data interpretation complexity. Additionally, the number of steps and mathematical operations required to answer each question were considered as selection criteria for the questions administered to the participants (Stillman, 1996).

The GRE data interpretation questions range in complexity from a 1-step process, without mathematical operation, that can be found on basic graphs to a multiple step process, with mathematical operations, that include interpreting relationships between the multiple graphs presented. In addition to the mathematical operations required to answer these questions, there are *elementary perceptual tasks/encodings* involved when reading the data presented in the typical graphs that the GRE questions use. These tasks, which are reflected in the visual mapping process of the visualization pipeline (North, 2012), determine the following on each graph: position along a scale (common or nonaligned), length, direction,

angle, area, volume, curvature, shading and color saturation (Cleveland & McGill, 1984). The questions selected for the infographic interpretation are found on Appendix A. The same amount of questions, reflecting the same mathematical complexity (number and type of operations) was selected as the "as-is" non-infographic GRE question. In other words, these were not represented *infographically*, but instead were presented to the participant as they appeared in the source.

Design of infographics for the GRE questions

There are various design frameworks and processes used to create visualizations. Literature on design of visualizations not only points out the existing disconnect between the design process and the design choices but also the difficulties in evaluating the effectiveness of the design. McKenna et al. (2014) present a framework that can be used by the designer as a guide for the process of creating visualizations specific to a problem or application. The framework consists of four design activities: *understand, ideate, make* and *deploy*. These activities are then addressed as determined by their *motivation*, or specific purpose (is the motivation to brainstorm or to solve a specific problem), and their *outcomes* (i.e., the specific result of the design activity). There are two methods used to further describe the outcomes of the design activity, *generative* and *evaluative*. The generative method allows for specifying a range of outcomes (is the designer addressing a single or multiple tasks/issues/ideas). The evaluative method describes the level of formality of the tools used in the process (based on designer preference versus controlled empirical research). Each design activity in relation to the design process for the infographics of the GRE questions is outlined in Table 4.

Question #	Problem and Answer	Solution Steps	Mathematical Operation(s) Involved
P1C14	If there were 38 child safety organizations and the funds contributed to these organizations in September 1989 were evenly distributed, how much did each charity receive? a. $12,000,000 b. $9,400,000 c. $2,500,000 d. $250,000 e. $38,000	1. determine how many $ were contributed in Sept 1989 for child safety 2. Divide $9.4million by 38 = $250,000	1. find and (÷)
K2B51	What was the total enrollment (in hundreds) for colleges in 2000? a. 110 b. 125 c. 135 d. 140 e. 150	1. determine totals from bars for year 2000 2. add 15+55+40=110	1. (+) (+) (+)
K2A48	In 2004 only 10,000 residents of town W lived in homes with efficient appliances and good maintenance. How many gallons per day were used by these residents for the three daily household purposes requiring the most water? a. 110,000 b. 160,000 c. 270,000 d. 300,000 e. 460,000	1. determine the top 3 water usage sources 2. add 9+10+11=30 3. multiply 30*10,000 residents	1. (+) (+) and (*) 1.9+10+11=x x*10,000
K4D46	The pie chart shows the relative quantities of five types of fruit sold every season at a market stand. if the total revenue for all the fruit sold one season was $7520, how much more revenue was collected that season for pears than for apples? a. $113 b. $200 c. $376 d. $752 e. $1,128	1. determine percentages for apples and pears 2. subtract 20%-15%=5% 3. multiply 5%*$7520	1. (-) and (*)

Pine Trees	The chart shows the portions of all four types of trees on main street. If there are 140 tress altogether on Main Street, how many of the trees are pine? a. 40 b. 24 c. 21 d. 15 e. 14	1. determine what percentage of trees are pine 2. multiply 140 by 15%	1. (+) (-) then (*)
K1B5	In 2009 how many millions of dollars were revenues from frozen food operations? a. $950 mil b. $800 mil c. $600 mil d. $350 mil e. $300 mil	1. determine total revenue from total revenue bar chart: $3billion from 2009 food related 2. add frozen and snack foods from pie chart 3. multiply responses	1. (+) and (*)
K1D41	In 1998 what were the approximate profits from the sales of cordless drills? a. 50,000 b. $70,000 c. $80,000 d. $90,000 e. $100,000	1. determine the total profits from all hardware tools in 1998: 30%(x)=60,000 X=200,000 2. multiply percentage of cordless drills 35%*200,000	1. (*) algebraic (*) for percentage
K2D58	Which answer is greater than the amount of revenue, in millions of dollars, earned by team X through food sales in 1997? a. 7 b. 10 c. 12 d. 13 e. 18	1. determine the food sales (venue) revenue for team x from bar chart = 40mil 2. multiply by percentage of food sales 35% 35% of 40 mil = 14 Question asks for the answer greater than 14	1. (*)

E5A51	The percentage of internet users in the US is 51%. How many internet users are there in the United States (in millions)? a. 5 b. 15 c. 20 d. 48 e. 35	1. determine the total internet users worldwide 29%(x) = 27.3mil 2. from total, multiply 51% to determine US users	1. (*) algebraic (*) for percentage
K5C55	If the stadium is sold out, what is the gross income on the sale of all the Upper-Deck tickets? a. $300,000 b. $350,000 c. $375,000 d. $400,000 e. $425,000	1. determine the number of seats that are upper deck 25% of 50,000 2. determine the price for upper deck seating $30 3.multiply 25%*50,000*30	1. (*) (*)

Table 3. Selected GRE Questions and Solutions

Activity Framework Adapted from McKenna et. al., (2014)

Activity	Motivation	Outcome
Understand	*"Gather, observe and research available information to find the needs of the user."*	1. Considerations: not as rigid as constraints but relate to aesthetics or usability
Ideate	*"To generate good ideas for supporting the understand outcomes."*	1. Structuring data in various forms (i.e., sketches, wireframes, prototypes)
Make	*"To concretize ideas into tangible prototypes."*	1. Making design choices that result in concrete prototypes of previously potential forms
Deploy	*"To bring a prototype into effective action in a real-world setting in order to support the target users' work and goals."*	1. The resulting "usable visualization system"

Creating Infographic Representations from Selected GRE Questions Using the Activity Framework

Understand	- Data presented in each question was written in table format to determine the data involved - This data was separated in two categories, text (i.e., titles, legend items) and numerical data (i.e., years, dollar amount, percentages). - In addition to the text and numeric data category the need for a relationship layer to aid in visualization is determined	- Considerations: Which graphic design guidelines will be used to display information? o Based on a minimalist design and incorporating aspects of the visual mapping step (from the visualization pipeline) 1. apply glyph: use icons that will convey idea of GRE question 2. use attributes from visual mapping to layout data with visually
Ideate	- What are the potential ways to present the data structured in the understand phase. - List different icons to represent the theme used in each question	- Select potential icons - Sketch representations of data: combine text and numerical data with different icons to illustrate the data (the layer that incorporates/adds a visual relationship)
Make	- Decide which tools to use for the creation of infographics from selected sketches	- Create infographics on vector and online drawing software - Consult with graphic design team to create infographics - Make adjustments to draft drawings to reach a final version
Deploy	- Present infographics to users	- Infographics are uploaded to a platform where users can access visual representations to complete a task

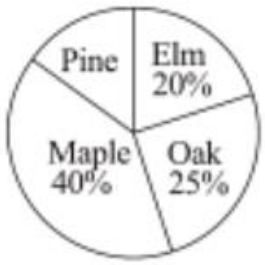
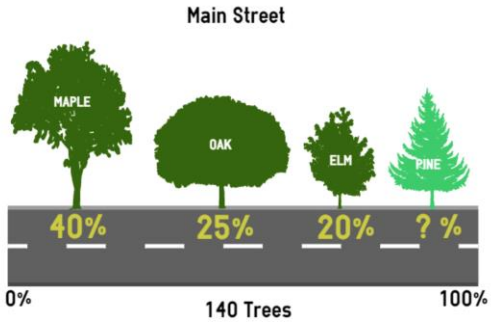
Table 4. Activity Framework Applied to Creation of GRE Infographic Representations

The next step in the design of the infographic borrows from the visual mapping element of the visualization pipeline. Though the visualization pipeline (the process of turning raw data into a visual representations intended for interaction with its end user) is computational based (North, 2012), the visual mapping element focuses on the presentation of the data. As summarized by North (2012), visual mapping consists of turning data points into visual *glyphs* and assigning them *visual properties* (i.e., spatial position, size, color, orientation, shape, texture, motion, blink, density and transparency) which can then be used to elicit desired outcomes of the problem requirements (in this case the GRE questions), and user insights. The GRE questions selected were visually mapped using the visual glyphs listed above, and the common attributes of position, length, color, orientation and shape.

This study presented infographics as an alternative visualization display in order to determine their effectiveness at presenting data required to answer standardized questions (see Table 5). This was performed using a simple mathematical task that uses common graphs to answer questions and comparing it to the same task with the data presented instead *infographically* (see Appendix A for a complete list of questions converted to an infographic).

The task consisted of answering a set of selected questions from the data interpretation section of the Graduate Record Examinations standardized test. This study hypothesized that participants would respond to questions more accurately when presented with data via an infographic display as opposed to data via traditional graphs (non-infographic display) as those used on the Graduate Record Examinations (GRE) portion of data interpretation multiple choice questions. The GRE questions selected only included addition, subtraction, multiplication and division. The GRE questions were chosen to control

for differences due to math performance; therefore, no effects from math difficulty were predicted.

Chart GRE Question	Converted into Infographic
 TREES ON MAIN STREET	

Question: Pine Trees

The chart shows the portions of all four types of trees on main street. If there are 140 trees altogether on Main Street, how many of the trees are pine?

a. 40
b. 24
c. 21
d. 15

Table 5. Example GRE Question and Infographic Counterpart

Method

Participants

Psychology undergraduate students from the research pool participated voluntarily in this research in order to partially fulfill a course requirement. A total of 77 participants completed the entire questionnaire (female = 48, male=29).

Design

A repeated measures design was used to measure performance across different graph(ic) types (infographics and non-infographic chart/graphs from example GRE questions). The independent factors were the type of graph(ic) and the level of math difficulty (easy and medium). The dependent variable was performance as measured by the correct responses on the GRE type questions.

Measures

A total of 20 GRE type questions were selected and separated into easy and medium difficulty. For each level of difficulty, five GRE questions were converted into an infographic display while the other five questions kept the same graph/chart (i.e., non-infographic) display adapted from the GRE sample question. Participants were presented with a questionnaire that included twenty GRE data interpretation type questions (5 infographic easy, 5 infographic medium, 5 non-infographic easy and 5 non-infographic medium). Questions were presented in the order taken from the GRE source material which alternated display type and difficulty level. All participants were given the same questionnaire with all of the questions listed in the same order (See Table 6).

Question Order	Type of Display	Math Difficulty Level
1	Non-Infographic	Medium
2	Non-Infographic	Medium
3	Infographic	Medium
4	Infographic	Easy
5	Non-Infographic	Easy
6	Non-Infographic	Easy
7	Infographic	Easy
8	Infographic	Easy
9	Non-Infographic	Easy
10	Non-Infographic	Easy
11	Infographic	Medium
12	Infographic	Medium
13	Non-Infographic	Easy
14	Non-Infographic	Medium
15	Infographic	Easy
16	Infographic	Medium
17	Non-Infographic	Medium
18	Non-Infographic	Medium
19	Infographic	Medium
20	Infographic	Easy

Table 6. Question Order with Display Type and Math Level

The GRE questions were multiple-choice type with one correct answer. Scores for the GRE questions counted as the total number of correct responses.

The preference questionnaire asked the participants to select their preferred type of graph (infographic versus non-infographic) on a series of characteristics (See Table 2).

The demographic portion consisted of two questions. Participants were asked their gender and their level of math completed.

Procedures

Participants were scheduled to meet in the designated lab space at the time of their pre-scheduled appointments using the online system on campus. The questionnaire was

posted on Qualtrics.com. This online platform is used for data collection. The participants were given 45 minutes to complete the questionnaire.

Upon arrival, the Qualtrics page was loaded on all computers and ready for each participant. Each scheduled appointment allotted for a maximum of three participants per timed session. A brief scripted introduction was given and questions were answered. Once participants were ready to begin, they were given the code to access the questionnaire.

Results

Summary of Findings

The data collected was analyzed using a factorial 2 (type of display) x 2 (math difficulty) within-subject analysis of variance (ANOVA) to determine whether the type of display, infographic or non-infographic, influenced the participants' responses to the GRE type data interpretation questions. This study hypothesized that there would be a difference in scores on the multiple choice questionnaire based on the type of display. A within-subjects ANOVA was used to analyze differences among four experimental conditions:

1. scores on easy GRE questions using an infographic to display data
2. scores on medium difficulty GRE questions using an infographic to display data
3. scores on easy GRE questions with the non-infographic used to display data
4. scores on medium difficulty GRE questions using the non-infographic to display data

A total of 77 participants completed the entire questionnaire. Results indicated a significant main effect for type of display, F (1, 75) = 36.091, $p < 0.001$, partial $\eta^2 = .325$, where the questions with infographic displays had significantly higher scores ($M_{\text{infographic}}$ = 2.303, SE = .112 and M_{regular} = 1.727, SE = .097), a significant main effect for level of difficulty, F (1, 75) = 30.573, $p < 0.001$, partial $\eta^2 = .290$ where the easy math level questions had significantly higher scores (M_{easy}= 2.357, SE = .110 and M_{medium} = 1.674, SE = .113), and a significant interaction between type of display and level of difficulty, F (1, 75) = 26.381, $p < 0.001$, partial $\eta^2 = .260$ where participants scored higher on medium difficulty questions that used the infographic display versus the regular graph display ($M_{\text{infographic medium}}$ = 2.18, SD = 1.335 and regular graph medium difficulty M = 1.01, SD = 1.057). For the

comparison between gender (see Figure 3), a significant difference in scores was also noted, F (1, 75) = 15.245, $p < 0.001$, partial $\eta^2 = 169$. Overall, males scored higher than females ($M_{\text{males}} = 2.379$, $SE = .147$ and $M_{\text{females}} = 1.651$, $SE = .114$). In other words, the type of display had a significant effect in the way the participants interpreted the information. Higher scores were noted on the questions that used infographic displays compared to the questions that used the regular graphs. Additionally, the type of display and the level of difficulty combined also influence the responses. The scores were higher for the infographic displays on the medium difficulty questions and no significant difference was found on the easy level questions based on display type.

Participants were also asked to rate their confidence after each GRE type question. A Pearson's correlation value was calculated as a control for evaluating the relationship between the accuracy of the response chosen by the participant and their understanding of the math performed. There was a significant positive correlation between the average confidence score and their average number correct across all four conditions, r=0.395, N=77, $p < 0.001$.

Results show that scores on the infographic medium difficulty level were significantly higher than the scores on the non-infographic medium difficulty level questions ($M = 2.18$, $M = 1.01$) respectively. In other words, the type of display did not have an effect on performance for the easy level questions. However, the infographic display significantly aided performance on the medium difficulty questions.

Type of Display and Math Level	Mean	Standard Deviation
Infographic Easy	2.21	1.162
Infographic Medium	2.18	1.335
Non-Infographic Easy	2.30	1.225
Non-Infographic Medium	1.01	1.057

Table 7. Descriptive Data for Questionnaire Scores

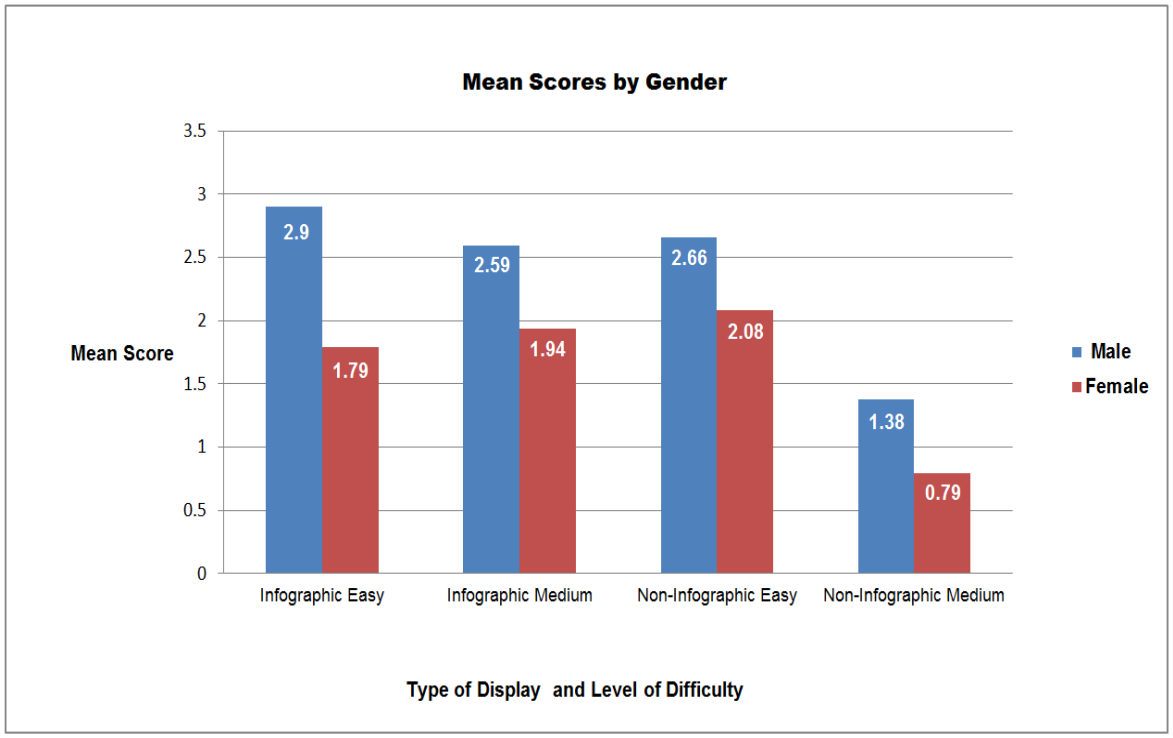


Figure 3. Mean Scores

Discussion

This study used two types of visual displays to present information required to answer GRE data interpretation questions. The questions were categorized into four groups based on type of display and level of math difficulty. Statistical analysis demonstrated that the infographic display assisted with the medium difficulty GRE questions. However, no significant difference was found between the scores of the easy GRE questions, regardless of the type of display. This is commonly found in instructional research (Clarebout et al., 2011), where problem solving for simple tasks does not benefit from supplemental material and performance is related to prior knowledge. In other words, participants' performance is related to their prior knowledge in solving math problems of the level required for this study. Though the GRE questions used were at a high school math level, math performance in general may have had an effect in the resulting scores. This could be better controlled for in future studies.

Results also demonstrated that scores were higher among the male population. This occurred regardless of the type of display or level of difficulty. There is evidence of gender differences among math performance; however it has been shown that the size of the effect of those differences is minimal (Hyde, 2014). Overall, performance was low in both males and females, however the size of this effect in this study was large (partial $\eta^2 = .169$).

The design of the infographics incorporated visual tools commonly used in design for visual perception. The tools used were color, size, hierarchy and icons for visual representations of the message in the GRE questions. The goal of these tools was to emphasize the information that was required in order to answer the question

correctly. Qualitative data was not collected, however, this may be the next step to determine how participants used the visuals to their advantage.

Design Improvements

Interpreting the results of this study highlights areas for improvement. The questionnaire focused on multiple choice questions that required a certain level of math skill. The GRE data interpretation questions were selected based on the fact that the purpose of this study was not to test math skill. This required the selection of the less difficult math level questions. This may have contributed to the non-significant difference for the easy question category across both types of displays.

Infographic displays are used to convey messages in an efficient manner. Incorporating a qualitative component that would ask the participant to interpret the visuals as they are presented with them may provide insight to the qualities of the infographic that assist with message interpretation. Selecting a series of infographic displays based on similar qualities to then perform a qualitative questionnaire with GRE type questions that are medium math level difficulty may help determine the role of the infographics' assistance to answer the questions. The qualitative component can be performed before and after the calculation process in order to collect information related to "how" the infographics are being used to complete the given task. In other words, the task of this study was for participants to answer questions by performing math (i.e., calculation process), the qualitative component would ask the participant to express their thought process while they decipher each infographic component independently before and after performing the mathematical task.

Finally, other questionnaire improvements may be considered. Participants were exposed to the same questions and in the same order. Question randomization to reduce participant fatigue was not incorporated in this study.

Practical Implications

Understanding the aspects of data interpretation advances the use of infographics. Some of these include the increased use of infographics as a tool in the classroom, to educate the general public and even to influence behavior. In the classroom, infographics are used as a creative way to have students work on a specific message. The collection of data, the analysis and interpretation of data and the display of the data are the main focus of such infographics assignments (Calacademy, 2015).

Critical information is often fragmented (and hidden away in unappealing web pages) and discourages the public from understanding the presented message. Government agencies are increasingly using infographics as a didactic tool to communicate this critical information to the public in a concise and thought-provoking manner. These infographics also present an "action" element used to influence behavior. The Centers for Disease Control and Prevention (CDC) developed a sharable gallery of infographics. They state that they encourage "the strategic use of infographics when the message is more visual in nature and requires more than data or charts to communicate successfully to the target audience" (CDC, 2015). The challenge here may be standardizing the message when the target audience includes various sociocultural characteristics and evaluating the effectiveness of these infographics for the public to interpret the intended message. However, these infographics provide a nexus for presenting the potentially complex and dull information in an efficient manner.

A study designed to evaluate *Street Infographics* as an intervention presented socially and locally relevant information using the pre-existing street signs. The goal of the study was to encourage community members to "gain local knowledge reflect on their perception and even foster social interaction" (p. 133), (Claes & Moere, 2012). Participants were interviewed before and after the implementation of the street visualization signs. The visualization consisted of four main pieces of information: the title "who lives in the street", the percentage of permanent residents, the percentage of students and nationalities (Belgian, non-Belgian). The important finding of this study includes the "curiosity, personal reflection, social interaction, perceptional changes, discussion among residents and the increase of public knowledge of social issues" (p. 138) comments collected from the participants. This summarizes the behaviors that can result from effective use of visualization. These behaviors may be innate to humans in general, however, in eliciting these, the cultural and social environment should be considered. Additionally, enculturation and socialization have a great deal of influence on human behavior. Taking into consideration sociocultural influences related to data interpretation can assist infographic design to achieve effectiveness.

It may be that our language is evolving to include infographics as a way to convey messages. However, there appears to be a lack of evaluation on whether the message of the infographics is conveyed effectively. Additionally, there seems to be a discrepancy between the amount of text used in an infographic…if the images are the ones conveying a message, then how much text is really necessary? These are areas that would benefit from future research regarding data interpretation and visual displays.

Conclusion

The main driver throughout this research has been the question of "interpretation". It was hypothesized that the infographics presented would abet the interpretation of the GRE question resulting in a correct response. The mathematical task chosen was then used to design the infographic in an attempt to visually highlight the components required to perform the task. It was in this process that the importance of "relationships" emerged. According to Bertin (1967/1983) "Perception consists of defining the relationships established within the image or among images, or between the image and the real world" (p. 2). He goes on to state that these relationships are what makes graphic visualizations unique compared to other visualizations. The selection of icons used in the infographics presented pieces of information in the GRE questions that would convey a mental image, thus propitiously benefitting from the "relationship" conveyed by the icon. Granted, the relationship notion relates to regular graphs as well, however it is the level of the relationship (an actual image of the subjects used in the question to create the infographic versus spatial and size relationships used in regular graphs) that would assist interpretation.

While standardizing the design of graphical displays may not be the important next-step, sufficient guidelines do exist for the different elements that are typically included in infographics (in other words, the existing guidelines for creating charts and graphs). The question is whether using bits and pieces of these guidelines produce an overall optimal infographic. The results of this study indicate that "visual representations" may assist data interpretation. However, evaluating the various levels of this visual assistance could result in efficient and effective graphical displays that convey the intended message and perhaps even positively influence behavior. Therefore, it is not the "what" or the "how" but instead the

“why” of the relationships portrayed by the items (such as icons or images) and the design guidelines used in an infographic that make interpretation easier or more accurate. As technology and its exponential advances provides the means to organize and decipher growing amounts of data, individual variances that make standardization so difficult become more manageable – particularly if these are analyzed using a data visualization filter. Future research of infographic interpretation should continue to examine and evaluate these relationships and how they relate to completing specific tasks accurately and efficiently.

References

Anderson, E. (2012). Evaluating visualization using cognitive measures. *Proceedings of the 2012 BELIV Workshop*, 1-4.

Anderson, E., Potter, K., Matzen, L., Shepherd, J., Preston, G., et al. (2011). A user study of visualization effectiveness using EEG and cognitive load. *Computer Graphics Forum*, *30*(3), 791-800.

Ashcraft, M. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*(1), 75-106.

Ashcraft, M., & Krause, J. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, *14*(2), 243-248.

Ashwin, C. (1984). Drawing, design and semiotics. *Design Issues*, *1*(2), 42-52.

Bateman, S., Mandryk, R., Gutwin, C., Genest, A., McDine, D., et al. (2010). Useful junk?The effects of visual embellishment on comprehension and memorability of charts. *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, 2573-2582.

Bertin, J. (1983). *Semiology of Graphics.* (W. J. Berg, Trans.). Madison, WI: The University of Wisconsin Press. (Original work published 1967)

Borkin, M., Vo, A., Bylinskii, Z., Isola, P., Sunkavalli, S., et al. (2013). What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2306-2315.

California Academy of Sciences (2015). Infographics in the Classroom Teacher Toolkit. Retrieved May 8, 2015, http://www.calacademy.org/infographics-in-the-classroom-teacher-toolkit

CDC (2015). Social Media Tools, Infographics. Retrieved May 01, 2015, http://www.cdc.gov/socialmedia/tools/infographics.html

Claes, S., & Vande Moere, A. (2013). Street infographics: Raising awareness of local issues through a situated urban visualization. *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*, 133-138.

Clarebout, G. , Horz, H. , Elen, J. , & Schnotz, W. (2011). Compensation mechanisms when interacting with learning aids. *World Journal of Education*, *1*(1), 119.

Cleveland, W., McGill, R. (1984) Perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79* (387), 531 -554.

Dur, B. (2014). Data visualization and infographics in visual communication design education at the age of information. *Journal of Arts and Humanities*, *3*(5), 39.

Educational Testing Service. (2010). *Graduate Record Examinations Math Review Chapter 4: Data Analysis*. Retrieved from Educational Testing Service website.

Friendly, M. (2009). Downloaded from http://www.math.yorku.ca/SCS/Gallery/milestone/

Grolemund, G., & Wickham, H. (2014). A cognitive interpretation of data analysis. *International Statistical Review*, *82*(2), 184-204.

Isola, P., Parikh, D., Torralba, A. (2011). Understanding the intrinsic memorability of images of images. *In Advances in Neural Information Processing Systems* (pp. 2429-2437).

Jones, D. L & Tarr, J. E., (2007). An examination of the levels of cognitive demand required by probability tasks in middle grades mathematics textbooks. *Statistics*

Education Research Journal, 6(2), 4-27.

Katz, J. (2012). *Designing Information: Human Factors and Common Sense in Information Design*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Kosara, R., & Mackinlay, J. (2013). Storytelling: The next step for visualization. *Computer*, *46*(5), 44-50.

Leonelli, S. (2014). Data interpretation in the digital age. *Perspectives on Science*, *22*(3), 397-417.

Liu, S., Cui, W., Wu, Y., & Liu, M. (2014). A survey on information visualization: Recent advances and challenges. *The Visual Computer*, *30*(12), 1373-1393.

McKenna, S. , Mazur, D. , Agutter, J. , & Meyer, M. (2014). Design activity framework for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 2191-2200.

Menon, V. (2010). Developmental cognitive neuroscience of arithmetic: Implications for learning and education. *ZDM*, *42*(6), 515-525.

National Academy of Sciences. (1995). Emerging needs and opportunities for Human Factors research.

North, C. (2012). Information visualization. In Salvendy, G., *Handbook of Human Factors and Ergonomics*, (pp. 1209-1236). Hoboken, NJ: Wiley.

Ru, G., & Ming, Z. (2014). Infographics applied in design education. *2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA)*, 984-986.

Salvendy G. (2012). *Handbook of Human Factors and Ergonomics*. Hoboken, NJ: Wiley.

Shen, I., & Cheng, W. (2015). Gestalt rule feature points. *IEEE Transactions on*

Multimedia, *17*(4), 526-537

Springmeyer, R., Blattner, M., & Max, N. (1992). A characterization of the scientific data analysis process. *Proceedings of the 3rd Conference on Visualization '92*, 235-242.

Stein, M., & Smith, M. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School*, *3*(4), 268-275.

Stillman, G. (1996). Mathematical processing and cognitive demand in problem solving. *Mathematics Education Research Journal*, *8*(2), 174-197.

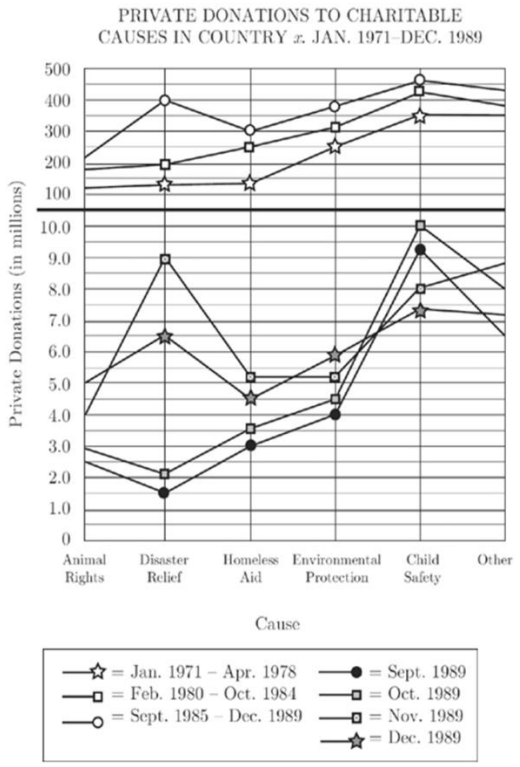
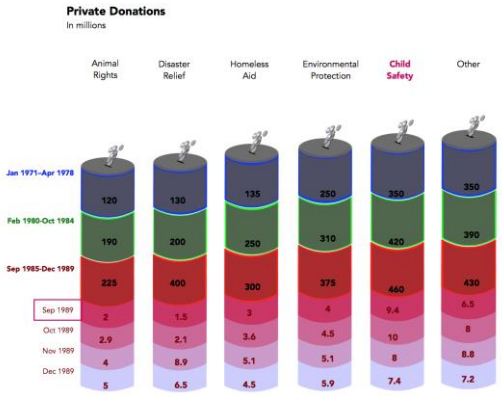
Tory, M., & Moller, T. (2004). Human Factors in visualization research. *Transactions on Visualization Research*, *10*(1), 72-84.

Tufte, E. (1983). *The Visual Display of Quantitative Information*. Cheshire, Conn.: Graphics Press.

Vickers, P. , Faith, J. , & Rossiter, N. (2013). Understanding visualization: A formal approach using category theory and semiotics. *IEEE Transactions on Visualization and Computer Graphics*, *19*(6), 1048-1061.

Wertheimer, M. , & Riezler, K. (1984). Gestalt theory. *Social Research*, *51*(1/2), 305-327.

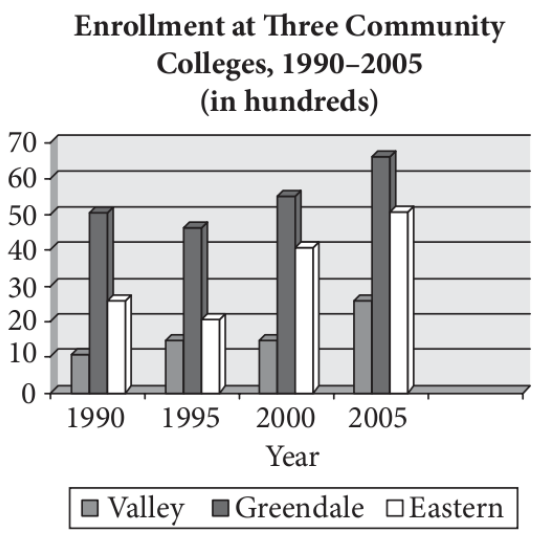
APPENDIX

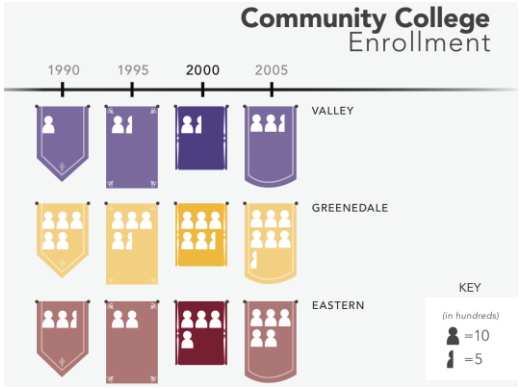
Chart GRE Question	Converted into Infographic
	

Question: P1C14

If there were 38 child safety organizations and the funds contributed to these organizations in September 1989 were evenly distributed, how much did each charity receive?

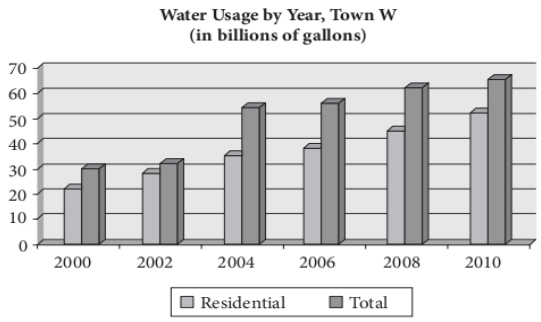
a. $12,000,000
b. $9,400,000
c. $2,500,000
d. $250,000
e. $38,000





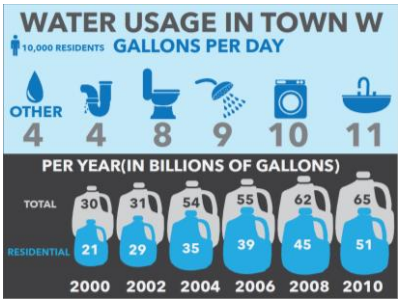
Question: K2B51
What was the total enrollment (in hundreds) for colleges in 2000?

a. 110
b. 125
c. 135
d. 140
e. 150



Daily Water Usage Statistics
(with Efficient Appliances and Good Maintenance)

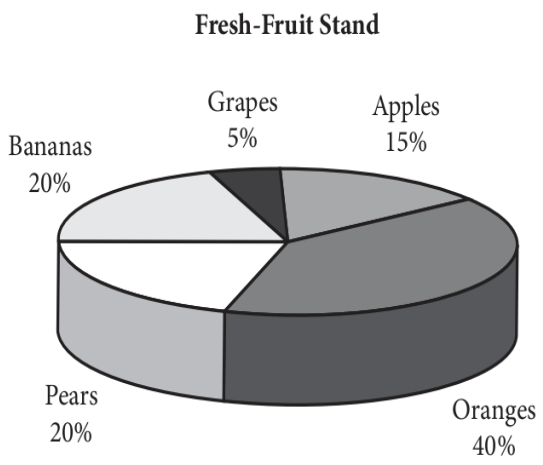
Use	Gallons per capita
Showers	9
Clothes Washers	10
Toilets	8
Leaks	4
Faucets	11
Other	4

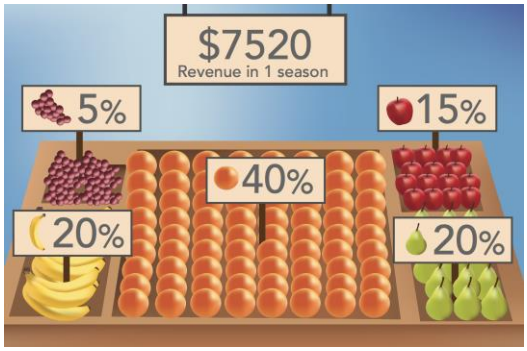


Question: K2A48

In 2004 only 10,000 residents of town W lived in homes with efficient appliances and good maintenance. How many gallons per day were used by these residents for the three daily household purposes requiring the most water?

a. 110,000
b. 160,000
c. 270,000
d. 300,000
e. 460,000

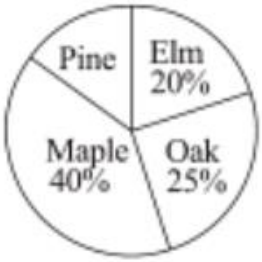




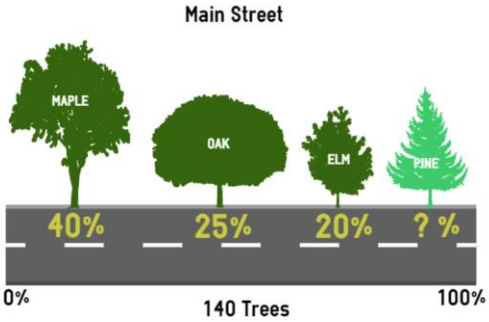
Question:K4D46

The pie chart shows the relative quantities of five types of fruit sold every season at a market stand. If the total revenue for all the fruit sold one season was $7520, how much more revenue was collected that season for pears than for apples?

a. $113
b. $200
c. $376
d. $752
e. $1,128



TREES ON MAIN STREET

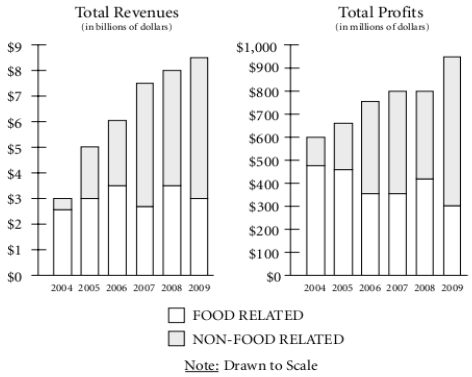


Question: Pine Trees

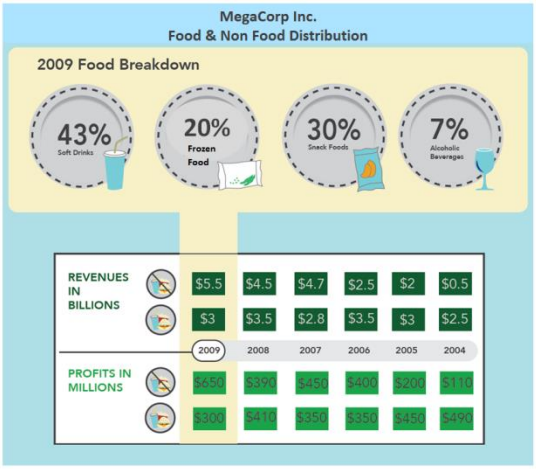
The chart shows the portions of all four types of trees on main street. If there are 140 trees altogether on Main Street, how many of the trees are pine?

a. 40
b. 24
c. 21
d. 15

MEGACORP INC. REVENUE AND PROFIT DISTRIBUTION FOR FOOD- AND NON-FOOD RELATED OPERATIONS, 2004–2009



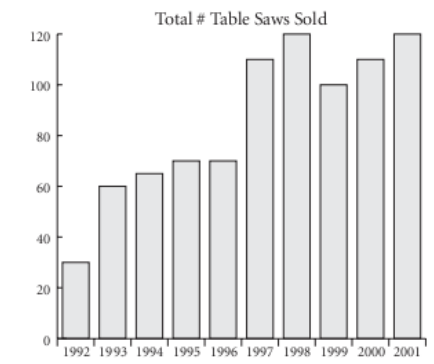
PERCENT OF REVENUES FROM FOOD-RELATED OPERATIONS IN 2009 BY CATEGORY

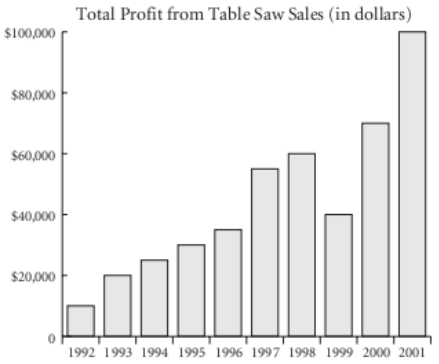


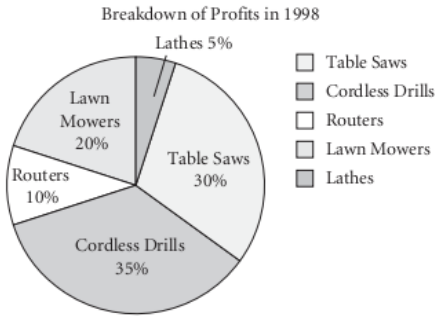
K1B5

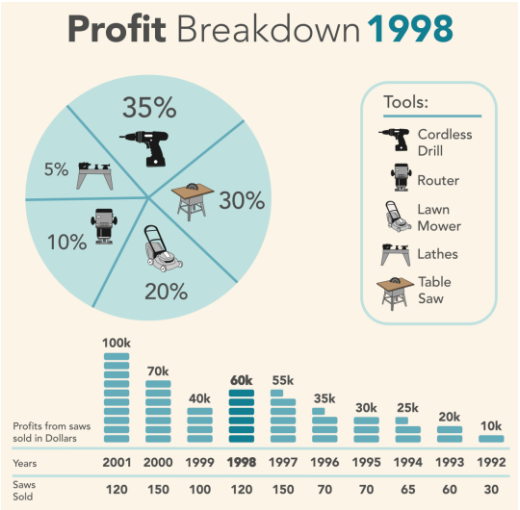
In 2009 how many millions of dollars were revenues from frozen food operations?

a. $950 mil
b. $800 mil
c. $600 mil
d. $350 mil
e. $300 mil





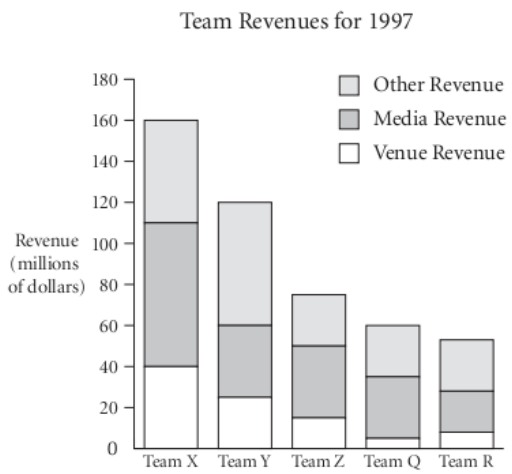


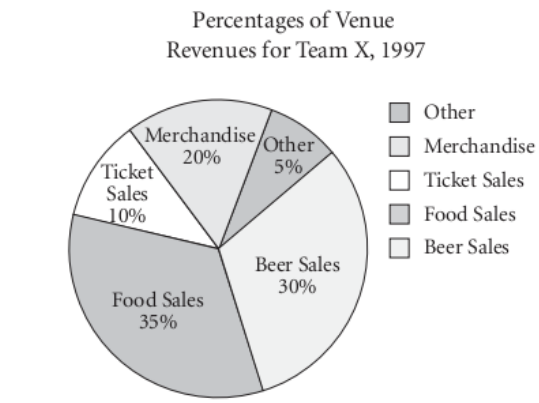


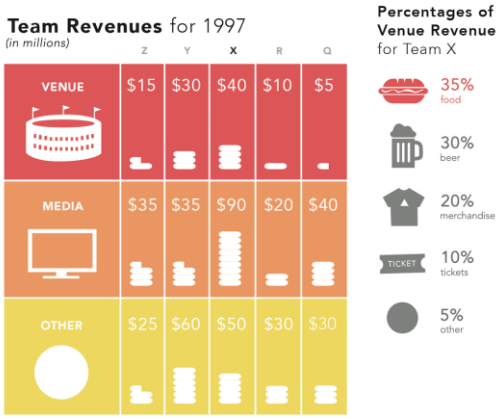
Question: K1D41

In 1998 what were the approximate profits from the sales of cordless drills?

a. $50,000
b. $70,000
c. $80,000
d. $90,000
e. $100,000





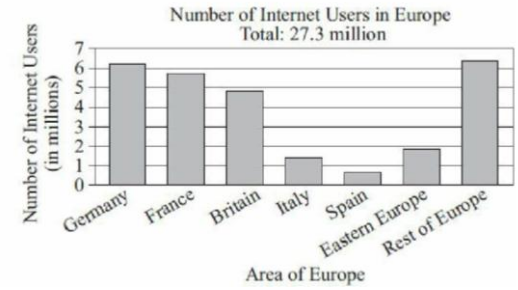


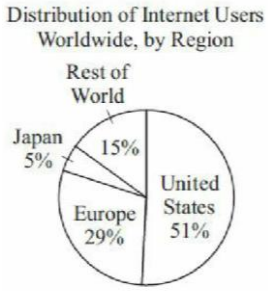
Question: K2D58

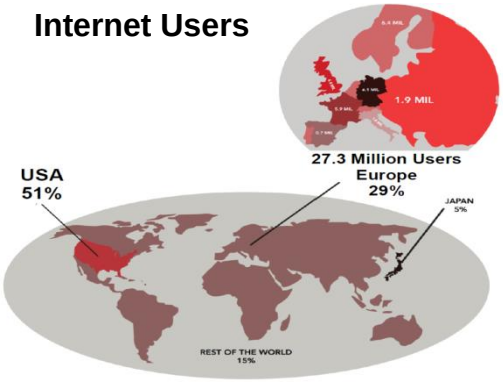
Which answer is greater than the amount of revenue, in millions of dollars, earned by team X through food sales in 1997?

a. 7
b. 10
c. 12
d. 13
e. 18

INTERNET USE IN YEAR *X*



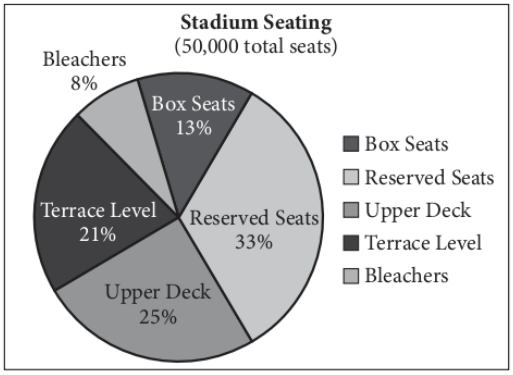


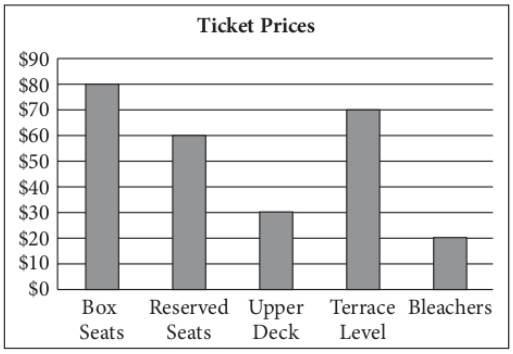


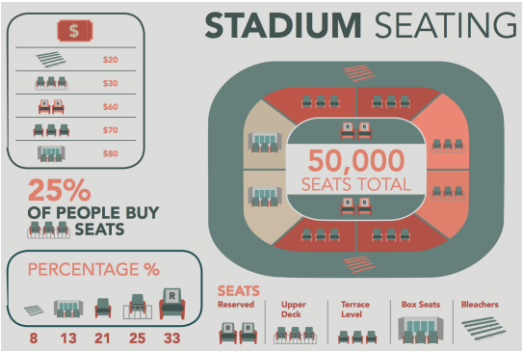
Question: E5A51

The percentage of internet users in the US is 51%. How many internet users are there in the United States (in millions)?

a. 5
b. 15
c. 20
d. 48
e. 35







K5C55

If the stadium is sold out, what is the gross income on the sale of all the Upper-Deck tickets?

a. $300,000
b. $350,000
c. $375,000
d. $400,000
e. $425,000