TOWARD A BETTER UNDERSTANDING OF VIEWERS' PERCEPTIONS OF TAG CLOUDS: RELATIVE SIZE JUDGMENT

by

Khaldoon Dhou

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Computing and Information Systems

Charlotte

2013

Approved by:

____
Dr. Mirsad Hadzikadic

____
Dr. Jing Yang

____
Dr. Mohamed Shehab

____
Dr. Ted Carmichael

____
Dr. Mark Faust

©2013
Khaldoon Dhou
ALL RIGHTS RESERVED

ABSTRACT

KHALDOON DHOU. Toward a better understanding of viewers' perceptions of tag clouds: relative size judgment. (Under the direction of DR. MIRSAD HADZIKADIC)

This dissertation focuses on viewers' perception of the relative size of words presented in tag clouds. A tag cloud is a representation of the word content of a source document where the relative frequency, or importance, of the keywords (i.e., tags) is depicted by presenting the most important tag words in a cluster called a tag cloud and varying visual characteristics of the tag words such as color, saturation, location and size. Although previous research has found that relative size is a strong visual factor for communicating relative importance of tag words, it is still unclear how viewers perceive the relative size of the words in tag clouds and how perceived size is influenced by other tag cloud characteristics. This dissertation looks at how viewers estimate the relative size of words given different characteristics such as decorations like (e.g., filled areas, boxes, and shadows), appearance of the words (e.g., varying the amount of narrow or wide letters), the typeface style (e.g., bold typeface), and location in the tag cloud (e.g., upper left vs. upper right quadrants). Significant under- and over-perception of the relative size of tag words were observed, primarily varying with the size of the target tag word. Word appearance had a modest effect on size misperception, while typeface style and location had a smaller effect. The results provide insight regarding the influence of surrounding tags on the perception of relative size of a tag word, as well as guidance to tag cloud designers regarding the influence of other presentation characteristics on perceived relative size.

ACKNOWLEDGMENTS

I would like to thank my doctoral supervisor and mentor, Dr. Mirsad Hadzikadic for his continuous patience, support and assistance in guiding me on this dissertation. Dr. Hadzikadic has been a great supervisor who has inspired and continues to inspire me. Beyond his guidance and inspiration, he was always available for me if I needed any help. His feedback has been invaluable to my advanced education and my life in general.

I would also like to take the time to thank my former supervisor Dr. Robert Kosara with whom I started this project prior to him joining Tableau Software. His valuable time, support, feedback and timely introduction to the field of information visualization are all much appreciated.

Furthermore, I would like to thank Dr. Mark Faust for his generous time and valuable help in the data analysis of information gleaned through the studies and for his valuable insights on this research ptoject. Dr. Faust was a wonderful person to work with and his overall feedback greatly shaped my dissertation.

I would also like to recognize and thank the Dr. Mohamed Shehab, Dr. Ted Carmichael and Dr. Jing Yang for their meaningful, thoughtful, and sincere comments and critiques that helped shape my dissertation and its presentation.

Certainly, I appreciate the assistance I received from my MSc supervisor Dr. David Casperson from the University of Northern British Columbia. It was with him that I started my work on chess perception, thus gaining many insights on perception, which was pivotal in making my decision to work on the perception of information visualization that I find very fascinating. Special thanks also to Dr. Andrew Kitchenham who was in my MSc supervision committee and helped me shape my research on perception and who also provided valuable input and feedback.

A heartfelt thank you to the Department of Computer Science at the University of North Carolina at Charlotte for funding me in the first year of PhD studies. Special thanks also to Charlotte Visualization Center for funding me after my first year while being in that center. Furthermore, I would like to express gratitude to Department of Software and Information Systems for funding me as a teaching assistant during the time I was with them.

Thank you also to the Graduate Student Support Specialist, Dora Bradley, for being supportive and helpful during the time of my studies. I would also like to thank the Office Manager in the Department of Software and Information systems, Gill Morgan, for all the kind attention and time.

For my parents: How can I offer a significant thank you for everything they did and are for me since the time I was a child. The encouragement and nurturing, correction and guidance have made me a better person. Special thanks also for my siblings, Salam and Basheer, my brother in law Mamoun and my nieces Layal and Lujain for always being with me with their hearts given the geographical separations between us.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Studies have shown that 90% of error in thinking is due to error in perception. If you can change your perception, you can change your emotion and this can lead to new ideas.
— Edward de Bono

Perception fascinates scholars from all different fields. They have discovered, among other things, that the way one presents information can greatly influence perception. For instance, the Coca-Cola glass bottle, with its unique shape, instantly became an icon. The unique tactile experience Coca-Cola offered ended with the introduction of plastic bottles and metal cans. These containers made it difficult for a blindfolded person to distinguish Coke from other soda brands [59, page 89]. Not only shape but also size and color intensify impact, increasing recall percentage in business magazines and, in case of size, the length of time users look at large advertisements [33, page 127]. Perception even attracted medical scholars: Their work has, in part, spurred the creation of drugs with different aesthetic qualities (e.g., shape and color) enabling consumers to differentiate one from another. Surprisingly, shape and color can evoke positive emotions and thus garner consumer loyalty [59, page 100]. Shape also ensures instantaneous brand recognition. In fact, automobile companies often user model shape as their defining feature [59, page 86]. Scholars have also explored size and obtained similar results. Plumert and Nichols-Whitehead showed that adults prefer size to color when making purchase decisions [70]. Even more, Aydinoglu, *et al* [13] found that labels of various sizes can be distinguished and acquire certain unique meanings among consumers.

Visualization is concerned with studying, understanding, viewing and communicating the visual representations of data. Perception is one of the essential aspects in information visualization [51]. Generally speaking, one must discern how eye and brain collaborate to process complex designs because misunderstanding of data can lead to serious consequences. For example, Elting, *et al* [34] found that the display of data affects the accuracy of decisions physicians make. The authors used different data displays: tables, pie charts, bar graphs and icons. The authors emphasized on the way the data is presented for the purpose of conveying the message. Schonlau and Peters [77] also noted that the display type affects the accuracy of the answers regarding recall on data. The authors used bar charts, pie charts and tables. Among the findings was the answers based on tables were more accurate than the ones in pie and bar charts. In a study by Spence [81], participants were required to compare the size of graphic elements. The author used a combination of lines, bars, pie and disk slices, cylinders, boxes and table entries. Each participant was asked to divide a line into two segments each of which is proportional to the size of one of the two elements he saw. He found that accuracy varied depending on the stimuli. Furthermore, Schapira, *et al* [76] evaluated the response of women to information about the risk of breast cancer where there were different display types. The authors found that some figures were easier to understand than other figures and some formats were even associated with risk.

A tag cloud is a visual representation of the word content of a text where the frequency of the words is communicated by the a tag cloud feature such as size. For example, size of a word in a tag cloud is proportional to its frequency in the original text - the higher the frequency of a term, the larger the typeface size. Tag clouds have become an important visualization tool for several purposes and in several fields. Figure 1 shows a tag cloud generated by TagCrowd [6], that is an online application for creating tag clouds.

Figure 1: Tag cloud example

Tag clouds are becoming popular navigational tools, as they provide a summary of the relative importance based on frequency of appearance in key words and users can use the tags to go to whatever block they want. One observes this on numerous websites, such as Flickr [10], which is a web application allowing people to host images and videos and Amazon [8], the well-known e-commerce company. Furthermore, tag clouds can be used to give users an impression of a content. Tag clouds are even used on social websites to provide an impression of a person and his interests ([39] as referenced by [73]). Tags can also be used in a search to locate specific items navigating through the text (e.g. article) as in the Google Chrome Browser [9]. Although the size of the words in tag clouds is a function of the frequency or importance of the words in the underlying database, the review of the existing literature did not reveal studies investigating how users perceive the relative size of words in tag clouds. Because typeface size is the strongest visual feature in the tag cloud [15] and it is a way of communicating the frequency in tag clouds ([45], [35]), perceiving relative typeface size is of fundamental importance for tag cloud designers as wrong judgment reflects a distortion in a user's perception, a critical issue.

Steele and Iliinsky [82] argued that size can effectively represent relative importance among entities and capture users' attention and conveying the relative size for very large and very small objects could be a challenging problem. This dissertation focuses on how users judge the relative size of typefaces in a tag cloud in the presence of other design elements already studied in the literature: Shape, font boldness, decorations and location in the tag cloud ([18]). This raises many questions: How do participants judge the relative size of words in tag clouds and how does that relate to relative size manipulation? Would larger word size affect the participants' judgment? Would decorations added to the words attract the users' attention and significantly alter their perception? What about the location? What if the typeface style is manipulated? In a recent study, Yau [93] highlighted the challenge of scaling words in tag clouds because the white space and view are less accurate than geometric shapes. This dissertation explores different parameters that interact with relative size in tag clouds, such as decorations, typeface style, word appearance, and location. Research in diverse fields has revealed the importance of such parameters. For example, decorations can be one graphic organizers used during instruction. McKnight emphasized the effectiveness of graphic organizers in empowering students to comprehending new information [65]. Likewise, design aesthetics and sight impressions are crucial in marketing and can help in creating an awareness of a certain brand [46]. Marketers must convey a visual impression to customers to convince them to purchase products.

Tag clouds use a function of frequency of key terms to determine how tags will be presented in the tag cloud. The typeface size is a primary visual characteristic in a tag cloud, however, tag cloud cloud designers use several other types of presentation characteristics that could interact with the size to influence perceived size. For example, Bateman, *et al* showed the interaction of different visual features in tag clouds (Figure 2). This clearly shows the interaction between the size of the word and the area of the tag, which makes the perception of size in a tag cloud a critical issue. Olt Aicher, a designer and a typographer emphasized on the importance of size in typography to make it appealing to the eye [52, page 56]. Typeface also received attention in business: Sulak [87] used eye tracking technology and other data collection techniques in an attempt to understand how font contributes to persuasion in business communications.

The goals of this dissertation are the following: (1) Seeing how the relative typeface size is perceived when presented in a tag cloud; (2) Seeing what other commonly used

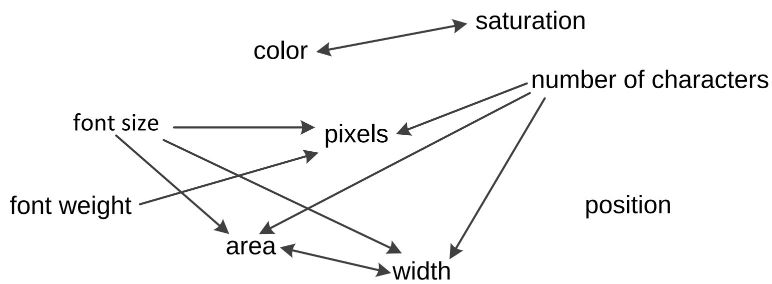


Figure 2: Interdependencies between different features in a tag cloud adapted from [15]

tag cloud presentation characteristics interact with perceived size. This dissertation is organized as follows: Chapter 2 provides an overview of applications of tag clouds in different fields. Chapter 3 reviews related published work in psychophysics and tag cloud evaluation. Chapter 4 provides hypotheses for this research. Chapter 5 provides a background on typography and shows how design guidelines were followed in the experimental design. Chapter 6 presents experiment 1, which studies the influence of distractors, appearance of words and decorations in typeface size ratio judgment. Chapter 7 presents experiment 2, which studies the influence of typeface style and location in typeface size ratio judgment. Chapter 8 provides a general discussion of the research findings in experiments 1 and 2. Chapter 9 provides summaries and conclusions.

Figure 3: Tag cloud of chapter 1 generated by `www.tagcrowd.com`

CHAPTER 2: APPLICATIONS OF TAG CLOUDS

Knowledge without application is like a book that is never read

— Christopher Crawford

Today's fast-paced world necessitates the immediate delivery of useful information to targeted customers. The bombardment of news, advertisements and mail, in our information-driven world, forces viewers to skim only a fraction of the information provided. The rest quickly fades out to the background. Tag clouds have drawn the attention of researchers from multiple fields. It is a visual tool that has gained fame with its power to convey chunks of information that when conjoined constitute a meaningful message. Besides computer science and engineering ([11], [20], [24], [23], [30], [31], [38], [48], [72], [79], [89]), tag clouds are used and studied in Systems Science and Engineering ([28], [88]); Industrial Engineering ([92]); Environmental Engineering ([41]); Sociology ([45]); Media and the arts ([64], [66]); Public Health ([37], [85]); Psychology ([25], [60], [91]); Linguistics ([27]); Education ([35], [57], [68], [84]); Public and Policy Administration ([71]); Management and Leadership([43], [49], [69], [74]); Library science ([12], [90]); Nutrition ([40]); Architecture ([62]); e-learning ([19]); Rhetoric and Composition ([67]); Nursing ([26]); and Geography ([58]). A review of the literature revealed different purposes for tag clouds. Two of the main purposes are decision making and analysis and exploration. Desai [30] used tag clouds as a way of comparing the tag behavior of different subpopulations when making a decision. In that study, end users could visually explore tag clouds belonging to different groups of users, which was more beneficial than just looking at the average rating of a certain product.

Matsunaga [64] used word clouds to design guidelines for electronic forms where different word cloud features have different meanings. For example, word size of the words in the cloud indicates the importance of the word and the color is to make the categories distinct from each other. Examples are Figure 4 and Figure 5.

Turning to analysis and exploration, in the field of industrial engineering, Wu [92] used a word cloud as a way to explore an open-ended question regarding whether Adaptive Cruise Control creates problems or safety concerns for drivers. The author conducted an exploratory analysis of participant comments using a word cloud. In linguistics, Critel [27] used tag clouds to analyze data in a study of students' participation in a writing classroom. In that work, methods were implemented to determine characteristics enabling users to perceive typeface in business documents. Levy [58] uses word clouds to illustrate changes between social groups, in terms of intergenerational differences and geographic commonalities: word clouds, frequency counts and transcript interactions provided a picture of problems facing Moldovans in their daily life. Dhone [31] used tag clouds as one way to search videos from a repository in her work on the Video Library Management Software Toolkit for the Nevada Climate Change Portal (VLMST for NCCP). Provost [72] used tag clouds as a way of providing context for the corpus for better understanding, not only showing the most frequent terms. While researching social media use in March 2011 during

Japan's neclear crisis, Stirratt [85] used Radian6 with widgets including a tag cloud. Radian6 provided access to some analytic tools and allowed the researcher to ascertain why a certain word appears in a tag cloud. Liu [60] used tag clouds to visualize the most frequent keywords and synonyms for four open-ended questions regarding the bullying of women in the Pharmaceutical/Biotechnology/Medical Device (PBMP) industry in the United States. LeNoue [57] conducted a study to gather information on the use of social networking sites by educators in education and learning. The author used Wordle to create weighted lists from some of the responses such as being asked to identify the challenges associated with the implementation of educational social software in the respondents occupational setting. Carmean [19] used tag clouds for not only word frequency analysis, but also summary construction.

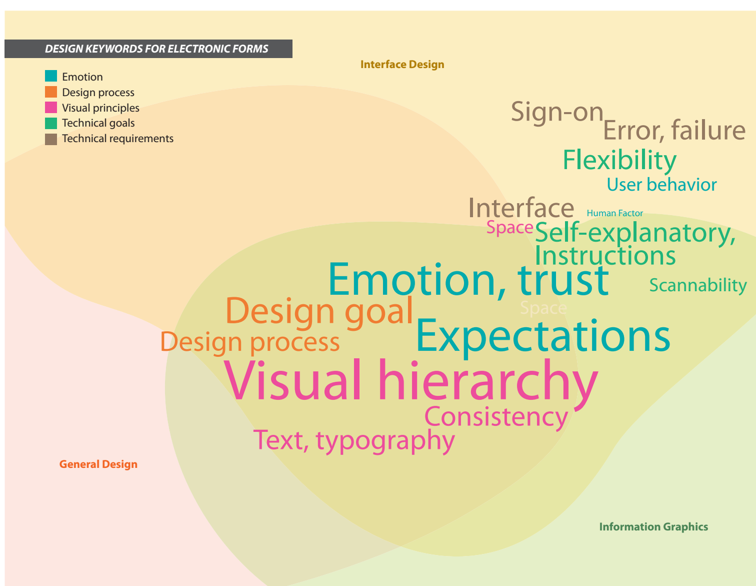


Figure 4: Example of a tag cloud adapted from [64]

In another application of location-based social media sites, Thompson [88] used a tag cloud as a way to visualize venue type (e.g. hotel, coffee shop or gallery).

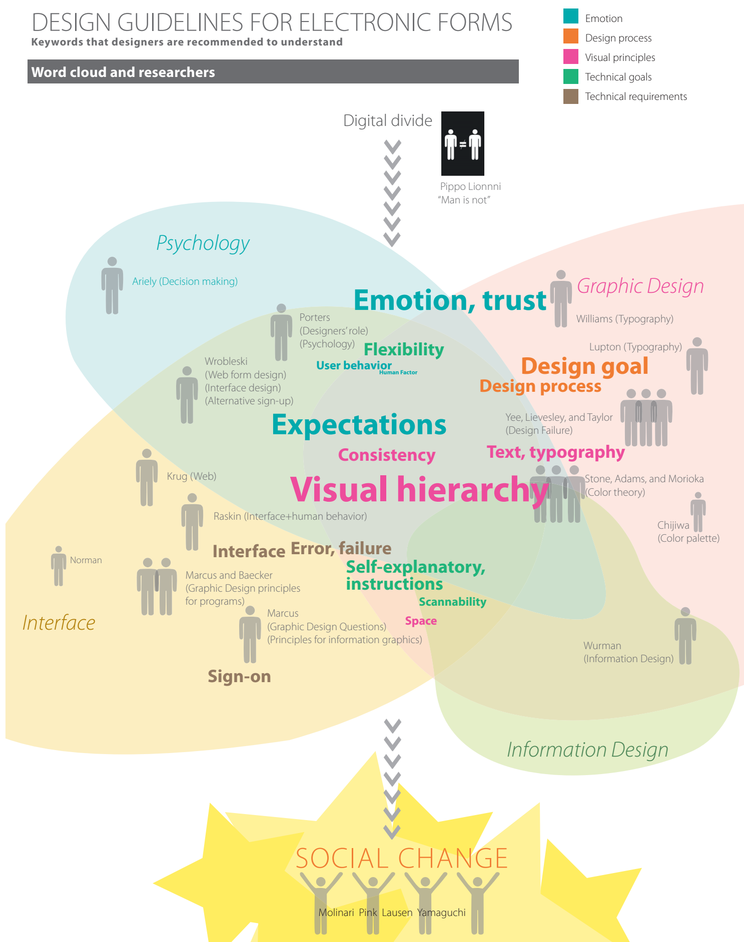


Figure 5: Example of a tag cloud adapted from [64]

Similarly, Shankar [79] used a tag cloud to visualize the tags for locations in their data set. One of the applications of tag clouds is the visualization of the documents. Hoyt [45] used a tag cloud to look for relevance and visualize codes representing quotations from articles. The author used the size of the words in a tag cloud as a feature representing the frequency of words where the higher the frequency of the word, the larger the size of that word in a tag cloud. Hayman [43] represented interviews in a tag cloud to determine which information recurred. Kantarakias [50], a researcher in leadership emphasizes on the importance of a tag cloud in his work on leadership for communicating and evaluating ideas and messages.

In a study aiming to predict crowd violence, White [91] used *NVivo 9* to visualize data in different words including tag clouds in their preliminary data analysis. Emanuel [35] used tag clouds in research on education for visualization and data analysis of participants' responses. Angel [9], performed qualitative analysis using tag clouds to gain more insights. In Nursing, subjects submitted responses to four questions asking about the strengths, weaknesses, opportunities and threats in the CCIRES (Collaborative Center for Interactive Reviews and Evidence Summaries) program [26] to evaluate and ultimately prove the program. For each of the four categories, tag clouds were generated via the NVIVO text frequency query.

Kuo, *et al* [54] expanded typical tag cloud analysis to include the construction of summaries. They described the use of PubCloud to summarize the results of the queries from PubMed, navigate the summarized results and compared the results of PubCloud with those of PubMed. The study demonstrated that tag clouds could be used both to summarize results and navigate relevant information. Sinclair and Cardew-Hall attempted to answer the question of whether tag clouds were useful for people seeking information from a folksonomy dataset [80]. The authors found that users preferred to use a tag cloud for general tasks and a search interface for specific tasks. The authors also observed that users answers most of the questions using a tag cloud.

Lee, *et al* introduced SparkClouds, a visualization combining both tag clouds and sparklines to show trends [56]. The authors found that with spark lines, SparkClouds, provided an overview of the trends with little additional space, and, with tag clouds, the layout was compact and aesthetically pleasing. This confirms the importance of studying aesthetics in tag clouds.

Figure 6: Tag cloud of chapter 2 generated by `www.tagcrowd.com`

CHAPTER 3: RELATED WORK

Aesthetics have substantial political consequences. How one views oneself as beautiful or not beautiful or desirable or not desirable has deep consequences in terms of one's feelings of self-worth and one's capacity to be a political agent

— Cornel West, *Breaking Bread: Insurgent Black Intellectual Life*

Psychophysics has to do with lawful relationships between physical aspects of objects and events (e.g., luminance, sound pressure, weight, size) and perception (e.g., brightness, loudness, heaviness, size). One general finding from this literature is Steven's power law, which indicates that for many magnitude judgments in perception, a ratio comparison is a perceptual constant no matter what part of the magnitude scale one is operating at. For example, a light that is perceived to be twice and bright as another light will continue to be perceived as twice as bright if both lights are moved up or down the brightness scale in a proportionately equal manner. This general finding works in many perceptual magnitude domains, and suggests that tasks that encourage participants to make ratio judgments of the relative size of 2 objects should yield reliable results.

In Steven's Power Law: $J = \lambda D^{\alpha}$ where

J = judged magnitude

D = actual magnitude

α = exponent

λ = scaling constant.

Cleveland, *et al* [21, 22] conducted an experiment to test previous work indicating that people's judgment of the area of a circle is proportional to the area raised to a power less than one. Participants judged circles including and not including maplike grid ticks, labels, scale and border. For the maplike stimuli, subjects were informed that an anchor circle represents toll charges of $100 and asked to give the dollar representation of three other circles without bringing the term area. For the other stimuli with no maplike grid ticks, the area of the anchor circle was called 100 units and subjects were asked to estimate the areas of the other three circles compared to that circle. According to the authors, higher exponents were obtained when the circle was available in the presence of other circles as opposed to shown to participants and removed before they make judgments.

Mates *et al* [63] measured the ability of participants to compare the areas of squares and rectangles. The authors proposed a formal model of area perception and its estimation. The stimuli used were filled and unfilled figures (a square and a rectangle) on the screen. According to the results, rectangles were judged as bigger than squares and they also found that the perception of contour influenced the users judgment of area.

Research on perception of rectangles seems to be applicable to size judgments of words because the overall shape of the words is rectangular with a dominant width axis. Based on the related work, it is predicted that people might over-judge the size of words as they do for rectangles, but it is still unclear how this will translate into a judgment of the relative ratio of sizes of 2 compared words.

This research is focused on relative size judgment of words in tag clouds. Relative size judgment where people judge the size of a scaled word compared to another word attracted researchers from various backgrounds and for various purposes. One design approach is to combine both large and small scales. Changing the scale can serve many purposes. One approach is conveying fantasy such as when a cartoon figure shrinks in size to be able to enter from the door. In addition, exaggerated size in advertising is used to draw the attention of the viewer to a certain element in the advertisement.

The application domain used in this research is the tag cloud environment. The review of the literature revealed very few studies in the evaluation of tag clouds. The way one presents information, that is, color, typeface and typeface size contributes profoundly to the formation of users' perception. Scholars from different fields have researched various visual characteristics studied in this research of tag clouds. Börner [16] demonstrated the usefulness of knowledge of graphic variables including position, form (size, shape and orientation/rotation), color (value or lightness, hue or hint, saturation or intensity), texture and optics. In addition, Klemmer [53] outlined three basic tools for visual design that powerfully convey information: typography, layout and color.

Understanding the visual features of tag clouds helps one to understand the different kinds of information that can be obtained from them. This dissertation is inspired by many studies in the area of tag cloud perception, for example, the study of Bateman, *et al* [15] investigating the most visually important features of tag clouds. The findings showed that font size, font weight, saturation and color wield a great visual influence. Other features are less likely to capture viewers' attention, for example, number of pixels, width and area. In a similar, yet different vein, Halvey and Keane evaluated the use of tag presentation techniques. The properties they found important included: alphabetization, tag position, the use of larger fonts. The authors found that bigger font sizes were associated with faster completion of tasks. Moreover, the results showed that participants were able to identify words in the upper left portion the most quickly. Their findings illustrate the necessity of studying these properties and the ways in which together they guide our understanding of perception in tag clouds.

Studying such visual features is essential if one is to effectively design and evaluate tag clouds. For instance, Rivadeneira, *et al* [73] described two studies evaluating how effective tag clouds can be for different tasks such as searching, browsing, impression formation and recognition. To build tag clouds, authors used font weight, font size, font color and word placement. The authors found that the recall of the words of larger size was higher than that for words of smaller sizes. Furthermore, the recall of the words in the upper left portion was significantly higher than that in other portions. Clearly, word size and position are important factors in users' perception and need to be studied further to gain a better understanding of tag cloud perception.

Hearst and Rosner [44] examined why designers create tag clouds and how they expect tag clouds to be interpreted. The authors used a qualitative method to assess current use, as well as the advantages and disadvantages of tag clouds as perceived by users.

Some research studies focused on the perception of tag cloud features. Zhang, *et al* [94] explored how the font size and tag location influence the Chinese perception of tag clouds. The findings demonstrated that recall for tags with larger font was significantly bigger than that for tags with smaller fonts in all three locations.

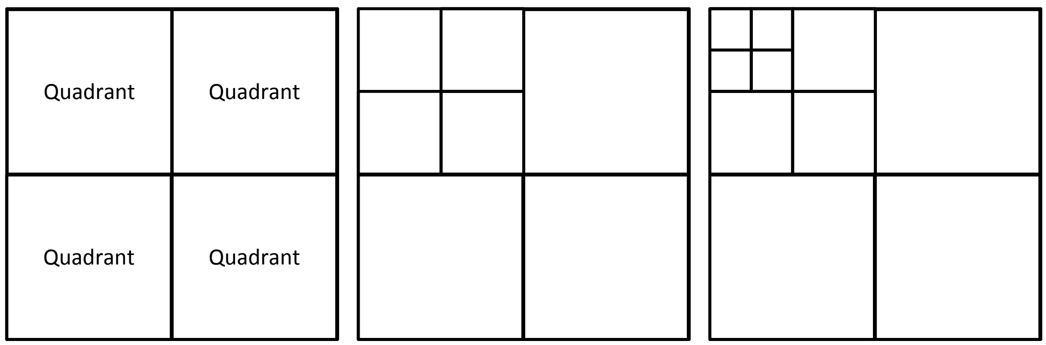


Figure 7: Three levels of quadrantizing a square adapted from [94]

Zhang, *et al* also found that the recall of tags in different locations depends on font size with a greater effect for larger fonts. Moreover, the recall effect was more significant in upper quadrants. Results also showed that eye fixation with tags of larger font size was significantly longer than that with tags of smaller font size. In addition, participants spent less time on the outer tags.

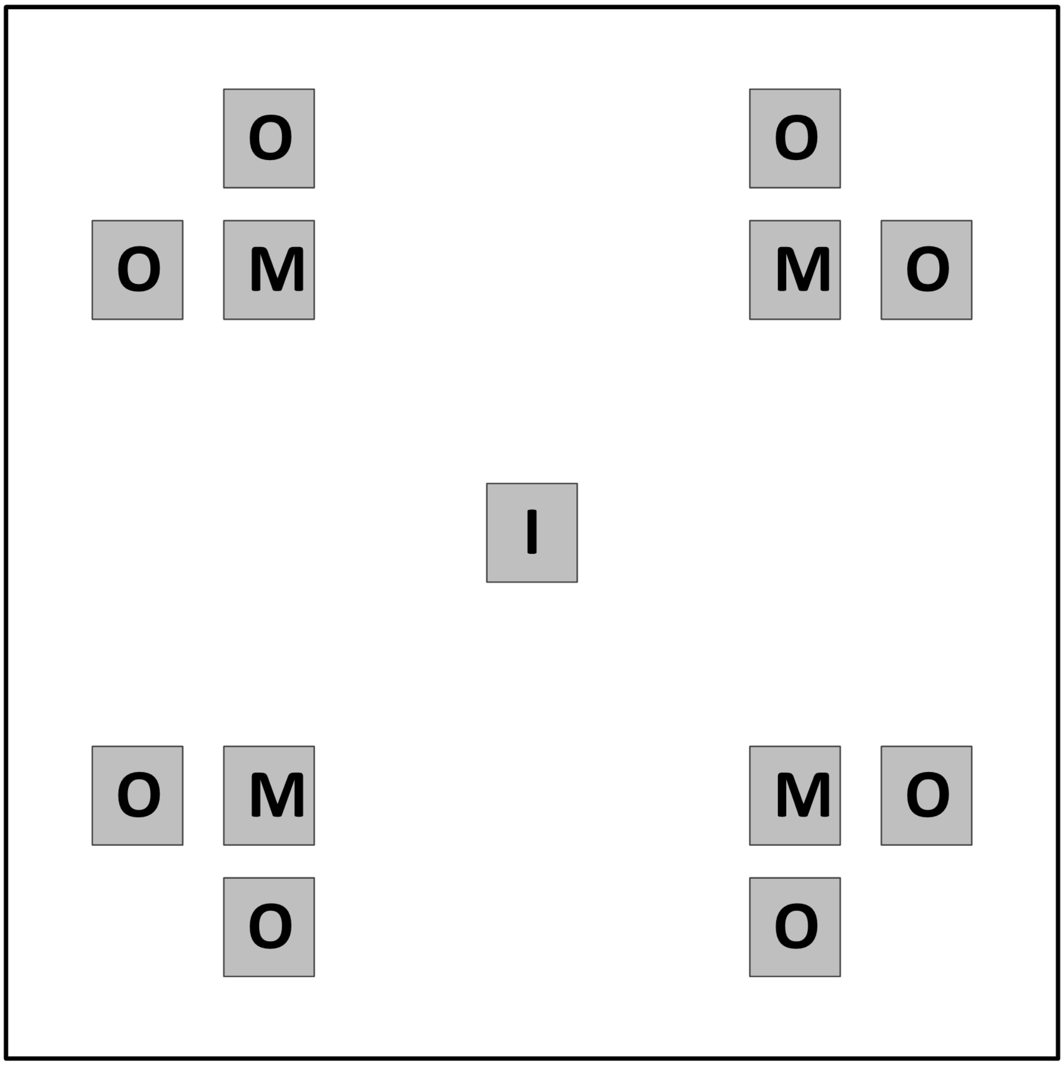


Figure 8: Locations of tags in the tag clouds in the study by Zhang, *et al* [94]. I= Inner, M=Middle, O=Outer. This figure is adapted from [94]

Figure 9: Tag cloud of chapter 3 generated by `www.tagcrowd.com`

CHAPTER 4: HYPOTHESES

The only relevant test of the validity of a hypothesis is comparison of prediction with experience

— Milton Friedman

This dissertation proposes that the characteristics of the words influence users to make different judgments on the relative size of the words. Specifically, it hypothesizes the following, derived from a review of the literature in related fields:

H1: As the relative typeface size increases, the accuracy of the result will increase. This is motivated by Weber's observation that the just noticeable difference between two weights is directly proportional to the magnitudes of these weights: $\Delta w = cw$ [95, page 131], where Δw is the just noticeable weight increment, w is the standard weight and c is a constant.

H2: Participants have differences in typeface size ratio judgment of target pairs that appear in tag clouds as opposed to target pairs that appear on the screens just by themselves. This hypothesis is supported by the unifying factor of proximity [55], which indicates that elements that are presented together are seen as a related pattern (Please refer to Figure 10). Seeing words as a related pattern is probably going to make participants judge their sizes differently.

H3: In the tag cloud environment, there is no statistical difference when users

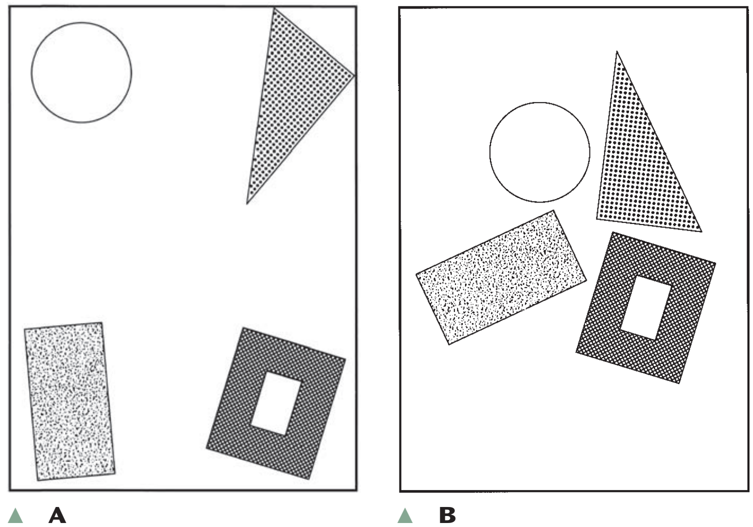


Figure 10: Isolated (A) versus close (B) items adapted from [55]

make relative typeface size judgments of a single pair of words with ascenders and descenders compared to a pair words with no ascenders and descenders. This hypothesis is supported by the results of the Cleveland, *et al* study [21] where maplike information had negligible effect on people's judgment on circle's areas. However, it is expected that the group with ascenders and descenders will have a slightly bigger relative size judgment. That is supported by the argument of Klemmer [53] which states that the higher the x-height of the typeface, the easier to read that typeface at smaller point sizes on devices with low resolution.

H4: The relative typeface size judgment of the word pairs with more ink in the target word than the anchor word is not noticeably different than the relative typeface size judgment of the pair of words with almost the same ink between the target and anchor words. This is supported by the finding of Bateman, *et al* that there is a small effect of the pixel count [15].

H5: Bold text has a strong factor that influences the participants overestimate of the typeface size ratio judgment of relative words in a tag cloud. This is supported by the study of Bateman, *et al* [15] who studied the importance of different tag cloud characteristics in capturing the attention of the viewer. Their research findings indicated that the font size and font weight (e.g. bold text) were the most influential to the viewer.

H6: Participants overestimate the size of the words in the upper left portion of the tag clouds. This is supported by the finding of Halvey and Keane [42] that the information found in the upper left portion of the tag cloud or list would result in the quickest identification. H6 is also supported by the study of Rivadeneira, *et al* [73] where the recall of the words in the upper left portion of the tag cloud was the highest. This seems logical given that in the current society people most often scan the documents from left to right and from top to bottom. The study of Zhang, *et al* [94] also supports this hypothesis as they found that the recall of the tags in the upper left quadrant was significantly better than the recall of tags in the lower quadrants.

All these hypotheses are validated in Chapter 8.

Figure 11: Tag cloud of chapter 4 generated by `www.tagcrowd.com`

CHAPTER 5: TYPOGRAPHY AND DESIGN

Readers usually ignore the typographic interface, gliding comfortably along literacy's habitual groove. Sometimes, however, the interface should be allowed to fail. By making itself evident, typography can illuminate the construction and identity of a page, screen, place, or product

— Ellen Lupton, *Thinking with Type: A Primer for Designers: A Critical Guide for Designers, Writers, Editors, & Students*

In this chapter, a background on typography will be provided to aid in understanding the research explored in this dissertation. This chapter will also review design guidelines offered by other researchers and how this research utilized them in designing experiments. For more information and background on typography, please refer to [17], [36], [53], [61] and [86].

5.1 Typography

The words font and typeface represent of different things in typography and they are often misunderstood. The word typeface as defined by Felici means "a collection of characters-letters, numbers, symbols, punctuation marks, etc.-that are designed to work together like the parts of a coordinated outfit", while a font is "a physical thing, the description of a typeface-in computer code, photographic film, or metal-used to image the type" [36]. Figure 12 illustrates the difference between a word's typeface and font.

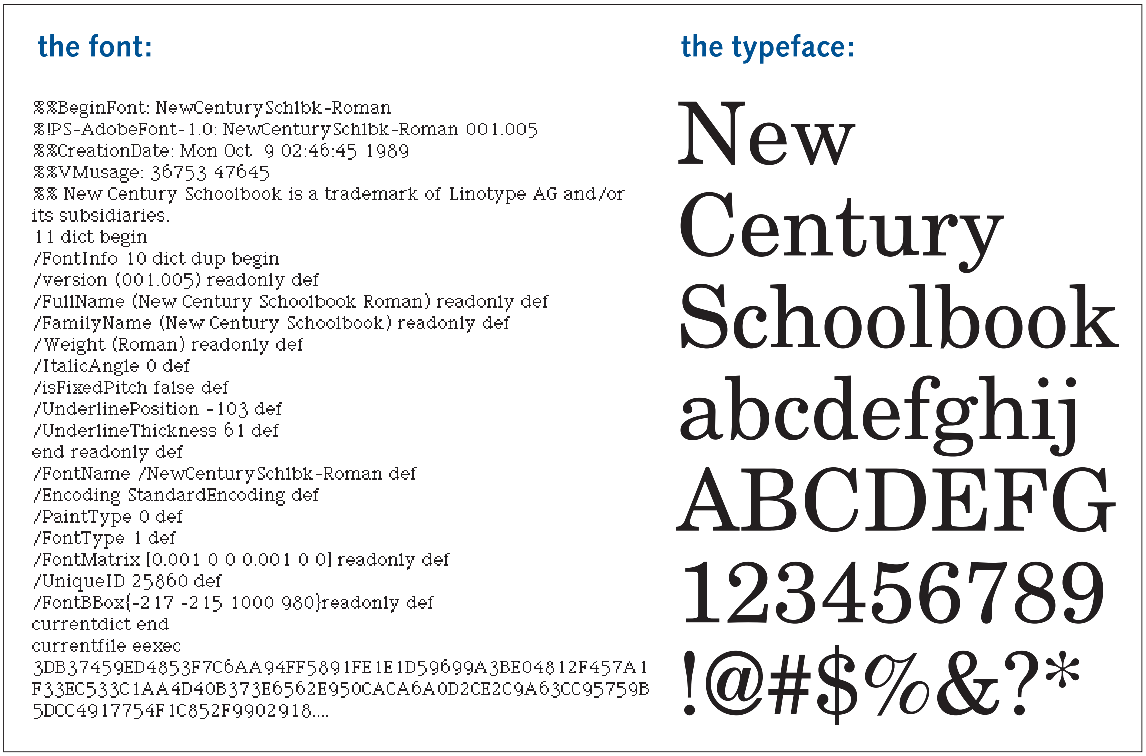


Figure 12: The difference between the words font and typeface adapted from [36]

A monospaced typeface is one in which all letters, numbers and symbols have the same width. In proportional typeface, the characters can have dissimilar widths. A popular computer typeface, Courier, is actually a typewriter typeface, and although used in computer word processing, its characters are identical in width. Monospacing left little or no room for further expression, but the idea of having words with letters varying in width was appealing to researchers. Figure 13 shows an example of monospaced and proportional typefaces. It is noticeable that words with the same number of characters always have the same length in monospaced font; the proportional typeface allows for far more variance.

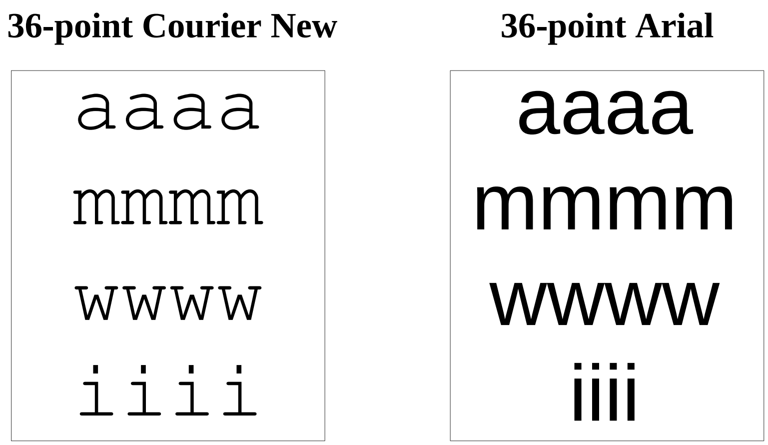


Figure 13: Example of monospaced (Courier New) versus proportional typeface (Arial). In monospaced typeface all characters have the same width while in proportional typeface characters vary in their width

When the letter is cast on its block, there is a space around the letter and that block height is defined from just above the apex of the tallest character to the limit of the lowest character (Figure 14). Such gaps prevent letters from overlapping over and touching each other. The point size is defined as the height of the virtual blocks.

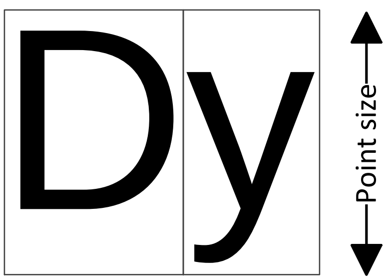


Figure 14: Point size defined as the height of the virtual block

In the typeface design (Figure 15), the baseline is an invisible line where most characters reside. The mean line is an invisible or imaginary line that resides at the top of the non-ascending small letters. Ascenders are defined as the parts of the small letters that extend above the mean line. Descenders are defined as the parts of the small letters that stretch below the baseline. The x-height or corpus size is the difference between baseline and mean line. em is a unit of width used in typography and it measures the horizontal space. One em is equivalent to the type size. For example, when the type is 16 point, the em is 16.

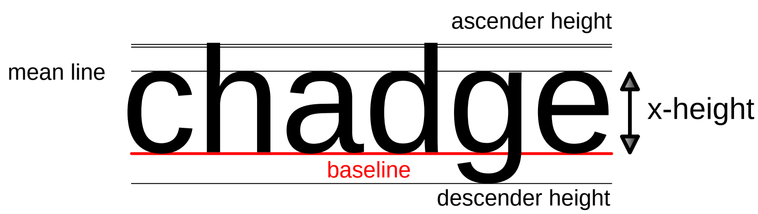


Figure 15: [An example showing some definitions in typeface design. Baseline is where most of the characters reside. The mean line resides at the top of the non-ascending small letters. x-hight is the difference between the mean and baseline

5.2 Key Practices Followed in the Experiment Design

This research followed some of the key practices to readable typography offered by Smashing Magazine [61]. Below are the key practices as explained by [61] and how they were followed in the design developed specifically for this dissertation:

(a) User-Friendly Headers

As is often the case, headers were used to capture the attention of the reader. This was important to ensure contextual understanding of the directions or information provided. User friendly headers are a crucial component in typography. Their importance extends to web and print typography. Additionally, user friendly headers are not only part of the basic text hierarchy, but also critical in scannable content, as well. It is important to understand the importance of the header size as otherwise, this can lead to distraction or misunderstanding. This key practice was followed in the instructions provided to the participants performing the experiments 1 and 1 (See Figures 16, 17 and 18). In the current design, headers were effectively used to convey the purpose (i.e. instructions of the experiment). The design considered the size of the headers so that participants do not lose their focus when reading the instructions and to pay appropriate attention to the information conveyed. Space was also utilized between the header and the text below it to insure that the information could be digested efficiently.

(b) Scannable text

Making a document scannable requires the judicious use of headers. The hierarchy and focus points of the headers are helpful to the navigation of the reader through the content. Focus points are those elements or objects within the text document that are supposed to capture the attention of the viewer. Examples are graphical elements and buttons. There are factors that make text easily scannable such as header size and position, text size, height and contrast and all these factors interact with each other. For this study, the documents were made to be scannable by having the header size and position distinguished from the body of the text. Red dots were also used as focus points to drag the attention of participants to the target pairs of words to be compared. For example, in the experiment instructions, the words ”dealizer” and ”citounst” have red circles indicating that these are the words to be compared (See Figure 21)

(c) White Space

White space, or that part of the document that does not have markings of any kind helps the users eyes travel more easily through the document. This factor also offers a break between various elements in the text such as graphics and text. The white space in this design was used to help easy navigation through the text. The instructions for the subjects in the experiment utilized white space to provide a separation between different elements in the layout and therefore help users pay attention to the meaning of the text (See Figures 16, 17 and 18)

(d) Consistency

Consistency in the hierarchy is important to a user-friendly layout. Often, consistency is viewed as a factor contributing more to usability. However, it also contributes to the readability of the written word. Consequently, it is important to have consistency in size, color, font and spacing between all headers that has the same importance. The consistency in the current design of tag clouds was applied by choosing words with certain identical characteristics within the same tag cloud. For example, in figure 22, all words in the tag cloud have bold font. Consistency was also applied to factors chosen in the design, for example having all the words in the tag cloud with the same decoration (See Figures 19 and 20)

(e) Density of Text

The text density is the amount of words placed in a certain area. Density has a significant impact on the readability of the text. It is affected by manipulating the spacing options such as text size and spaces between letters. There should be a balance between these options to help readability and scanability of the text. The text in written materials for the study was both readable and scannable as a balance between the spacing options such as the line height, spacing and text size were carefully maintained.

(f) Emphasis of Important Elements

The emphasis of particular elements in the written document is yet another key factor that can be utilized to provide a focus point for the consumer. Focus points are especially vital in Web typography. Highlighting links, using bold text for important items and displaying quotes are just a few examples. Emphasizing these objects helps providing focus points the writer desires the reader to attend to. The emphasis of important elements in this design was achieved via bolding header information to make it attractive to the eye of the user and to distinguish it from other parts (See Figures 16, 17 and 18). Dots were also used so that users could identify the target pair of words to be compared easily (See Figures 20 and 21).

(g) Organization of Information

Organized information can enhance readability. This factor was utilized to make the information easy to find and read. For example, in the instructions provided at the beginning of the experiment (Figures 16, 17 and 18), there were different blocks for instructions, tag cloud example, answer bar where users choose the answer they think it is right.

(h) Clean Graphical Implementation

Clean graphical implementation can be a challenge. While a graphic can bolster and expand on the written information, implementation can be problematic. In order for a graphic to be helpful, the author should utilize an appropriate amount of white space between the graphic element and the text of the document. Decoration can also be used to define a graphic element such as a clear border around an image. This type of decoration will help provide a noticeable division between the image and the text.

A factor, such as a border, should never disrupt the text. Icons and illustrations, for example, would best be benefited by white space versus a border as the border would diminish attention to the text. In the experiments provided by this research, space was used to separate the tag cloud image in the instructions from the text (Figures 16, 17 and 18). In addition, only simple borders were around the tag cloud to make it separate and draw the attention of participants to it (Figures 16, 17, 18, 19, 20 and 21). Furthermore, boxes around the questions requiring a response from the participant were employed (Figures 19, 20 and 21). The goal was that the presented text not be disrupted by anything.

(i) Use of Separators

Separators are a straightforward way define specific portions of information. This factor can separate major elements such as headers or lesser text into manageable portions. An extremely simple type of separator is a single, unadorned, line. Typically a simple line such as this is used to divide other elements and factors in the hierarchy. However, simple lines can also provide more subtle means of dividing sections of test or information. Boxes are used to divide content and to guide the users through the provided layout. In short, separators are used to divide the text into different parts in an organized way. This was necessary in this design to distinguish one part from another such as tag cloud, scroll bar, instructions so that the participant was alerted to look for new information. In addition, boxes around the questions were provided and they were kept simple so that users will be able to distinguish them from the rest of the items in the design (Examples are on Figures 16, 17, 18, 19, 20 and 21).

(j) Good Margins

It is ironic that the white space in a document guides the reader to the text and not away from in. The lack of white space results in a kind of chaos as the reader does not know where to begin (or end). Good margins are a very commonly used white space element. Effective use of whitespace along the sides of text will ensure the reader pays attention to the important content of the text in an article or document. Margins can also be used to help separate different elements in the design from each other. Margins were used in the instructions (Figures 16, 17 and 18) to make a separation between the tag cloud and the question asking users how much one word is bigger than another. Thus, the content of the page was further defined.

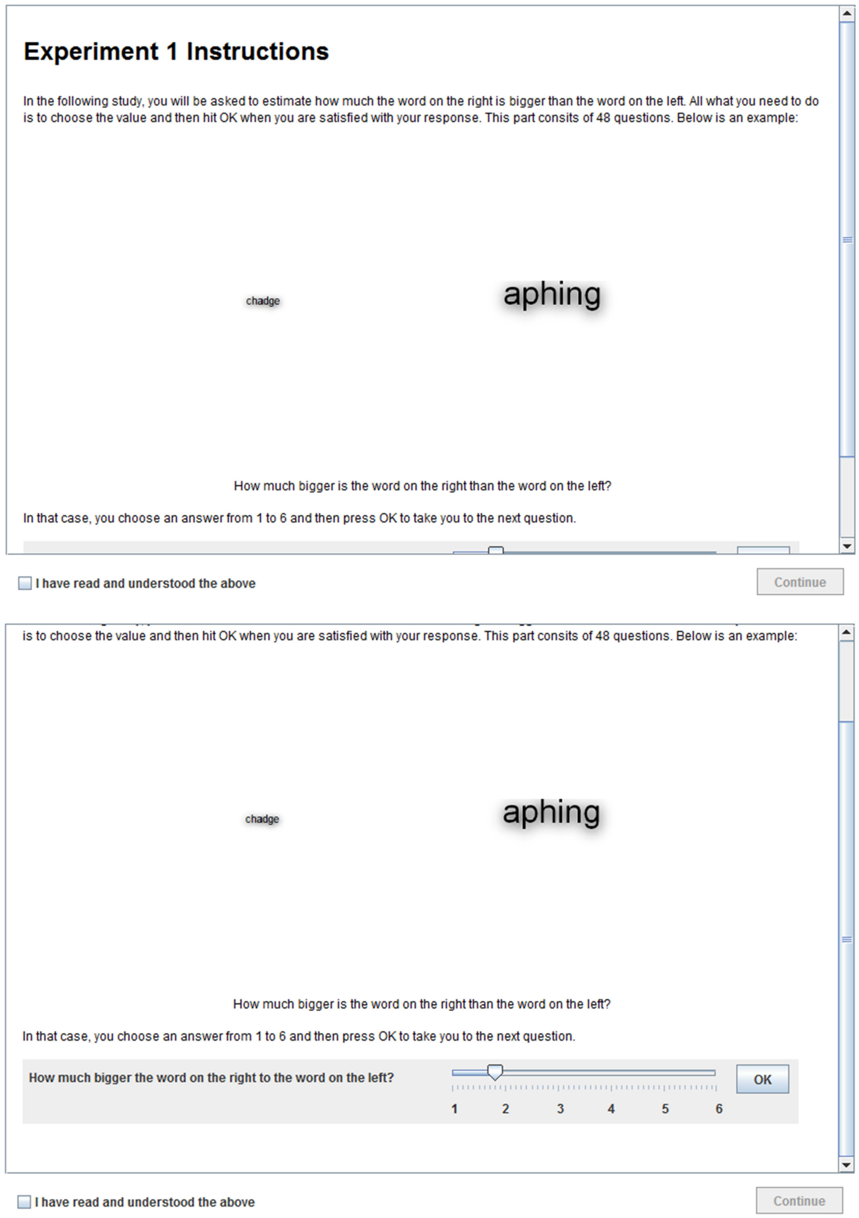


Figure 16: Instructions for experiment 1 - tag cloud absent display type

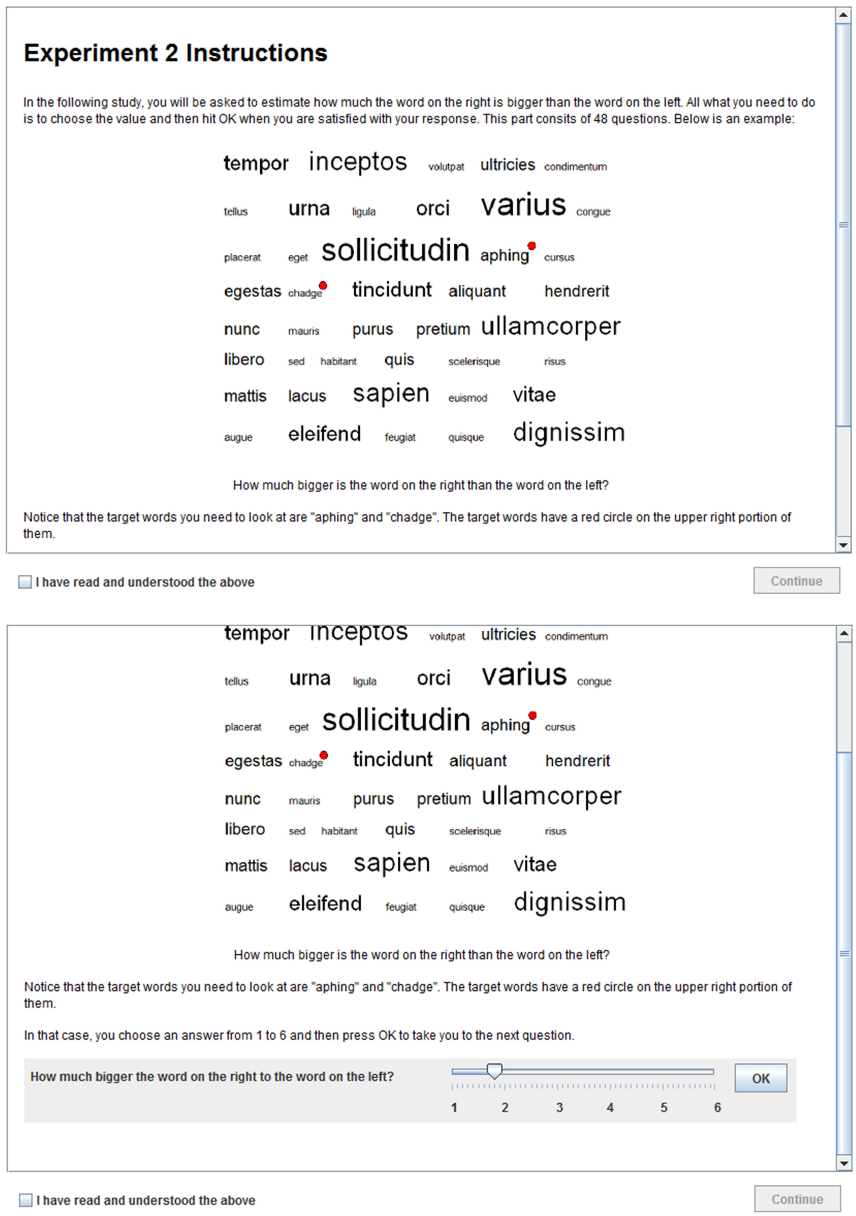


Figure 17: Instructions for experiment 1 - tag cloud present display type

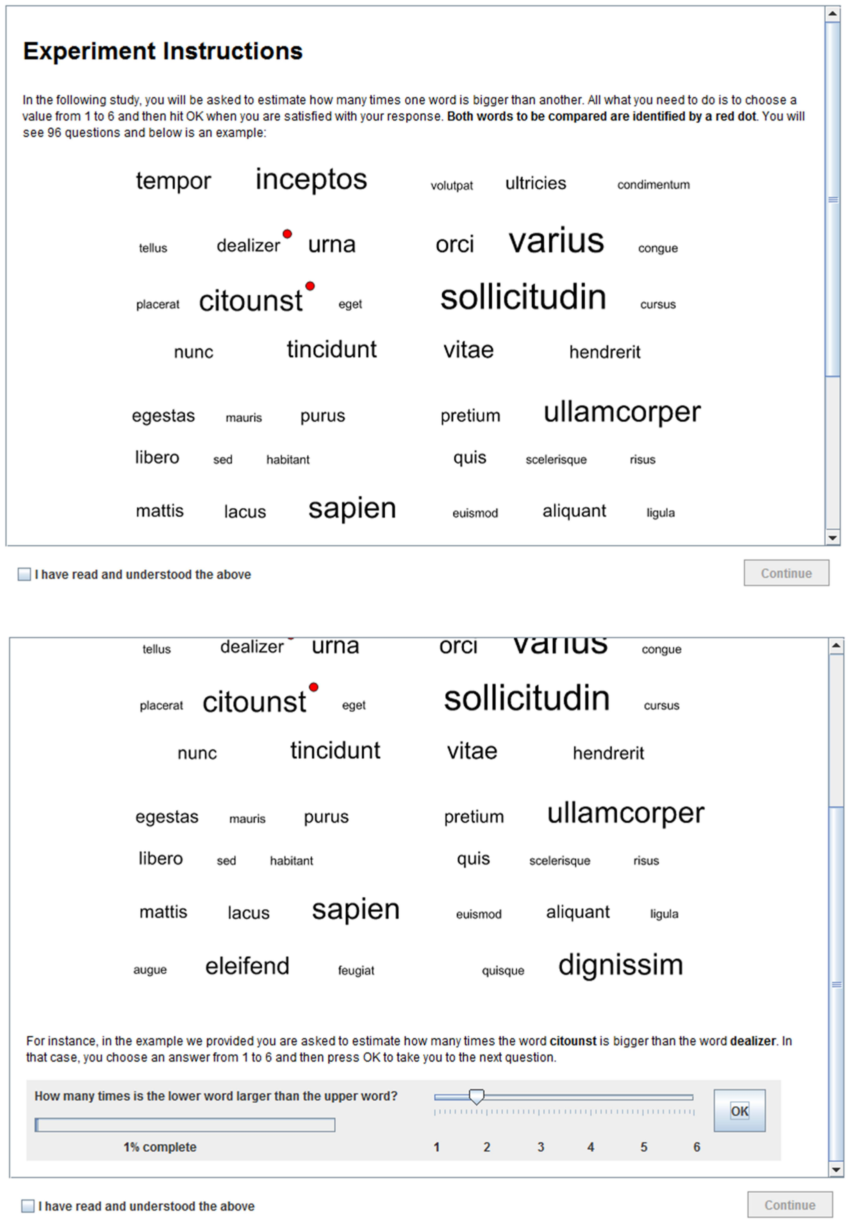


Figure 18: Instructions for experiment 2

Figure 19: Example of a screen shot in experiment 1 - tag cloud absent display type

Figure 20: Example of a screen shot in experiment 1 - tag cloud present display type

Figure 21: Example of a screen shot in experiment 2

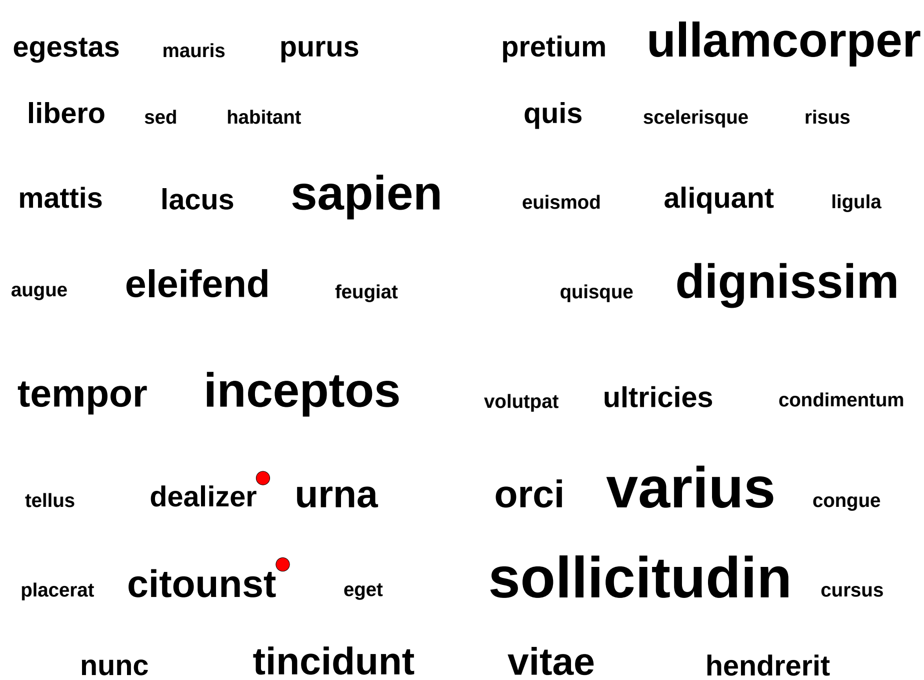


Figure 22: Example of a tag cloud from experiment 2

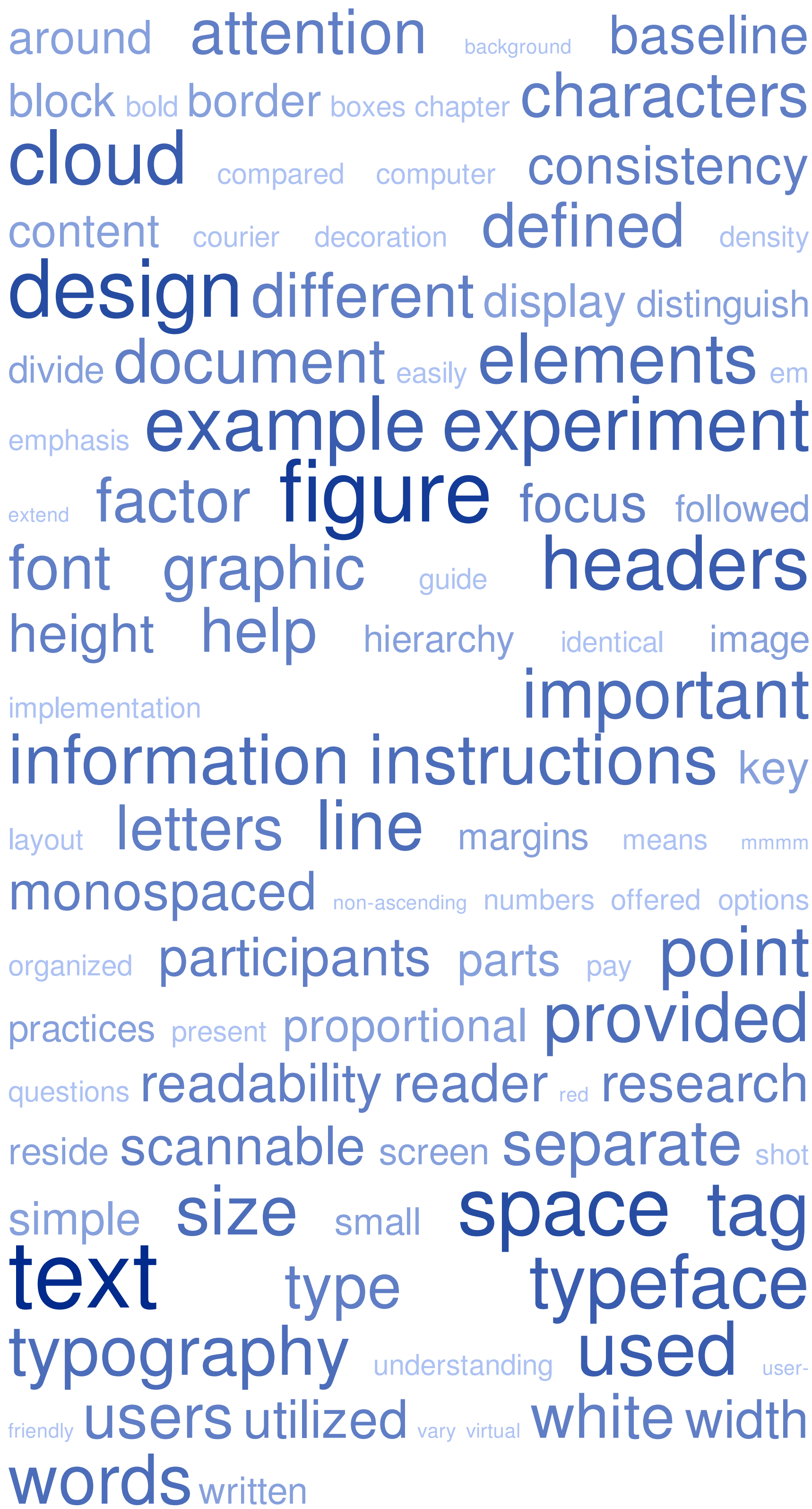


Figure 23: Tag cloud of chapter 5 generated by `www.tagcrowd.com`

CHAPTER 6: EXPERIMENT 1: THE INFLUENCE OF WORD APPEARANCE, DISTRACTORS, AND DECORATIONS ON TYPEFACE SIZE RATIO JUDGMENT

A garden is a complex of aesthetic and plastic intentions; and the plant is, to a landscape artist, not only a plant - rare, unusual, ordinary or doomed to disappearance - but it is also a color, a shape, a volume or an arabesque in itself

— Roberto Burle Marx

The focus of experiment 1 is to attempt to understand how users perceive the relative size of the words given different characteristics and without including any effect of the semantics. Tag cloud creators manipulate characteristics of words in a tag cloud, e.g. typeface size and text box decoration such as shading and boxes around the tags, or words. Typeface size has been found to be the most influential feature in capturing the viewer's attention [15]. Furthermore, size has been used for the purpose of identifying information within a tag cloud [42]. In fact, the size of the words in a tag cloud is one of the main features used to communicate frequency of tags in a source document ([35], [45]). The literature review did not reveal any study that examines the relative size or any manipulation of typeface size ratio judgment of words in tag clouds. Decorations such as boxes or frames around the words have been of interest in the creation of tag clouds [5]. The perception of decorations such as boxes and frames around the words has also been studied in psychophysics [63]. The mere appearance of words, in their basic form, is necessarily a key element in a tag cloud as the tag cloud consists of words. Since letter identification is a critical issue in research [75], this study asserts that the appearance of words can play a role in judging the relative size in tag clouds. In this chapter, the dissertation explores whether the manipulation of tag clouds display type (tag cloud present, tag cloud absent), decorations and appearance of words have an effect on the typeface size ratio judgment. The goals of experiment 1 are the following:

- Providing a first look at how veridical (i.e. accurate) relative size judgments of words in tag clouds are
- Understanding the influence of additional background words in the tag cloud on relative size judgments
- Understanding the influence of different types of letters in the words (e.g., wide vs. narrow letters, and letters with ascending and descending features)
- Understanding the influence of framing decorations (e.g., shading or boxes) around the words being compared.

It is possible that the tag cloud of background words will distract the attention [29] or the addition of tag cloud background words could lead to framing effects [32]. Having one element on the screen (e.g. target pair of words to be compared while the rest of tag aloud is absent) is a way to emphasize that element and attract the attention of the viewer to it [55, page 64]. In this dissertation, the interest was on testing whether the distractors, appearance of words and decorations have a significant effect on the size ratio judgment. The nature of a tag cloud is to communicate relative frequency (and importance) of words in the target document through variation of relative typeface size of the words in the tag cloud. Therefore, it is crucial that the viewer be able to accurately perceive relative size of words in the tag cloud. With this impression comes an indication of relative importance of the word in the original document from which the tag cloud was generated.

Extensive work in visual perception has indicated that the human perceptual system is particularly good at making ratio comparisons. This is the basis of so-called Steven's scaling, after the researcher who popularized the idea that the human perceptual system dealt with ratio comparisons of relative magnitudes (e.g., relative size, weight, or brightness) in a stable and robust manner. That is the reason of having participants make a ratio judgment (e.g., how many times larger is the target word than the comparison word) of pairs of words. Steven's scaling suggests that for simple perceptual intensities like brightness, loudness or 1-dimensional size such as length of line, the ratio between two perceptual perceived intensities is constant no matter where on the scale one is [83]. For example, if one is light is perceived as twice as bright, the relative brightness would stay the same if you double the brightness of these two lights.

6.1 Method

6.1.1 Participants

In this experiment, 65 participants were recruited from the Amazon's Mechanical Turk [1]. Participants were paid a base rate of $.50 and $0.04 for each correct answer and for an answer to be correct, it has to be within the 15% of the actual value.

Participants were not given any feedback on the accuracy of individual judgments and they were paid at the end of the study according to the accuracy. Since there were 96 questions, the maximum possible would be 4.34 (0.50 + 96 * 0.04 = $4.34). Participants were informed that the decision to participate in the study was completely up to them and if they decided to be included in the study, they could stop at any time. There were 41 male and 24 female participants. Before beginning the experiment, each participant needed to indicate that he read and understood the consent form provided, which states that he might participate if he was able to comfortably communicate in English and had 20/20 full color vision or corrected to 20/20. Participants were only asked about gender and to explain how they made the decisions.

6.1.2 Materials

Participants in the study were asked to view 96 screens containing a target word pair each: a target word and an anchor word. They were asked to judge how much bigger the target word appeared compared to the anchor word. The anchor word was always smaller in size (always 12 pt.) than target word that varied in typeface sizes that ranged as follows: 18, 24, 30, 36. The design involved manipulation of four independent variables - tag cloud display type, appearance, decoration and relative size:

1. The presence or absence of a background tag cloud in addition to the target word pair was manipulated. Half of the trials had a background tag cloud of 48 words including the target pair, while the rest only had the target and anchor words on a blank screen.

2. Appearance of the letters in the target word pair was varied by varying the amount of letters with ascending and descending portions (e.g., h, c, p, letters ascending, non and descending parts), and varying the amount of narrow and wide letters (e.g. w, i, letters that are wide versus narrow). See Figure 24 for the three word pairs used in this experiment.

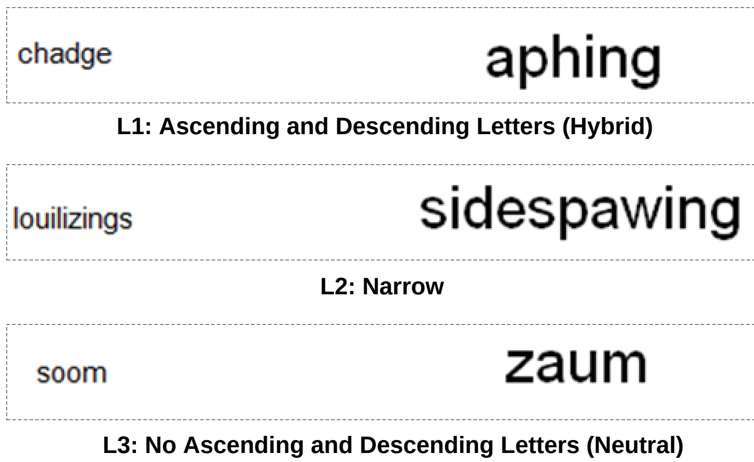


Figure 24: The appearance of words in a target word pair: L1: words have approximately the same number of ascending and descending letters (hybrid) L2: larger word that has wider letters (narrow vs. wide) L3: no ascending or descending letters in either word of target pair (neutral)

3. Target word pairs differed in terms of decorations such as shading and boxes around the words (See Figure 25). For this manipulation, all words in a tag cloud including the pair of words to be judged have the same decoration.

4. Relative typeface size of words in a pair was manipulated. The smaller word in a pair was always 12 pt (the anchor word). The larger word was 18, 24, 30 and 36 pt (See Figure 26).

For the present experiment, only three pairs of words were chosen (See Figure 24 for the pairs used in this experiment). 48 target pairs were constructed by taking the

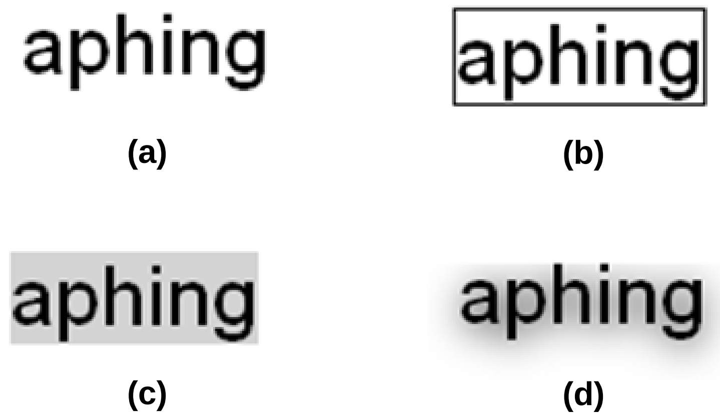


Figure 25: List of decorations in target word pairs: (a) no decoration, (b) boxes, (c) filled areas, (d) shadow. Target word pair always has the same decoration like the rest of the words in a tag cloud

3 pairs of words presented in Figure 24 (originally from [3] and [4]) and creating 4 decoration levels of each pair at 4 relative sizes of the larger word. The smaller word of each was always of size 12 pt. Each of these 48 displays created were presented both with and without a background of words resulting in a total of 96 displays that were presented in a unique random order to each participant without time restriction. The displays were presented to the participant first with a tag cloud absent display type in the first block of trials in a randomized order and in the second block of trials, with a tag cloud present display type also displayed in a randomized order.

Because the effect of semantics that can exist when English words are used, words from *Lorem Ipsum*, which is a modified piece of Latin text commonly used as filler in layout designs [2] were used in the experiments in this dissertation. Latin text is assumed to be unfamiliar to the vast majority of readers, while still retaining roughly the word lengths and letter frequencies found in English. Using words from this sample, as well as nonsense words from other sources [3, 4], ensured that users would

Figure 26: An example of the relative sizes of the target pair in the tag cloud present display type: (a) 12 vs. 18, (b) 12 vs. 24, (c) 12 vs. 30, (d) 12 vs. 36

focus on the sizes of the words themselves and not the semantics.

The criteria for choosing the pair of words in this experiment was the following:

1. The two words in the target pair have the same number of letters. For example, the two words 'chadge' and 'aphing' have 6 letters

2. The two words in the target pair must have the same number of ascending and descending letters

3. The two words in the target pair must have the same width when they are at the same size and when using variable length width typeface. The exception is for the wide vs. narrow appearance where one word was wider than the other

The Final Design was a two display type (tag clouds absent, tag clouds present) by 3 appearances (pair of words with no ascending and descending letters, pair of words in which the larger word has wider letters, and pair of words with ascending and descending letters) by 4 decorations (no decoration, boxes or borders, filled areas and shadow) by 4 sizes (typeface 12 pt. vs. typeface 18 pt, typeface 12 pt. vs. typeface 24 pt, typeface 12 pt. vs. typeface 30 pt and typeface 12 pt. vs. typeface 36 pt) (See Figure 27).

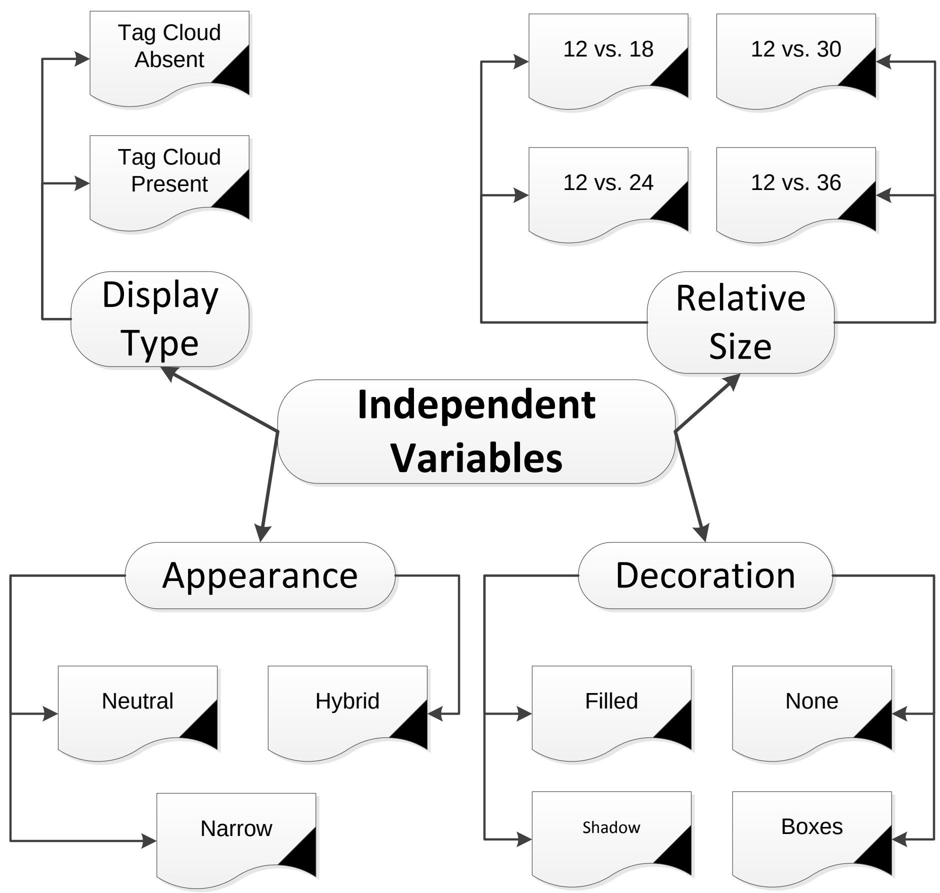


Figure 27: Independent variables in the first experiment

For the tag cloud present display type, 42 words including the pair of words to be compared were distributed randomly in the tag cloud where the smaller word of the target pair (anchor word) was always on the left part of the tag cloud. The frequency of the words in a tag cloud varied based on their sizes. Table 1 shows the frequency of the words in the tag clouds used in this experiment and including the target pair of words.

Table 1: Frequency of word sizes in the tag cloud

Word Size	**Frequency**
Font 12	18 times
Font 18	12 times
Font 24	6 times
Font 30	4 times
Font 36	2 times

For the tag clouds absent display type (Example: Figure 28), the smaller word (anchor word) was on the left of the screen and the bigger word was on the right of the screen. The pair of words to be compared in the tag cloud present display type were marked by red dots (Example: Figure 29) as color is one of the ways to achieve emphasis via contrast on certain elements [55, page 59]. The screen size was always 800x600 pixels.

Figure 28: Screen shot of tag cloud absent display type

Each relative typeface size ratio judgment at each particular relative size was compared with the typeface size ratio squared at that particular relative size. For example, the typeface size ratio squared at relative size 12 vs. 18 is 2.25 ($(18/12)^2 =$

Figure 29: Screen shot of tag cloud present display type

2.25) and the typeface size ratio squared at relative size 12 vs. 36 is 9 ($(36/12)^2 = 9$). The reason why this measurement was chosen is because when the typeface size gets doubled, both the height and width get doubled. This measurement is close to the area of the tag, which is defined as the minimal box around the word in a tag cloud [14].

6.1.3 Procedure

In this experiment, a participant first completed an informed consent screen. The participant was then asked to review the experiment instructions for the first display type, the tag cloud absent display type where there was an example of one trial with instructions on how to perform the first block of the experiment, the tag cloud absent part. After the participant agreed that he read and understood the instructions, the first block started where a participant saw 48 displays for the tag cloud absent block.

Each display is a trial. In each trial, a participant was asked how much bigger the word on the right compared to the word on the left. For each trial, the participant was asked to choose an answer from a continuous scale of 1 to 6 where 1 meant the two target words have the same size and 6 meant the word on the right was 6 times larger than the word on the left by clicking OK and progressing to the next screen. Both the numeric response to the nearest tenth and the time response were recorded. Displays remained visible until a response was completed then a new display followed immediately. After the participant completed the first block of trials, he was presented with the instructions for the second block of 48 trials of the tag cloud present display type.

6.2 Results and Discussion

The size ratio judgment (See Figures 28 and 29) for each display type were submitted to a 2 display type by 3 appearance by 4 decoration by 4 size repeated measures ANOVA. Separate 3-way repeated measures ANOVAs for each display type (tag clouds absent and tag clouds present) were also conducted. All effects were reported as significant at $p < 0.05$.

6.2.1 Analysis of Full Design

There were 2 significant effects that included the display type factor. These will be reported in this subsection. Other significant effects not including the display type factor will be saved for the subsections of reporting results for each display type (tag cloud absent vs. tag cloud present separately).

The main effect of the display type was statistically significant, $F(1, 64) = 6.305$,

$p = 0.015$, $\eta^2 = 0.09$. This indicates that mean size ratio judgment with the tag cloud present ($M = 3.34$) was significantly larger than the mean size ratio judgment with the tag cloud absent display type ($M = 3.21$). The mean values of users' responses at each display types were compared with the mean value of the typeface ratio squared, that is, 5.375 as a test value. Two one sample t-tests were conducted. For the tag cloud absent display type, $t(64) = -22.108$, $p < 0.001$ and for the tag cloud present display type, $t(64) = -20.943$, $p < 0.001$. While both results were statistically different from the test value, the tag cloud present display was better in making the judgments closer to average typeface ratio squared.

Table 2: Mean size ratio judgment for the two displays: tag cloud present and tag cloud absent. The mean typeface ratio squared is the mean value of the typeface ratio squared of the 4 relatives sizes used in this experiment, $((18/12)^2 + (24/12)^2 + (30/12)^2 + (36/12)^2)/4 = 5.375$

Size Ratio Judgment	Mean
Tag Cloud Present	3.206
Tag Cloud Absent	3.339
Typeface Size Ratio Squared	5.375

There was also a significant interaction effect between the display type, appearance and relative size, $F(6, 384) = 2.239$, $p = 0.039$, $\eta^2 = 0.034$. This indicates that the interaction between the appearance and size was different at each display type (Figure 30).

Table 3: Mean values of the appearance for tag cloud absent and tag cloud present display types

Appearance	Mean (Tag Cloud Absent)	Mean (Tag Cloud Present)
No Ascenders and Descenders	3.039	3.211
Narrow vs. Wide	3.457	3.572
Ascenders and Descenders	3.123	3.233

For further information about mean values, please refer to Table 4. To further understand the interaction between display type, appearance and size, Figure 30 illustrates 2 interactions at each level of the display type. In the two interactions, participants always over judge the narrow vs. wide pair of words as an indication that varying letter sizes can lead to a higher size ratio judgment. For each display type, separate post-hoc tests were conducted comparing the wide vs. narrow level with no ascenders and descenders level at each relative size. All these tests showed that the main effect of the appearance was significant. Likewise, separate post-hoc tests comparing the ascenders and descenders level with the wide vs. narrow level at each relevant size revealed significant main effects of appearance in all the tests.

It can also be noted from Figure 30 that the size ratio judgment diverges as the size increases for tag cloud absent display type and it stays the same for the tag cloud present display type for words with ascending and descending letters and the words with no ascending and descending letters. In order to understand this effect, two-way ANOVAs were conducted between the two levels of appearance: words with no ascending and descending letters and words with ascending and descending letters at size 12 vs. size 36 for each display type. For the tag absent display type, the main effect of appearance was statistically significant $F(1, 64) = 9.715$, $p = 0.003$, $\eta^2 = 0.132$, while the main effect of appearance was not significant in the tag cloud present display. *This would seem to be an indication that the effect of ascending and descending letters dissipated after moving to a tag cloud environment.*

The significant interaction display type x appearance x relative size had a very small effect size and did not involve an appreciable qualitative change in the pattern of means between displays with and without a background tag cloud present. The primary effect of the addition of a background tag cloud on relative ratio size judgments of the *comparison word pair was a small increase in the mean size ratio judgment.*

Table 4: Mean values for the interaction between display type, appearance and relative size

Tag Cloud	Appearance	Size	Mean	Typeface Size Ratio Judgment
Absent	No Ascenders & Descenders	12 vs. 18	2.095	2.25
		12 vs. 24	2.61	4
		12 vs. 30	3.418	6.25
		12 vs. 36	4.031	9
	Narrow vs. Wide	12 vs. 18	2.272	2.25
		12 vs. 24	3.049	4
		12 vs. 30	3.843	6.25
		12 vs. 36	4.663	9
	Ascenders & Descenders	12 vs. 18	2.081	2.25
		12 vs. 24	2.696	4
		12 vs. 30	3.489	6.25
		12 vs. 36	4.225	9
Present	No Ascenders & Descenders	12 vs. 18	2.159	2.25
		12 vs. 24	2.796	4
		12 vs. 30	3.55	6.25
		12 vs. 36	4.339	9
	Narrow vs. Wide	12 vs. 18	2.415	2.25
		12 vs. 24	3.146	4
		12 vs. 30	4.009	6.25
		12 vs. 36	4.719	9
	Ascenders & Descenders	12 vs. 18	2.154	2.25
		12 vs. 24	2.876	4
		12 vs. 30	3.548	6.25
		12 vs. 36	4.355	9

6.2.2 Tag Cloud Absent Display Type

Analysis of 3x4x4 Repeated Measures ANOVA was conducted where there were three independent variables: appearance, decoration and relative size.

6.2.2.1 The Effect of Appearance

The main effect of appearance was statistically significant, $F(2, 128) = 72.587$, $p < 0.001$, $\eta^2 = 0.531$. Post-hoc tests were conducted between each two of the three levels of appearance (2x4x4 repeated measures ANOVAs where each time two levels of the appearance factor were picked). These post-hoc tests indicated that the mean size ratio judgment for narrow vs. wide ($M = 3.457$) was significantly larger than the mean size ratio judgment for no ascenders and descenders ($M = 3.039$), $F(1, 64) = 95.903$, $p < 0.001$, $\eta^2 = 0.6$ and ascenders and descenders ($M = 3.123$),

$F(1,64) = 78.021$, $p < 0.001$, $\eta^2 = 0.549$.

In addition, the main effect of appearance was statistically significant when picking the two levels in appearance: no ascenders and descenders and ascenders and descenders. That indicates that the mean size ratio judgment of the target pair of ascenders and descenders ($M = 3.123$) was significantly larger than the mean size ratio judgment of the target pair with no ascenders and descenders ($M = 3.039$), $F(1,64) = 8.916$, $p = 0.004$, $\eta^2 = 0.122$. *This demonstrates that when tag clouds are absent varying the width of the letters can play a role in overestimating the words sizes. In addition, words with ascending and descending letters seem to bias the perception of the viewers and lead them to judge words as bigger.*

6.2.2.2 The Effect of Decoration

There was also a significant main effect of the decoration, $F(3,192) = 3.608$, $p = 0.014$, $\eta^2 = 0.053$. Separate 3x2x4 repeated measures ANOVA tests were conducted where each time two factors of the decoration were picked. The mean size ratio judgment of the boxes ($M = 3.26$) was significantly larger than the mean size ratio judgment of the no decoration group (3.187), $F(1,64) = 6.029$, $p = 0.017$, $\eta^2 = 0.086$. In addition, the mean size ratio judgment of the boxes ($M = 3.26$) was significantly larger than the mean size ratio judgment of the shadow group ($M = 3.157$), $F(1,64) = 10.46$, $p = 0.002$, $\eta^2 = 0.14$. *This clearly shows the minor influence of the boxes around the words in capturing the attention of viewers. This agrees with the findings of Mates, et al [63] that contour can bias the users' judgment of the size of an area.* Furthermore, separate 3x2 repeated measures ANOVA were conducted where each time 3 levels of appearance and 2 levels of decoration were picked at each relative size. These tests showed that the mean size ratio judgments of boxes ($M = 2.837$, $M = 3.858$) were significantly larger than that of no decoration ($M = 2.714$, $M = 3.467$) at relative sizes 12 vs. 24, ($F(1, 64) = 4.841$, $p = 0.031$, $\eta^2 = 0.07$) and 12 vs. 30, ($F(1, 64) = 7.586$, $p = 0.008$, $\eta^2 = 0.106$), respectively. Interestingly, the mean typeface size ratio judgment of no decoration ($M = 2.17$) was significantly larger than that of shadow ($M = 2.06$) at relative size 12 vs. 18, $F(1, 64) = 5.211$, $p = 0.026$, $\eta^2 = 0.075$. In addition, the mean size ratio judgments of boxes ($M = 2.181$, 3.698) were significantly larger than these of shadow ($M = 2.06$, 3.512) at relative size 12 vs. 18, $F(1, 64) = 6.383$,, $p = 0.011$, $\eta^2 = 0.097$ and relative size 12 vs. 30, $F(1, 64) = 9.472$, $p = 0.003$, $\eta^2 = 0.129$. Finally, the mean size ratio judgment of filled areas ($M = 2.187$) was significantly larger than that of shadow ($M = 2.06$), at relative size 12 vs. 18, $F(1, 64) = 8.532$, $p = 0.005$, $eta^2 = 0.118$.

6.2.2.3 The Effect of Size

The main effect of size was significant, $F(3, 192) = 331.799$, $p < 0.001$, $\eta^2 = 0.838$. To further understand the effect of size, separate 3x4x2 repeated measures ANOVAs were conducted where each time 2 levels of the size independent variable were picked. When picking the relative typefaces 12 vs. 18 ($M = 2.243$) and 12 vs. 24 ($M = 2.939$), the main effect of size was statistically significant, $F(1, 64) = 168.738$, $p < 0.001$, $\eta^2 = 0.725$. The main effect of size was also statistically significant when picking relative typefaces 12 vs. 18 ($M = 2.243$) and 12 vs. 30 ($M = 3.702$), $F(1, 64) = 265.155$, $p < 0.001$, $\eta^2 = 0.806$. Likewise, the main effect was statistically significant when picking relative typefaces 12 vs. 18 ($M = 2.243$) and 12 vs. 36 ($M = 4.471$), $F(1, 64) = 342.394$, $p < 0.001$, $\eta^2 = 0.843$. Picking the relative typeface sizes 12 vs. 24 ($M = 2.939$) and 12 vs. 30 ($M = 3.702$) revealed that the main effect of size was statistically significant, $F(1, 64) = 183.884$, $p < 0.001$, $\eta^2 = 0.742$. Picking the relative sizes 12 vs. 24 ($M = 2.939$) and 12 vs. 36 ($M = 4.471$) also showed that the main effect of size was statistically significant, $F(1, 64) = 299.830$, $p < 0.001$, $\eta^2 = 0.824$. Finally, picking the relative sizes 12 vs. 30 ($M = 3.702$) and 12 vs. 36 ($M = 4.471$) revealed that the main effect of size was statistically significant, $F(1, 64) = 238.526$, $p < 0.001$, $\eta^2 = 0.788$. All these post-hoc tests indicated that there were statistically significant differences in the typeface size ratio judgment for each two relative typefaces used in this experiment.

6.2.2.4 The Interaction Effect (Appearance x Size)

The interaction between appearance and size was also significant, $F(6, 384) = 4.252$, $p < 0.001$, $\eta^2 = 0.062$. In order to understand the interaction, separate post-hoc tests were conducted (2x4 ANOVAs where each time 2 levels of appearance factor were picked along with 4 levels of decoration factor at specific relative size). Post-hoc tests indicated that the main effect of the appearance was always significant at each of the 4 relative sizes when the two levels of appearance 'no ascenders and descenders' and 'narrow vs. wide' were chosen. The main effect of appearance was also significant when choosing the 'narrow vs. wide' and 'ascenders and descenders' at each of the 4 relative sizes.

In addition, three 2x4x4 repeated measures ANOVAs were conducted where each time 2 levels of appearance, 4 levels of decoration and 4 levels of size were chosen. Tests revealed that the interaction effect appearance x size was only significant when the Wide vs. Narrow level was chosen with no ascenders and descenders (neutral), $F(3, 192) = 8.308$, $p < 0.001$, $\eta^2 = 0.115$ and with ascenders and decenders (hybrid), $F(3, 192) = 2.972$, $p = 0.033$, $\eta^2 = 0.044$

6.2.3 Tag Clouds Present Display type

Size ratio judgments for the displays with a tag cloud present were submitted to a 3x4x4 repeated measures ANOVA, including 3 types of appearance (no ascenders and descenders, narrow-wide letters, ascenders and descenders), 4 types of decorations (no decoration, boxes, filled areas and shadow), and 4 font sizes for the target word of the comparison pair (18, 24, 30, 36).

6.2.3.1 The Effect of Appearance

The main effect of appearance was statistically significant, $F(2, 128) = 45.674$, $p < 0.001$, $\eta^2 = 0.416$. In order to further understand the effect of appearance, separate 2x4x4 repeated measures ANOVAs were conducted where every time 2 levels of the appearance factor were picked. The main effect of appearance was significant when no ascenders and descenders and narrow vs. wide factors were picked, $F(1, 64) = 53.378$, $p < 0.001$, $\eta^2 = 0.445$. This indicates that the mean size ratio judgment of narrow vs. wide ($M = 3.572$) was significantly larger than the mean size ratio judgment of no ascenders and descenders ($M = 3.211$). Furthermore, the main effect of appearance was significant when ascenders and descenders and narrow vs. wide were chosen, $F(1, 64) = 53.708$, $p < 0.001$, $\eta^2 = 0.456$. This indicates that the mean size ratio judgment of narrow vs. wide ($M = 3.572$) was significantly larger than the mean size ratio judgment of ascenders and descenders ($M = 3.233$).

6.2.3.2 The Effect of Size

The main effect of size was significant, $F(3, 192) = 285.501$, $p < 0.001$, $\eta^2 = 0.817$. For further understanding of the effect of size, separate 3x4x2 repeated measures ANOVAs were conducted where each time 2 levels of the size independent variable were picked. When picking the relative typefaces 12 vs. 18 ($M = 2.149$) and 12 vs. 24 ($M = 2.785$), the main effect of size was statistically significant, $F(1, 64) = 154.4$, $p < 0.001$, $\eta^2 = 0.707$. The main effect of size was also statistically significant when picking relative typefaces 12 vs. 18 ($M = 2.149$) and 12 vs. 30 ($M = 3.583$), $F(1, 64) = 292.483$, $p < 0.001$, $\eta^2 = 0.820$. Likewise, the main effect was statistically significant when picking relative typefaces 12 vs. 18 ($M = 2.149$) and 12 vs. 36, $F(1, 64) = 392.929$, $p < 0.001$, $\eta^2 = 0.860$. Picking the relative typeface sizes 12 vs. 24 ($M = 2.785$) and 12 vs. 30 ($M = 3.583$) revealed that the main effect of size was statistically significant, $F(1, 64) = 269.074$, $p < 0.001$, $\eta^2 = 0.808$. Picking the relative sizes 12 vs. 24 ($M = 2.785$) and 12 vs. 36 ($M = 4.306$) also showed that the main effect of size was statistically significant, $F(1, 64) = 389.704$, $p < 0.001$, $\eta^2 = 0.859$. Finally, picking the relative sizes 12 vs. 30 and 12 vs. 36 ($M = 4.306$) revealed that the main effect of size was statistically significant, $F(1, 64) = 247.246$, $p < 0.001$, $\eta^2 = 0.749$. All these post-hoc tests indicate that there are statistically significant differences in the typeface size ratio judgment for each two relative typefaces used in this experiment. Figure 32 shows that overall participants started having nearly correct judgment and then they were under-judging the ratio size and that under-judgment was increasing as the target word gets bigger in size.

The empirical results (Figures 31 and 32) demonstrate general under-perception of typeface size ratio judgment, but as the size of the target word approaches the size of the anchor word, the typeface size ratio judgment gets closer to the correct size ratio squared. In addition, the mean typeface size ratio judgments look linear. The fact that the data is not upward curving like the actual stimuli size ratio squared values indicates that participants were most likely weighting a single width dimension more than the less perceptually salient height dimension, and less than the product of the 2 dimensions (i.e., the 2D size).

6.2.3.3 The Interaction Effect (Appearance x Size)

The interaction between appearance and size was also significant, $F(6, 384) = 2.286$, $p = 0.035$, $\eta^2 = 0.034$ when ascenders and descenders and narrow vs. wide were chosen, $F(3, 192) = 3.485$, $p < 0.017$, $\eta^2 = 0.052$. In order to explore the effect of the appearance at each relative typeface size, separate post-hoc tests were conducted (2x4 ANOVAs where each time 2 levels of appearance factor are picked along with 4 levels of decoration factor at specific relative size). Post-hoc tests indicated that the main effect of the appearance was always significant at each of the 4 relative sizes when the two levels of appearance 'no ascenders and descenders' and 'narrow vs. wide' were chosen. The main effect of appearance was also significant when choosing the 'narrow vs. wide' and 'ascenders and descenders' at each of the 4 relative sizes.

6.2.3.4 The Interaction Effect (Decoration x Size)

The interaction between decoration and size was likewise significant, $F(9, 576) = 2.197$, $p = 0.021$, $\eta^2 = 0.033$ (See Figure 33 B) which was not the case in the tag cloud absent display type. Looking at Figure 33, there was a notable divergence between the values of the decoration factor at the relative sizes 12 vs. 18 and 12 vs. 24.

Table 5: Mean values of the decoration for tag cloud present display types

Decoration	Relative Size	Mean	Size Ratio Judgment
No Decoration	12 vs. 18	2.207	2.25
	12 vs. 24	2.819	4
	12 vs. 30	3.704	6.25
	12 vs. 36	4.465	9
Boxes	12 vs. 18	2.298	2.25
	12 vs. 24	3.009	4
	12 vs. 30	3.683	6.25
	12 vs. 36	4.488	9
Filled Areas	12 vs. 18	2.285	2.25
	12 vs. 24	3.005	4
	12 vs. 30	3.7	6.25
	12 vs. 36	4.426	9
Shadow	12 vs. 18	2.182	2.25
	12 vs. 24	2.925	4
	12 vs. 30	3.722	6.25
	12 vs. 36	4.507	9

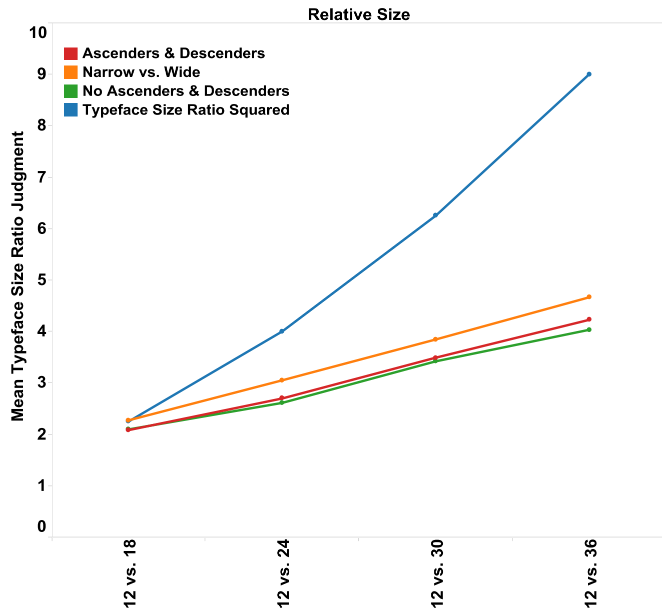
Separate 3x2x4 repeated measures ANOVAs where conducted where each time the three levels of appearance factor and 2 levels of decoration were picked at relative sizes 12 vs. 18 and 12 vs. 24. The main effect of decoration was significant when no decoration ($M = 2.819$) and boxes ($M = 3.009$) were chosen at relative size 12 vs. 24, $F(1, 64) = 10.052$, $p = 0.002$, $\eta^2 = 0.136$. In addition, the main effect of decoration was significant when no decoration ($M = 2.819$) and filled ($M = 3.005$) were chosen at relative size 12 vs. 24, $F(1, 64) = 8.904$, $p = 0.004$, $\eta^2 = 0.122$. The main effect of decoration was also significant when boxes ($M = 2.298$) and shadow ($M = 2.182$) were chosen at relative size 12 vs. 18, $F(1, 64) = 5.814$, $p = 0.019$, $\eta^2 = 0.083$. In addition, the main effect of decoration was significant when filled ($M = 2.285$) and shadow ($M = 2.182$) were chosen at the relative size 12 vs. 18, $F(1, 64) = 5.673$, $p = 0.02$, $\eta^2 = 0.081$.

6.2.3.5 The Interaction Effect (Appearance x Decoration x Size)

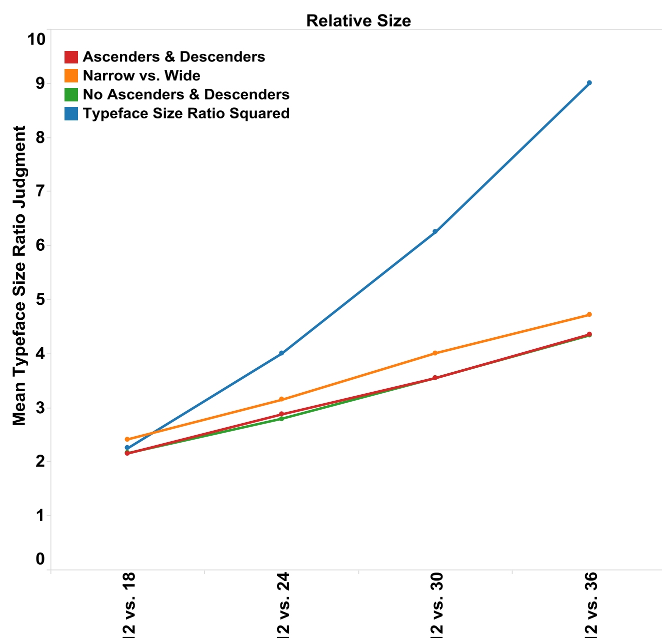
The interaction appearance x decoration x size was significant, $F(18, 1152) = 1.766$, $p = 0.025$, $\eta^2 = 0.027$. In order to understand this interaction, it was broken into four 3x4 repeated measures ANOVAs at each level of the decoration independent variable (Figures 34, 35, 36 and 37). This clearly shows that participants were over-judging the typeface size ratio when viewing pairs of words that vary in the width of the letters.

6.2.4 One Sample T-Tests

One sample t-tests are used to compare the mean of a sample to a particular value. The purpose of conducting the t-tests in the present experiment was to compare the mean size ratio judgment at each relative size for all the participants with the value of the typeface ratio squared at that particular size. The typeface size ratio squared at each particular size was calculated by dividing the target word typeface size by the anchor word typeface size and squaring the result. The reason why the typeface ratio squared was chosen is because when doubling the size of the letter, both the height and width of that letter are doubled. The mean size ratio judgment at each relative size was compared to typeface ratio squared at that particular size for both tag cloud present and tag cloud absent displays. One sample t-tests revealed that at relative size 12 vs. 30 the mean typeface size ratio judgment was not statistically significant than the typeface size ratio squared at that particular size, (2.25) for tag absent ($M = 2.15$) and tag present ($M = 2.243$) display types. For the rest of the relative typefaces used in the present experiment (12 vs. 24, 12. vs. 30 and 12. vs. 36), comparing the mean size ratio judgment of participants for each relative typeface with the typeface ratio squared for that particular size always revealed statistically significant effects. For the tag cloud absent display type, one sample t-tests indicated that the relative typeface size ratio judgments of sizes 12 vs. 24 ($M = 2.785$), 12 vs. 30 ($M = 3.583$) and 12 vs. 36 ($M = 4.306$) were significantly lower than the typeface size ratio squared at these particular sizes (4, 6.25, and 9), ($t(64) = -12.472$, $p < 0.001$), ($t(64) = -22.881$, $p < 0.001$) and ($t(64) = -37.105$, $p < 0.001$) respectively. For the tag cloud present display type, one sample t-tests indicated that the relative typeface size ratio judgments of sizes 12 vs. 24 ($M = 2.9394$), 12 vs. 30 ($M = 3.702$) and 12 vs. 36 ($M = 4.4712$) were significantly lower than the typeface size ratio squared at these particular sizes (4, 6.25, and 9), ($t(64) = -10.730$, $p < 0.001$), ($t(64) = -21.741$, $p < 0.001$) and ($t(64) = -34.645$, $p < 0.001$) respectively. Figure 38 shows the interaction between the display type and relative size and from the figure, it is noticeable that at the relative size 12 vs. 18, the overall judgment was very close to size ratio squared while the divergence gets bigger as the relative size increases.



(A) Tag cloud absent display type



(B) Tag cloud present display type

Figure 30: The interaction between appearance and size in tag cloud absent and tag cloud present display types. Participants were greatly influenced by varying the width of the letters in the target pair. There was also a minor influence of ascenders and descenders in the tag cloud absent display type at relative size 12 vs. 36 and this was dissipated when the tag cloud was added

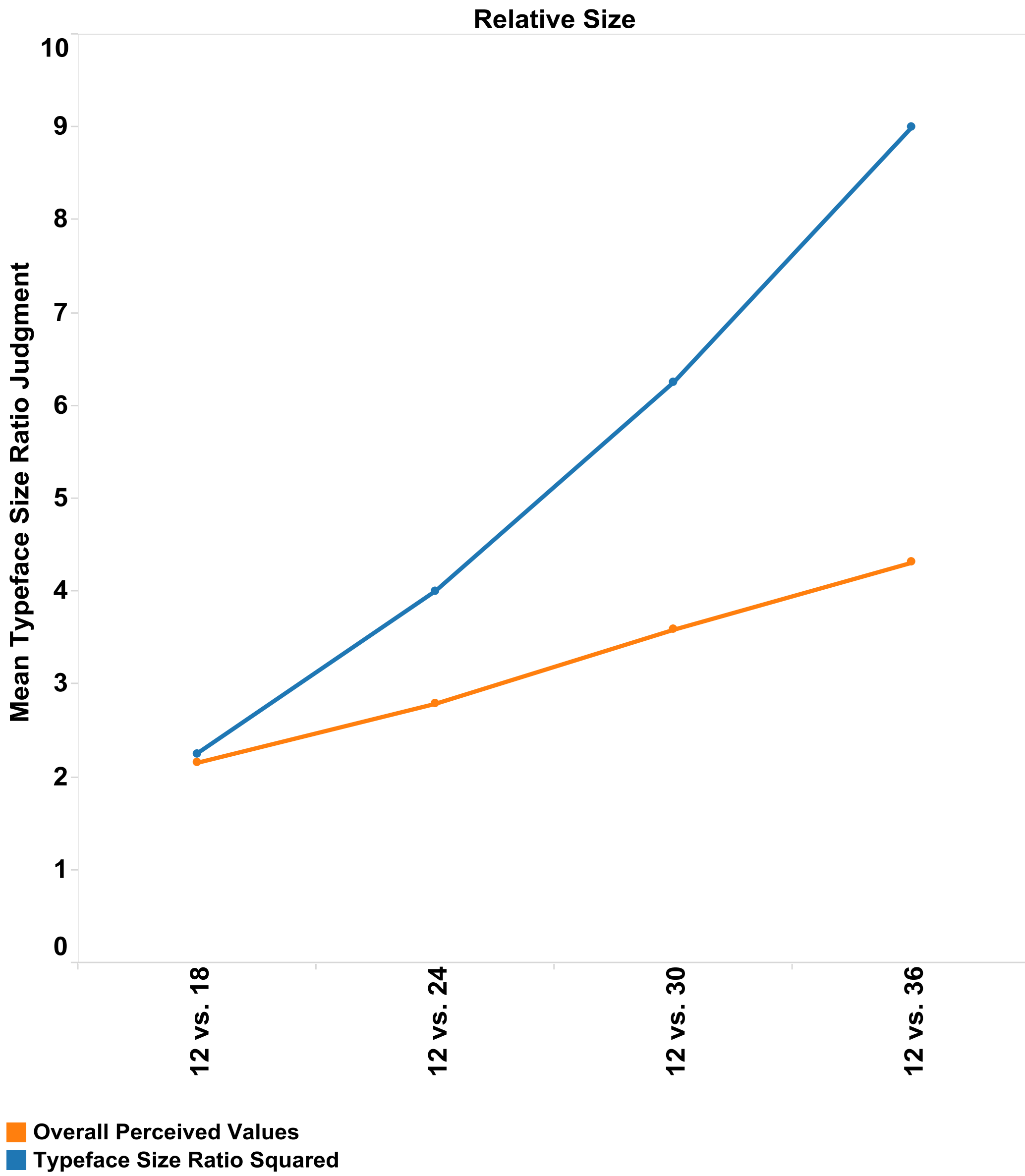


Figure 31: Relative mean size ratio judgment for the tag cloud absent display type as a function of target word typeface. The blue line is the typeface ratio squared for relative sizes used in this present experiment. The figure clearly shows that the typeface size ratio judgment gets closer to typeface size ratio squared as the size of the target word gets closer to the size of the anchor word

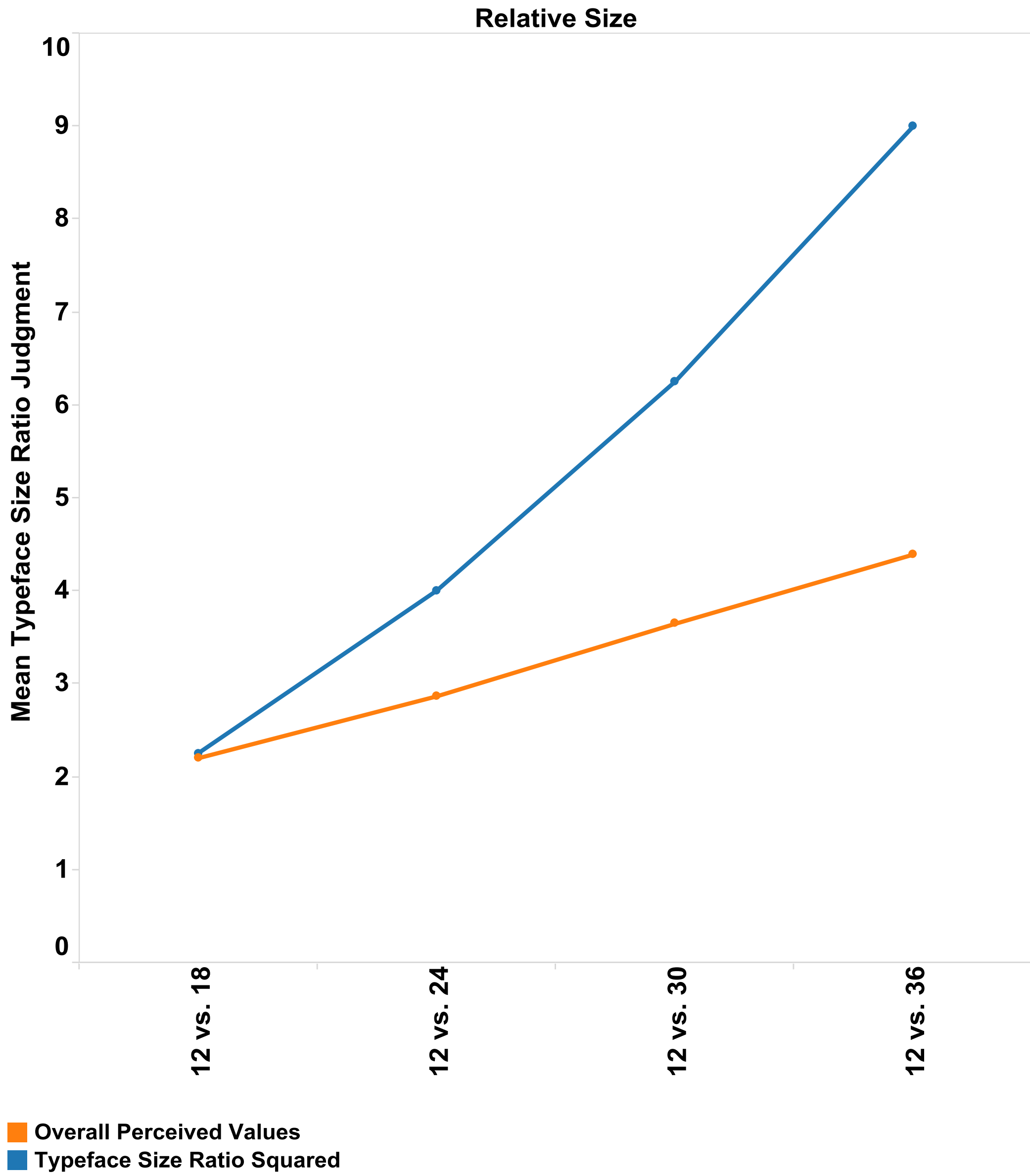
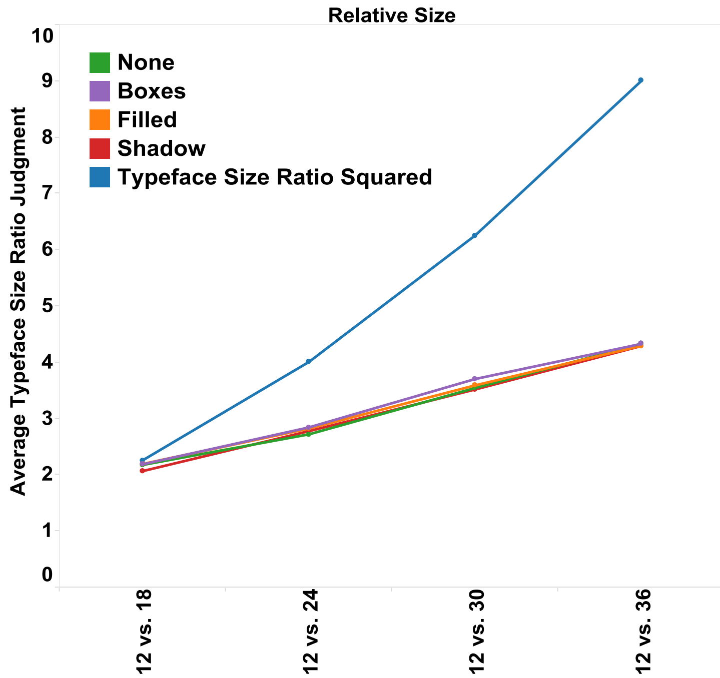
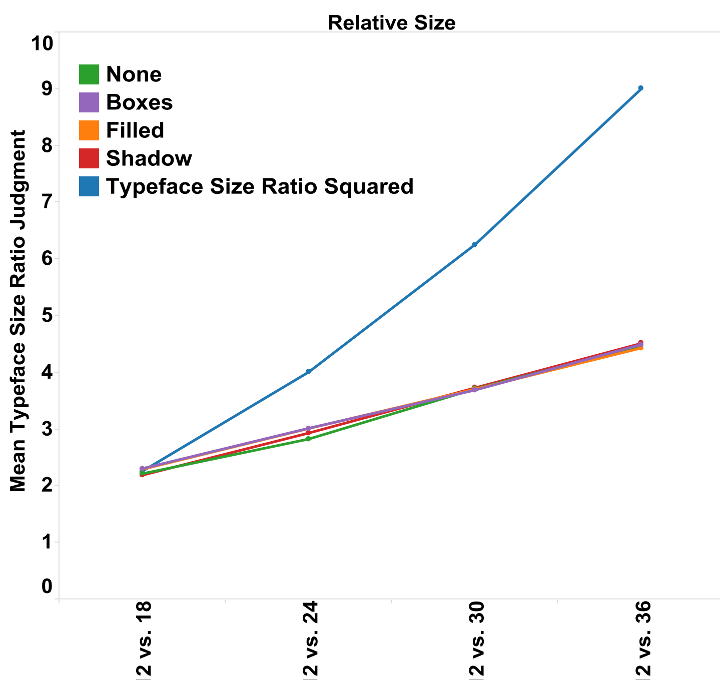


Figure 32: Relative mean size ratio judgment for the tag cloud present display type as a function of target word typeface. The blue line is the typeface ratio squared for relative sizes used in this present experiment. The figure shows that the mean typeface size ratio under-judgment increases as the target word size is increased



(A) Tag Cloud Absent Display Type



(B) Tag Cloud Present Display Type

Figure 33: The interaction between decoration and size independent variables for the tag cloud absent and tag cloud present display types. Significant minor influences were found

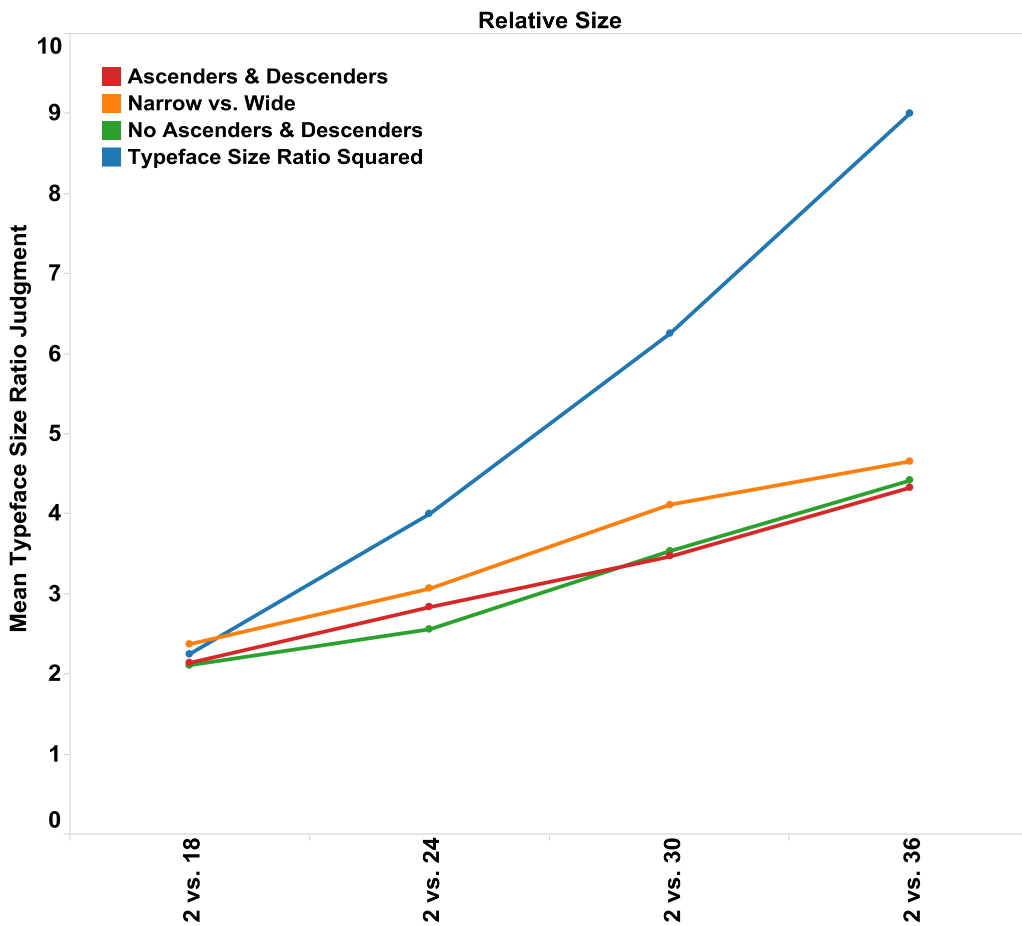


Figure 34: The interaction between appearance and size independent variables for the tag cloud present display type and the no decoration group

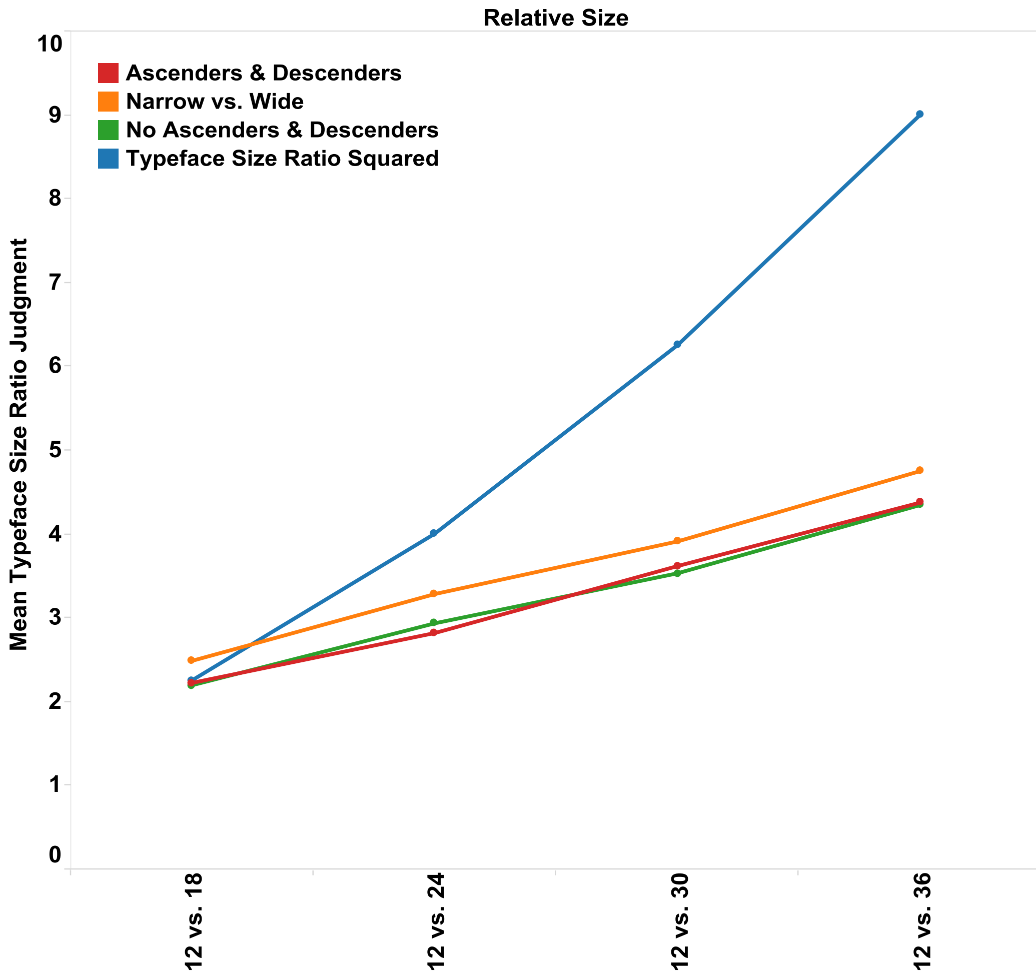


Figure 35: The interaction between appearance and size independent variables for the tag cloud present display type and the boxes group

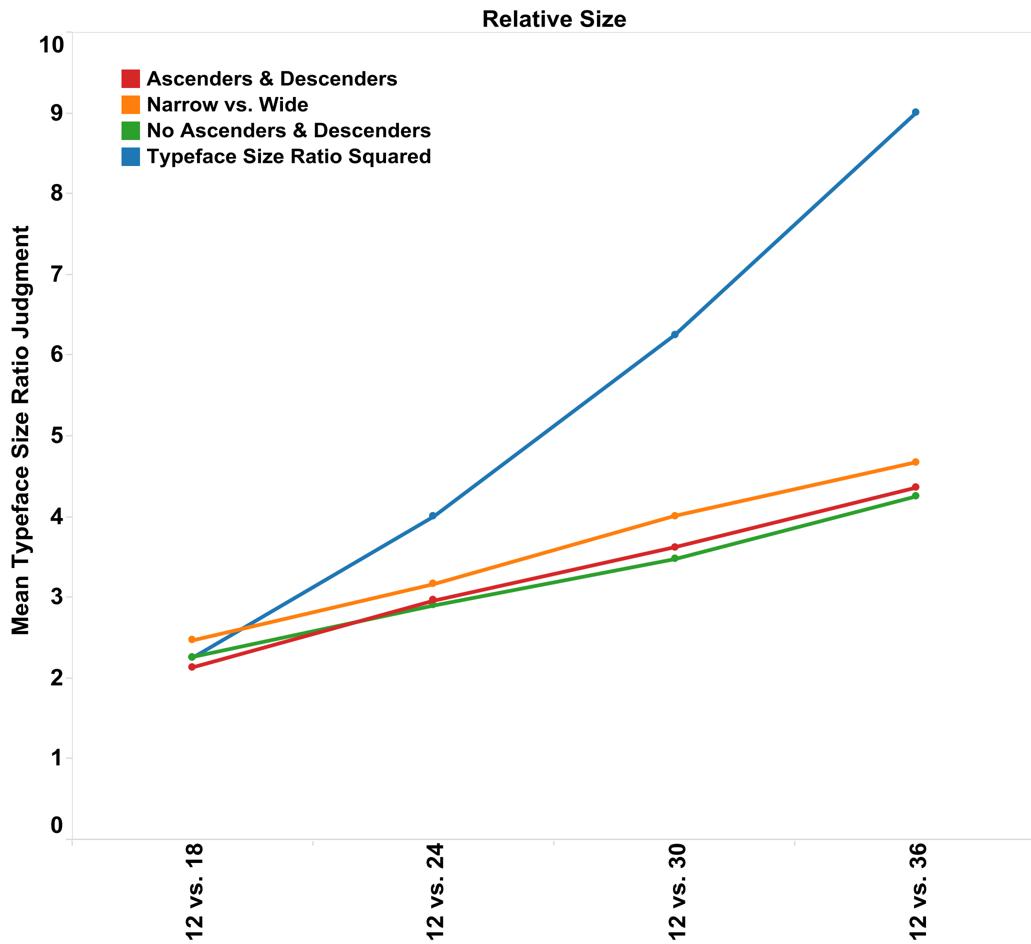


Figure 36: The interaction between appearance and size independent variables for the tag cloud present display type and the filled areas group

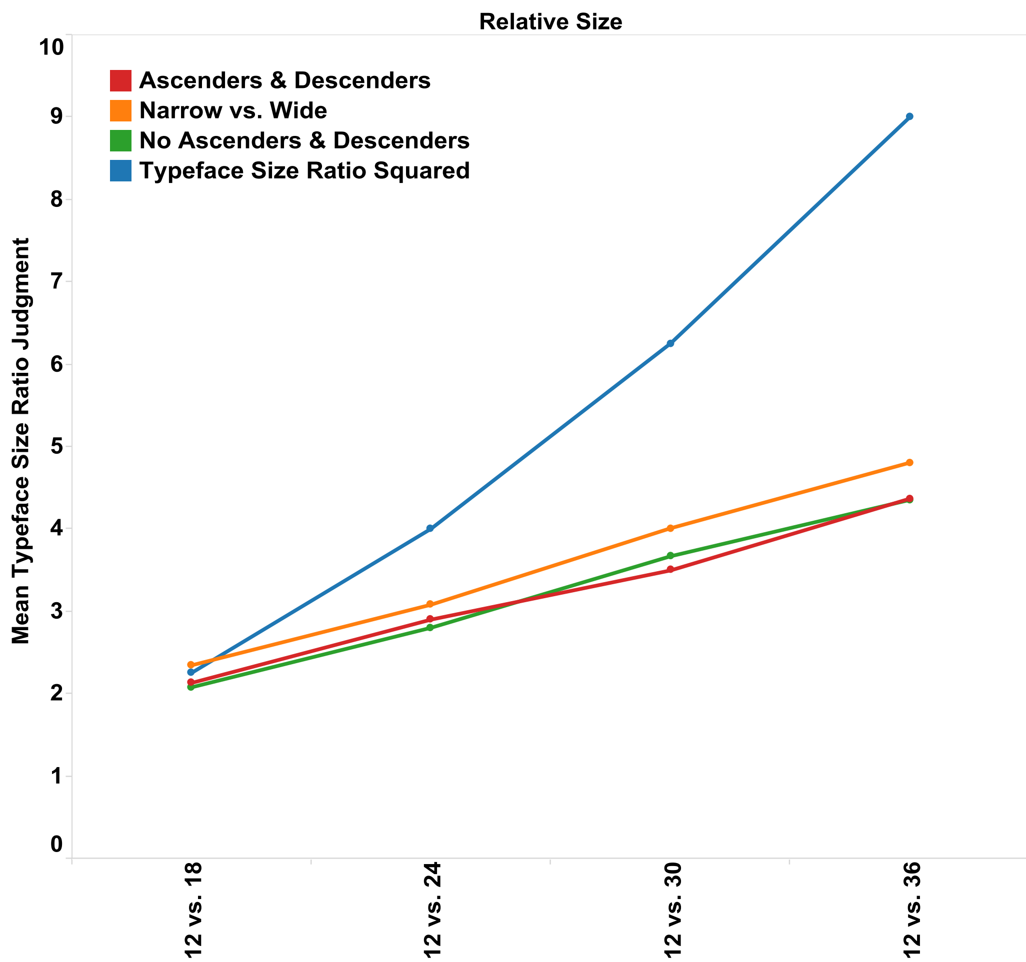


Figure 37: The interaction between appearance and size independent variables for the tag cloud present display type and the shadows group

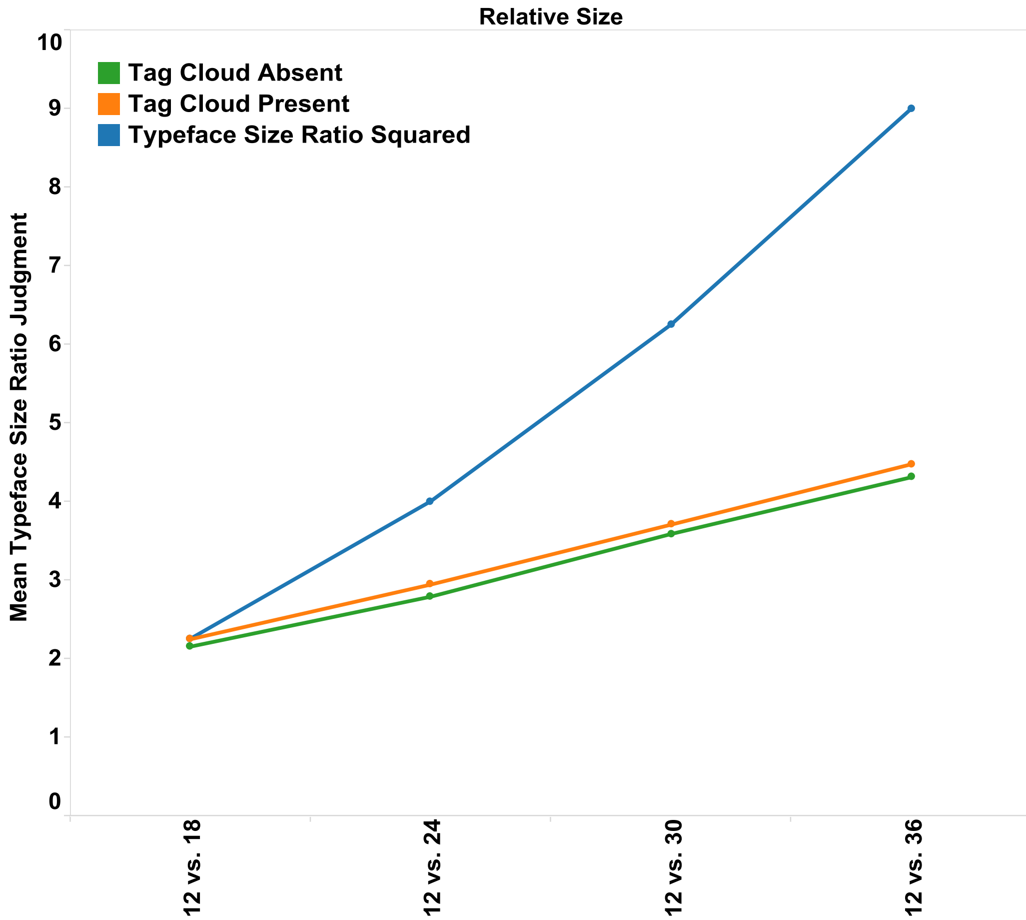


Figure 38: The interaction between display type and relative size. The mean typeface size ratio judgment for tag cloud absent and tag cloud present display types are represented by green and orange lines, respectively. The blue line is the typeface ratio squared size for relative sizes used in this present experiment

Figure 39: Tag cloud of chapter 6 generated by `www.tagcrowd.com`

CHAPTER 7: EXPERIMENT 2: THE INFLUENCE OF TYPEFACE AND LOCATION ON JUDGMENT OF RELATIVE SIZE

The writer operates at a peculiar crossroads where time and place and eternity somehow meet. His problem is to find that location

— Flannery O'Connor

This study is an extension to the previous study towards better understanding of users' perception of size ratio judgment in tag clouds. The focus on this study is on the perception of size ratio judgment given the typeface style (e.g. regular vs. bold) and the location of the words in a tag cloud (e.g. upper-left vs. upper-right). In addition, because the general under-judgment in experiment 1 might be due to an anchoring effect and because the anchor word used was of size 12 pt., the decision was to increase the size of the target word to 18 pt.

Along the typeface size, the typeface style is one of the most influential characteristics that captures the attention of the viewer in a tag cloud [15]. In addition, tag cloud designers control the locations of the words. The location of the tags in a tag cloud received attention in research [15, 73, 78, 94]. In this experiment, participants are asked to judge the typeface size ratio of a target pair. The target pair consists of two nonsense words obtained from [3] on Dec 4, 2012. The purpose of choosing nonsense words is to make sure that participants are focused on the relative sizes of the words themselves and not on the semantics of the words. The smaller word in the target pair is called the anchor word and the bigger word is called the target word.

7.1 Method

7.1.1 Participants

Data for this study was collected from 78 participants of UNC Charlotte students. Participants performed the experiment in the lab and they were informed that the one who performed the most accurately would be awarded a $25 Starbucks gift card. The purpose of that was to bring participants into focus on the task they were doing instead of putting random answers. Two of the participants entered the same answer for each question and thus, they were removed from the pool of participants. Two other participants reported that they did put random responses and thus, their responses were removed and not included in the analysis for the study, as well. The rest of the participants were 50 males and 24 females. Participants were only asked about their age and gender. Providing email was optional so that to reach a participant in case he wins. Two participants did not report their ages. The ages of the rest of 72 participants ranged from 18 to 50 with an average of 22.35 years old. Participants reported having 20/20 perfect or corrected to perfect vision besides being able to communicate comfortably in written English.

7.1.2 Materials

Each participant viewed 96 screens each of which had a question asking him to choose by how much one word (target word) is bigger than another word (anchor word). The smaller word (anchor word) was always presented in 18 pt., while the bigger word (target word) varied in size: 24, 30 and 36. Target pairs of words in this experiment exist in a display of a tag cloud and not just by themselves on the screen. The design involves a manipulation of 4 independent variables (IVs) and 2 control variables (CVs), as follows:

- Typeface style. Half of the trials had words in a tag cloud in a regular typeface, while the rest of the trials had words in a bold typeface. All the words in the tag cloud have the same typeface style manipulation as the target pair. Figure 40 shows two screen shots where the typeface style was manipulated

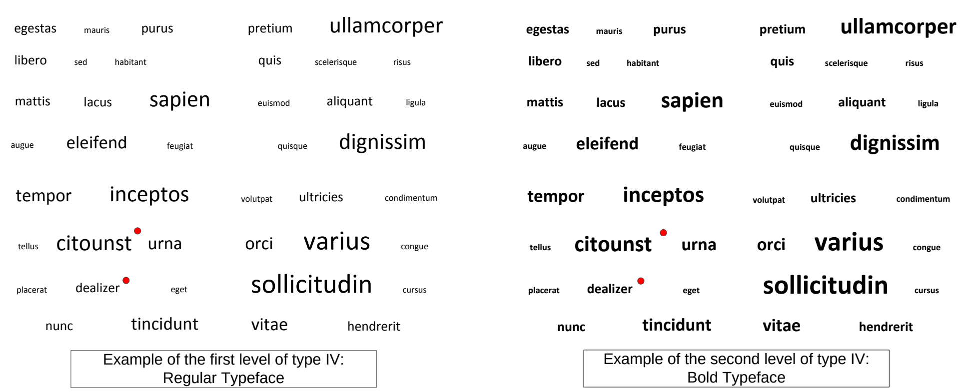


Figure 40: An example of a tag cloud where typeface style IV is manipulated. The left tag cloud has words in a regular typeface while the right tag cloud has words with bold typeface

- Horizontal. The location of the target pair was varied between the left and right portions of the tag cloud. Figure 41 provides an example of two tag clouds used in the experiment where the horizontal independent variable was manipulated
- Vertical. The location of the target pair varied between the upper and lower portions of the tag cloud. Figure 42 is an example of two tag clouds used in this experiment where the vertical independent variable was manipulated
- Manipulating the relative size of target pairs used in the experiment. The smaller word was always 18 pt. The larger word of the pair did vary in size:

Figure 41: An example of two tag clouds where the horizontal IV is manipulated. In the left tag cloud, the target pair is located in the left part of the tag cloud while in the right tag cloud the target pair is located in the right portion of the tag cloud

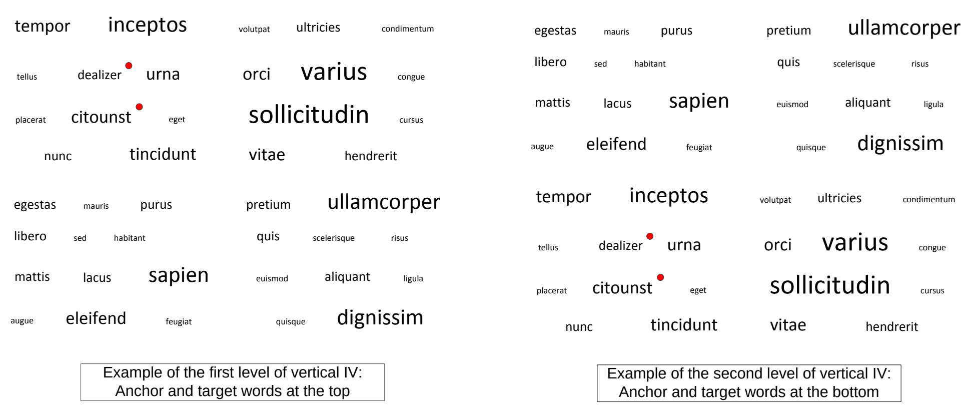


Figure 42: An example of two tag clouds used in this experiment where the vertical IV was manipulated. In the left tag cloud, the words in the tag cloud including the target pair were located in the upper part of the tag cloud while in the right tag cloud, the words in the tag cloud including the target pair were located in the lower part of the tag cloud

24, 30 and 36 pt. Figure 43 shows the relative typeface sizes used in this experiment.

In this experiment, there were two control variables (CVs):

- Changing the pair of words used in the experiment. In this experiment, two pairs of words were chosen:

    First pair: dealizer (smaller word) vs. citounst

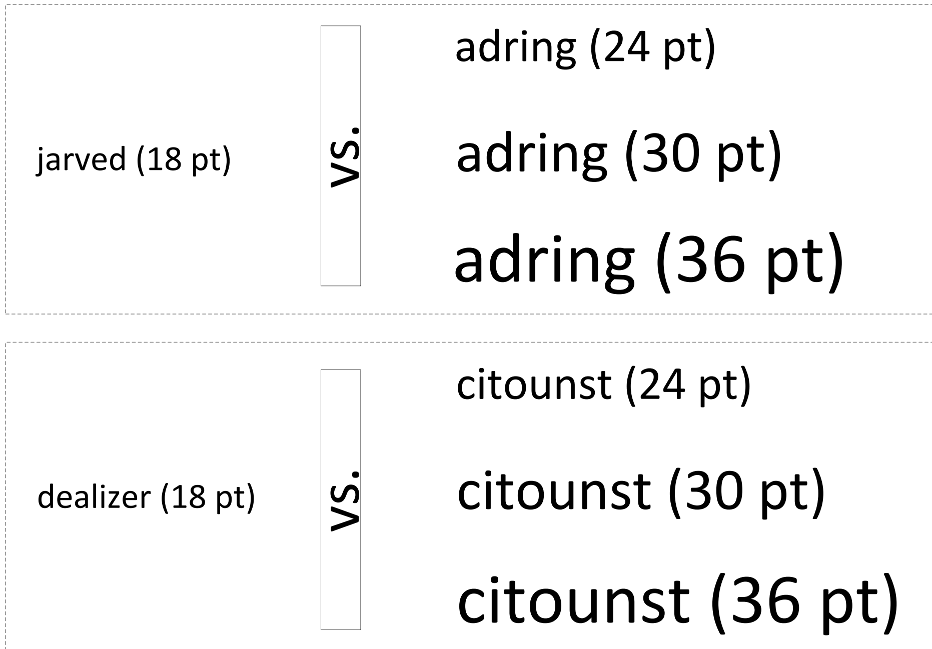


Figure 43: Typeface sizes used in the present experiment. The smaller word in the target pair always comes in size 18. The bigger word varies in size: 24, 30 and 36

Second pair: jarved (smaller word) vs. adring

The two pairs were originally taken from from [3]. Figure 45 is an example of two screen shots where the pair of words is manipulated.

- Swapping the locations of the pair of words to be compared (Example: Figure 46)

The 4 independent variables and the 2 control variables were manipulated to create 96 pairs of words where each pair was in a tag cloud. The dependent variable was the size ratio judgment for the target pair of words in a tag cloud.

Words in the tag cloud (excluding the target pair) varied in their frequency. The frequency of the size of the words in the tag cloud is as in Table 6.

The two target pairs of words were selected according to the following criteria:

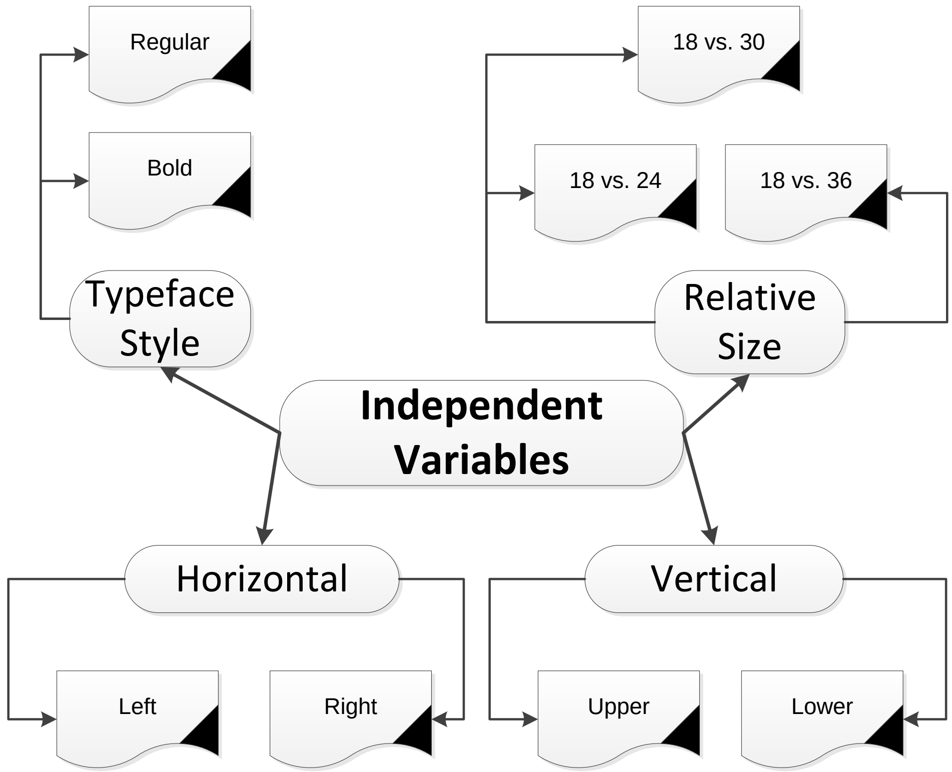


Figure 44: Independent variables in experiment 2

Table 6: The frequency of *Lorem Ipsum* words in the tag cloud

Typeface Size	**Frequency**
Typeface 12 pt.	17 times
Typeface 18 pt.	11 times
Typeface 24 pt.	6 times
Typeface 30 pt.	4 times
Typeface 36 pt.	2 times

(a) In each pair, the two words must have the same total number of ascenders and descenders.

(b) In each pair, the two words must have roughly the same width when they are at the same size and when variable width font is used. For example, the two words in each target pair have almost the same width (Example: Figure 47)

(c) In each pair, the two words much have the same number of letters. For example, the two words dealizer and citounst have 8 letters each.

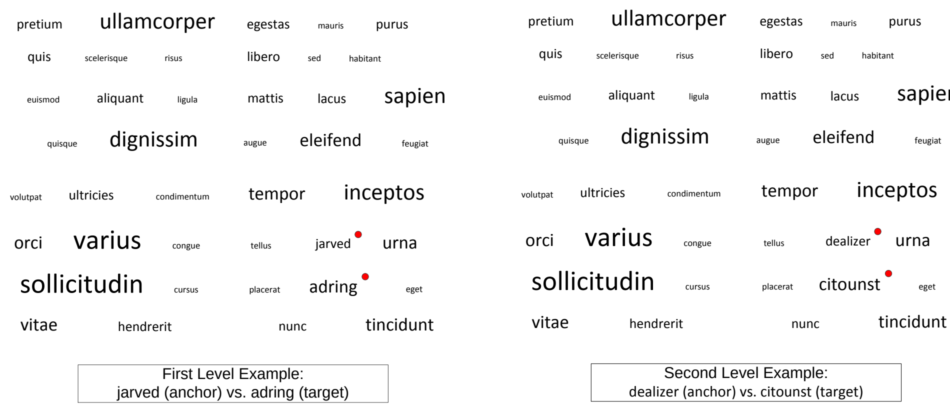


Figure 45: An example of manipulating pair of words CV used in the present experiment. In the left tag cloud the first pair of words (jarved vs. adring) was used while in the right tag cloud, the second pair of words was used (dealizer vs. citounst)

The rest of the words in the tag cloud are taken from *Lorem Ipsum*, that is a modified piece of Latin text commonly used as filler in layout designs. Latin text is assumed to be unfamiliar to the vast majority of readers, while still retaining roughly the word lengths and letter frequencies found in English. The reason to choose non-sense words and a Latin text is to eliminate the effect of the semantics since the interest is in how users perceive the size of the words in tag cloud without including any semantic effect. The resolution of the screen was 800x600 pixels.

Because changing the words in the neighborhood of the target pair might add a bias to the judgment, mirroring was performed, the purpose of which was to retain the words in the neighborhood of the target pair to avoid any bias in the perception. Mirroring was performed as follows:

First, dividing the tag cloud into 4 quadrants where there was a number of words in each quadrant and where the target and anchor words were located in one quadrant. The four quadrants were: Upper Left (UL), Upper Right (UR), Lower Left (LL),

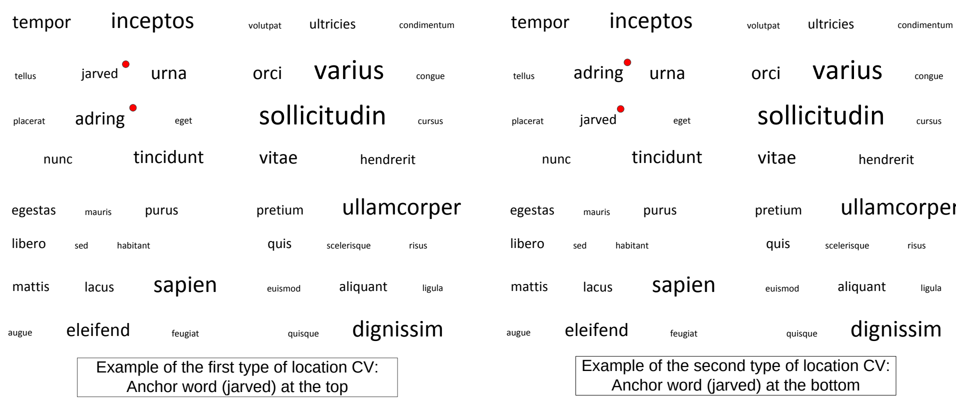


Figure 46: An example showing the manipulation of the locations of the target pair. In one manipulation (left screen) the smaller word (anchor word) is at the top while the other word is at the bottom. In the screen to the right, the locations of words are swapped where the anchor word is at the bottom and the target word is at the top

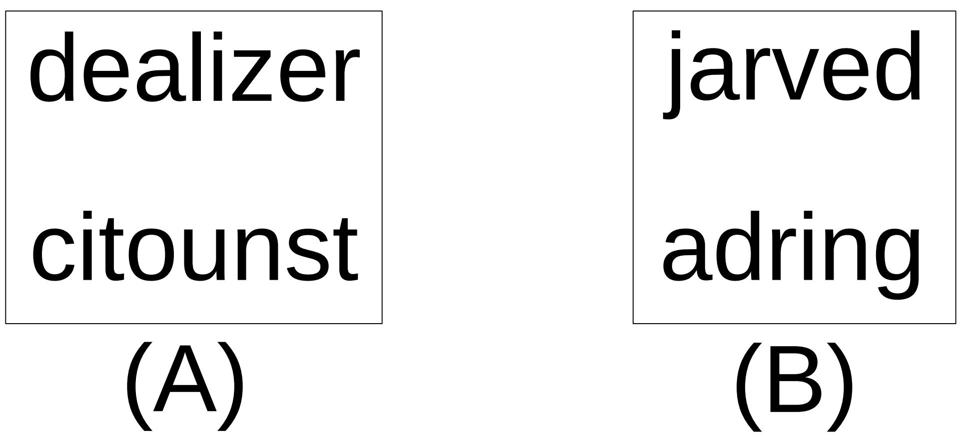


Figure 47: The target pairs chosen in experiment 2: the two words in each pair have the same number of letters, they have the same total number of ascenders and descenders and when they have the same length, their width is roughly the same

Lower Right (LR) (Figure 48). Second, varying the location of the word pairs by quadrant in the tag cloud. Figure 49 shows how the contents of the quadrants in a tag cloud change as the horizontal and vertical IVs are manipulated. Figure 50 is an example of how a tag cloud looks like as the vertical and horizontal locations of the target and anchor words are manipulated.

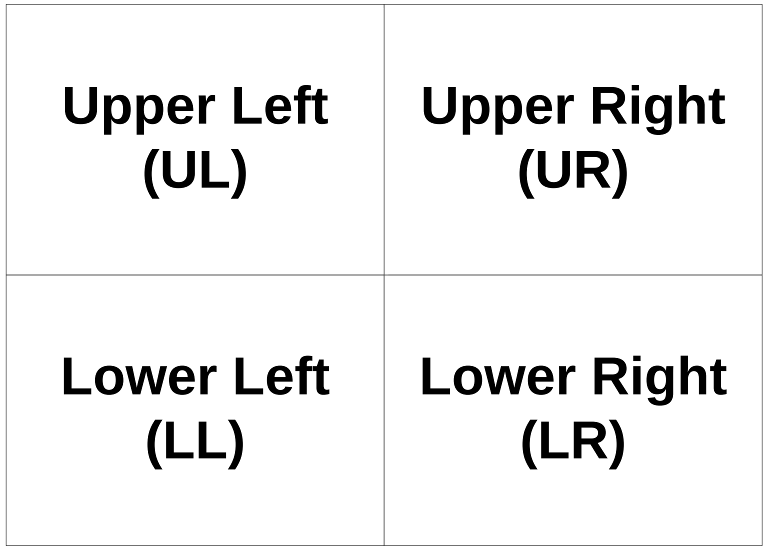


Figure 48: Quadrants in a tag cloud. The lines are invisible (imaginary) and they don't appear in the design

Each relative typeface size ratio judgment at each particular relative size was compared with the typeface size ratio squared at that particular relative size. For example, the relative typeface size ratio at relative size 18 vs. 24 was compared with the typeface size ratio squared at that relative size which is 1.778 ($(24/28)^2 = 1.778$). The reason why this measurement was chosen is because when the typeface size gets doubled, both the height and width are doubled. This measurement is close to the area of the tag, which is defined as the minimal box around the word [14].

7.1.3 Procedure

In the present experiment, participants first completed an informed consent. After that, a participant was asked to review the experiment instructions where there was one example on how to perform a trial. After the participant agreed that he read an understood the instructions, the experiment started where the participant saw 96 displays presented in a random order. Each one of the displays was a trial. In each

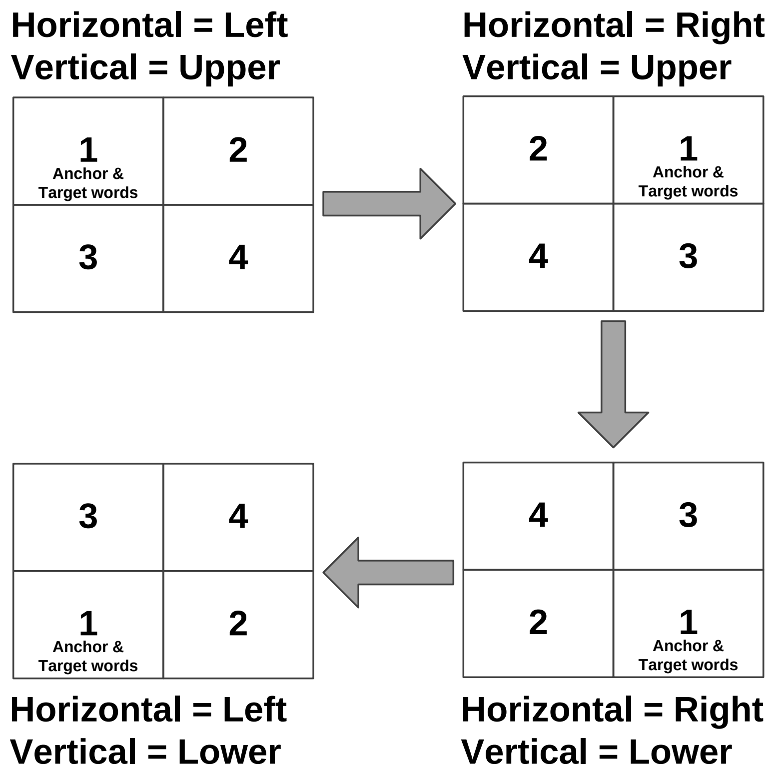


Figure 49: Changing vertical and horizontal locations of anchor and target words in a tag cloud. Every time the horizontal and vertical IV were manipulated, blocks change their location within a tag cloud.

trial, the participant was asked how much bigger one word to another. For each trial, the participant would choose an answer from a continuous scale that runs from 1 to 6 and hit OK to go to the next trial. Both the numeric and time responses (to the nearest tenth) were recorded. The screen stayed visible till the participant recorded an answer and then a new display followed immediately.

7.2 Results and Discussion

Before conducting statistical analysis, words that have the same typeface style, horizontal, vertical and relative size IVs were averaged and the average constituted the dependent variable. The average size ratio judgments were submitted to a 2 typeface style (regular vs. bold) by 2 horizontal location (left vs. right) by 2 vertical

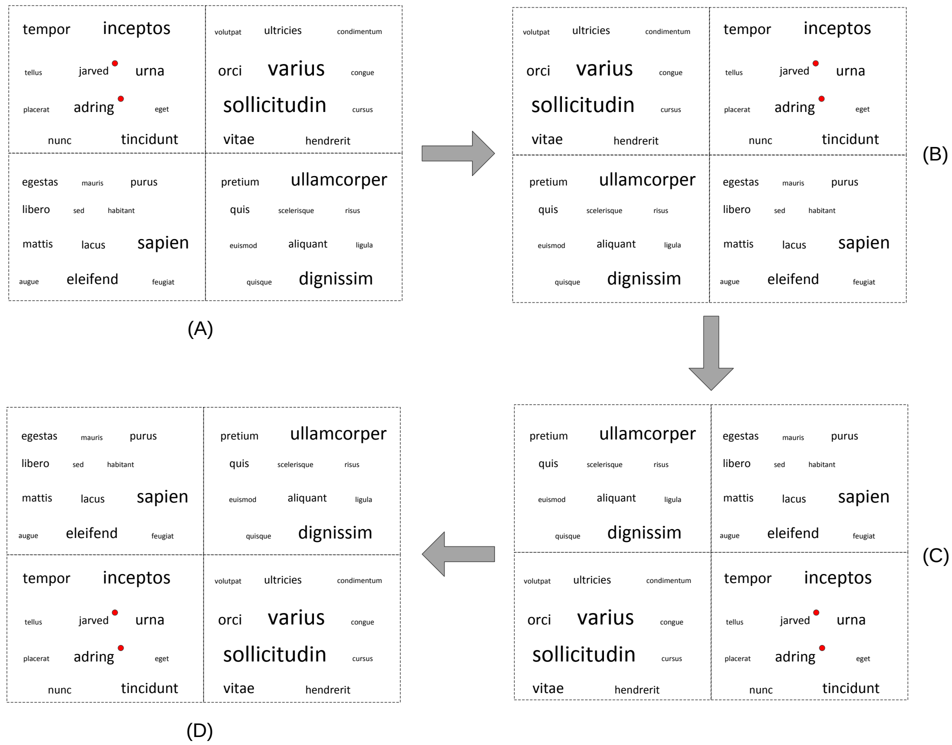


Figure 50: An example of how a tag cloud looks like after the vertical and horizontal IVs of the target pair are manipulated. All lines in the figure are imaginary and they don't show in the actual tag cloud

location (upper vs. lower) by 3 relative sizes (18 vs. 24, 18 vs. 30 and 18 vs. 36) repeated measures ANOVA. All the effects were reported as significant at $p < 0.05$.

7.2.1 The Effect of Typeface Style

The main effect of typeface style was significant, $F(1, 73) = 9.294$, $p = 0.003$, $\eta^2 = 0.113$, indicating that the mean size ratio judgment for target pairs in bold typeface ($M = 2.828$) was significantly larger than the mean size ratio judgment for the target pairs in regular typeface ($M = 2.772$). That result agrees with the finding of Bateman, *et al* that boldness had a strong visual effect in capturing the attention of the viewer.

7.2.2 The Effect of Size

The main effect of size was significant, $F(2, 146) = 141.964$, $p < 0.001$, $\eta^2 = 0.780$. Figure 51 shows the participants' judgment of the three relative sizes used in the experiment. Conducting some post-hoc tests of 2x2x2x2 repeated measures ANOVAs where each time two levels of size IV were picked revealed that the main effect of size was significant in all cases which indicates the strength of size IV in capturing the attention of the viewers. Mean size ratio judgment at relative size of 18 vs. 36 ($M = 3.519$) was significantly larger than mean size ratio judgment at relative size 18 vs. 30 ($M = 2.836$), $F(1, 73) = 191.463$, $p < 0.001$, $\eta^2 = 0.724$, and also significantly larger than mean size ratio judgment at relative size 18 vs. 24($M = 2.045$), $F(1, 73) = 273.888$, $p < 0.001$, $\eta^2 = 0.790$. Furthermore, mean size ratio judgment at relative size of 18 vs. 30 ($M = 2.836$) was significantly larger than the mean size ratio judgment at relative size 18 vs. 24 ($M = 2.045$), $F(1, 73) = 185.52$, $p < 0.001$, $\eta^2 = 0.792$. This finding confirms the findings of Bateman, *et al* that size is the strongest feature that captures the attention of the viewers of a tag cloud.

7.2.3 The Interaction Effect (Horizontal x Size)

The interaction between horizontal and size was significant, $F(2, 146) = 4.783$, $p = 0.010$, $\eta^2 = 0.061$ (Figure 52). Figure 52 indicates that the mean size ratio judgment of the target pair is higher in the right portion of the tag cloud for relative size 18 vs. 24, while the mean size ratio judgment of the target pair is higher in the left portion of the tag cloud for relative sizes 18 vs. 30 and 18 vs. 36. Post hoc tests (2x2x2 repeated measures ANOVAs where each time 2 levels of the size IV were

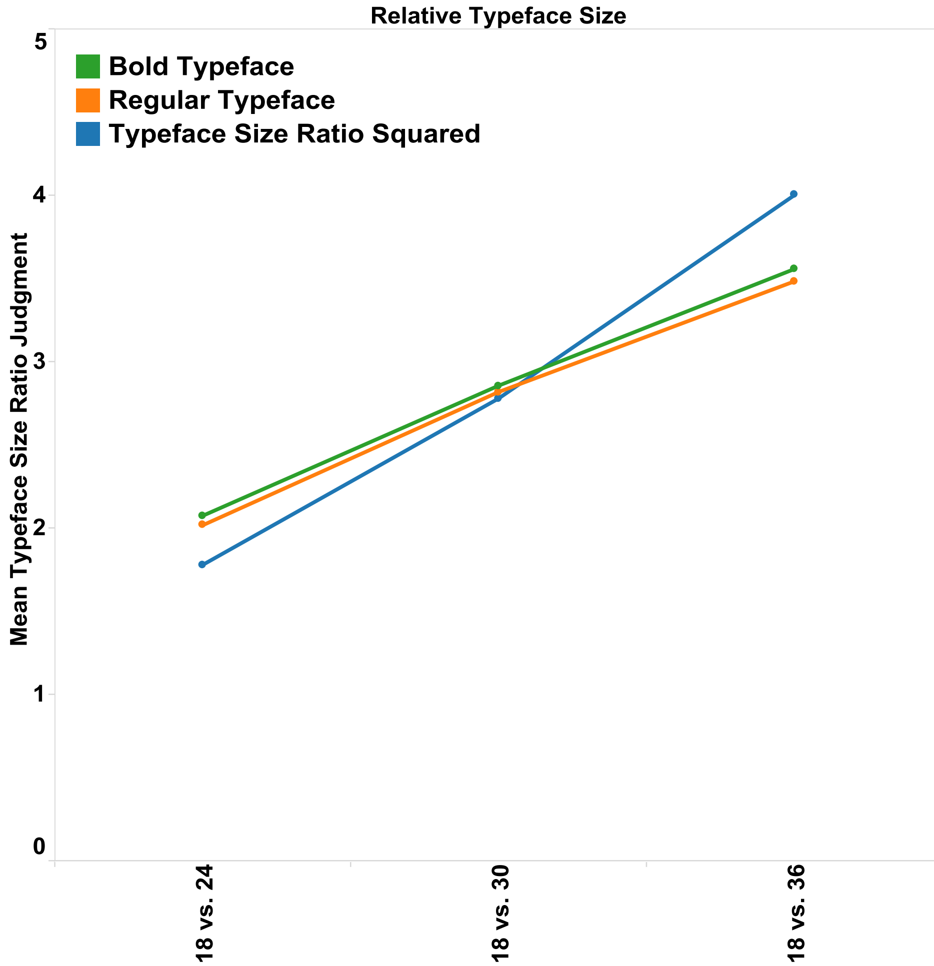


Figure 51: The interaction between the typeface style and relative size. The green and orange lines represent the mean size ratio judgments for bold and regular typefaces, respectively. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36.

picked) revealed that the main effect of horizontal IV was statistically significant at relative size 18 vs. 36, $F(1, 73) = 6.806$, $p = 0.011$, $\eta^2 = 0.085$. That indicates that mean size ratio judgment of the target pair in the left portion of the tag cloud ($M = 3.555$) is larger than the mean size ratio judgment of the target pair in the right portion of the tag cloud ($M = 3.484$) for the relative size 18 vs. 36. Table 7 shows the mean size ratio judgment at the horizontal and size interaction.

Table 7: The interaction between horizontal and size

Horizontal	Relative Size	Mean	Typeface Ratio Squared
Left	18 vs. 24	2.022	1.778
	18 vs. 30	2.856	2.778
	18 vs. 36	3.555	4
Right	18 vs. 24	2.067	1.778
	18 vs. 30	2.817	2.778
	18 vs. 36	3.384	4

7.2.4 The Interaction Effect (Vertical x Size)

There was also a significant interaction between vertical and size, $F(2,146) = 4.755$, $p = 0.010$, $\eta^2 = 0.061$ (Figure 53). In order to understand the effect, the levels of vertical (upper and lower) were compared with each other at each size, thus conducting separate repeated measures ANOVAs 2x2x2 (typeface style x horizontal x vertical) at each level of size independent variable. Mean size ratio judgment for the lower part ($M = 3.555$) was significantly higher than the mean size ratio judgment for the upper part ($M = 3.484$) at relative size 18 vs. 36, $F(1,73) = 6.103$, $p = 0.016$, $\eta^2 = 0.077$.

7.2.5 The Interaction Effect (Typeface Style x Horizontal x Size)

The interaction between typeface style, horizontal and size was significant, $F(2,146) = 4.818$, $p = 0.009$, $\eta^2 = 0.062$. In order to understand this interaction, it was broken into two interactions at each level of typeface style IV. For each of these two interactions, a 2x2x3 (horizontal x vertical x size) repeated measures ANOVA was conducted at each level of the typeface style. For the regular typeface style, the interaction between horizontal and size was statistically significant, $F(2,146) = 4.2$, $p = 0.017$, $\eta^2 = 0.054$ (Figure 54 A). For the bold typeface, the interaction between horizontal and size was statistically significant, $F(2,146) = 5.411$, $p = 0.005$, $\eta^2 =$

0.069 (Figure 54 B). The interactions in Figure 54 indicate that for the regular typeface, the mean typeface size ratio judgment for target pairs located in the left and right portions in the tag cloud tend to be almost the same at relative size 18 vs. 24 and 18 vs. 30 while there is a notable divergence at relative size 18 vs. 36. However, for the bold typeface, the horizontal and size interaction is mostly the opposite: while there is a notable divergence between mean size ratio judgment for target pairs located in the left and right portions of the tag cloud at relative sizes 18 vs. 24 and 18 vs. 30, the mean size ratio judgment for pairs located in left and right portions of the tag cloud is almost identical at relative size 18 vs. 36.

7.2.6 The Interaction Effect (Typeface Style x Horizontal x Vertical x Size)

The interaction between typeface style, horizontal, vertical and size was significant, $F(2, 146) = 3.38$, $p = 0.037$, $\eta^2 = 0.044$. In order to analyze the 4-way interaction, two 3-way interactions of three independent variables at each level of the forth independent variable were conducted. Type, Horizontal and Size IVs were analyzed at each level of the Vertical IV. Then each of the 3-way interactions was analyzed as two 2-way interactions at each level of the third IV. So, for the first time, the vertical IV was picked and the 4-way interaction was divided into two 3-way interactions at each vertical IV. Then, for each of the 3-way interactions, typeface style was picked and the horizontal and size interaction were analyzed at each typeface style IV. That resulted in four 2-way interactions. The interactions are in Figures 55, 56, 57, 58.

For the purpose of clarity, they are called: upper-regular (Figure 55), upper-bold (Figure 56), lower-regular (Figure 57), and lower-bold (Figure 58). For the upper-regular, the interaction between horizontal and size was not significant, while for the upper-bold, the interaction between the horizontal and size was significant, $F(2,146) = 6.296$, $p = 0.002$, $\eta^2 = 0.079$ (Figure 56). For the upper-bold part (Figure 56), separate paired samples t-tests between size ratio judgment of target pairs in the left and right portions in the tag cloud were conducted at each relative size: A paired samples t-test revealed a statistically significant difference between the mean size ratio judgment of target pair in the right part ($M = 2.141$) and left part ($M = 1.988$), $t(73) = -3.497$, $p = 0.001$ for bold typeface style in the upper portion of the tag cloud at relative size 18 vs. 24.

For the lower-regular part, the interaction between horizontal and size was significant, $F(2,146) = 3.386$, $p = 0.037$, $\eta^2 = 0.044$ (Figure 57). However, the interaction between horizontal and size was not significant in the lower-bold part (Figure 58). Looking at the mean size ratio judgment of target pairs in the lower-regular part (Figure 57), there is a notable divergence when participants judge target pairs at both right and left parts of the tag cloud at relative size 18 vs. 36. A paired samples t-test revealed a statistically significant difference between the mean size ratio judgment of target pair in the right part ($M = 3.414$) and left part ($M = 3.575$) at the relative size 18 vs. 36, $t(73) = 2.712$, $p = 0.008$ for target pairs that appear in the lower part of the tag cloud and in regular typeface.

7.2.7 One Sample T-Tests

One sample t-tests are used to compare the mean of a sample to a particular value. In the present experiment, three t-tests to compare the size ratio judgment at each relative size for all the participants with the value of typeface ratio squared at that particular size were conducted. For the relative typeface 18 vs. 24, on average, size ratio judgment of participants ($M = 2.0445$) was significantly larger than the typeface ratio squared at that particular size (1.778), $t(73) = 3.912$, $p < 0.001$. In addition, for the relative typeface 18 vs. 36, on average, the size ratio judgment of participants ($M = 3.5191$) was significantly lower than the typeface ratio squared (4). At relative size 18 vs. 30, the one sample t-test revealed that the mean typeface size ratio judgment was not statistically significant than the typeface ratio squared at that particular size, (2).

7.2.8 Paired Sample T-Tests

A paired sample t-test is used to determine whether the difference between the mean values of two conditions measured in the same way is significant. In the present experiment, separate paired sample t-tests between the mean values of typeface size ratio judgment of each two quadrants of the tag cloud (upper-left, upper-right, lower-left and lower-right) were conducted. Paired sample t-tests indicated that the mean typeface size ratio judgment of target pairs in the lower-left portion of the tag cloud ($M = 2.828$) was significantly different than the mean typeface size ratio judgment of target pairs in the upper-right portion of the tag cloud ($M = 2.781$), $t(73) = 2.163$, $p = 0.034$.

Separate paired sample t-tests were also conducted in order to understand if there were significant differences between the mean values of typeface size ratio judgment between each two quadrants at each relative typeface size (18 vs. 24, 18 vs. 30 and

18 vs. 36). These tests revealed that at relative size 18 vs. 24, the mean value of typeface size ratio judgment of target pairs in the upper-left portion in the tag cloud ($M = 2.001$) was significantly different than the mean value of typeface size ratio judgment of target pairs in the lower-right portion of the tag cloud ($M = 2.073$), $t(73) = 2.701$, $p = 0.009$. Furthermore, at relative size 18 vs. 36, and the mean value of typeface size ratio judgment of target pairs in the upper-left portion of the tag cloud ($M = 3.517$) and the mean value of typeface size ratio judgment of target pairs in the lower-left portion in the tag cloud ($M = 3.592$) were significantly different, $t(73) = 2.241$, $p = 0.028$. In addition, at relative size 18 vs. 36, the means values of typeface size ratio judgment of target pairs at both lower-left ($M = 3.592$) and upper-right ($M = 3.45$) portions of the tag cloud were significantly different, $t(73) = 3.733$, $p < 0.001$. Figure 59 shows the significant differences between mean values of typeface size ratio judgment in different quadrants.

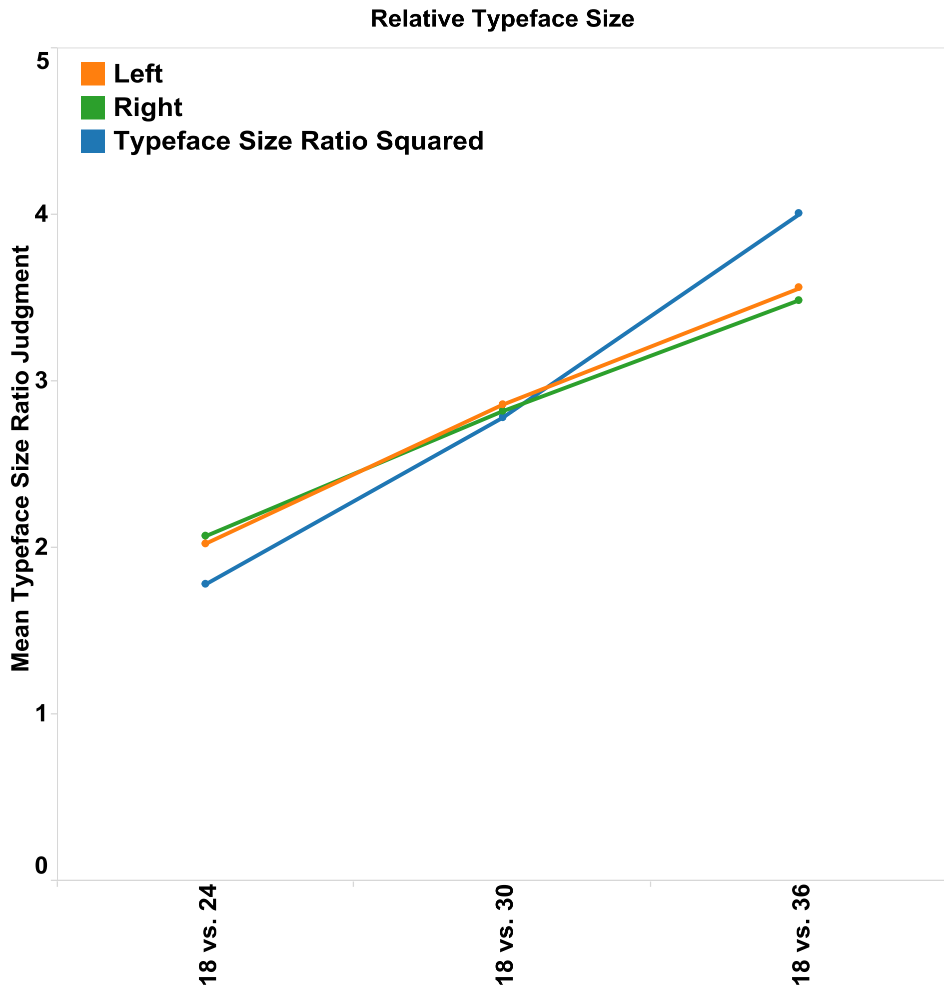


Figure 52: The interaction between horizontal and relative size IVs. The levels of the horizontal IV (left and right) are represented by orange and green lines, respectively. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36

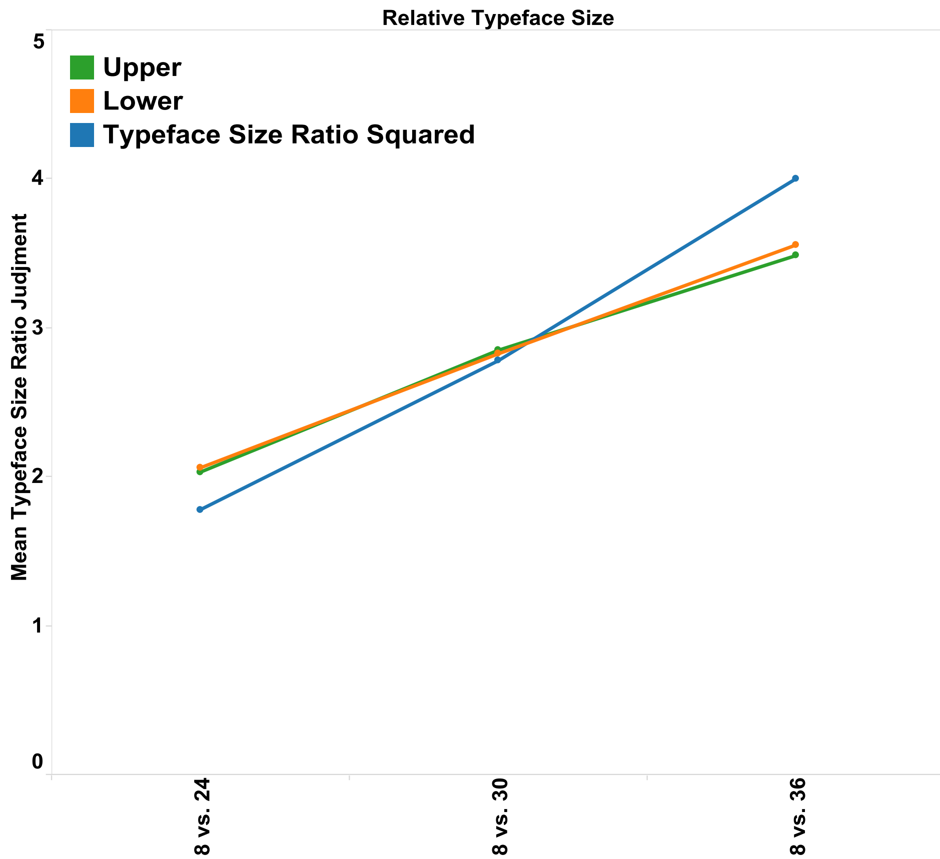
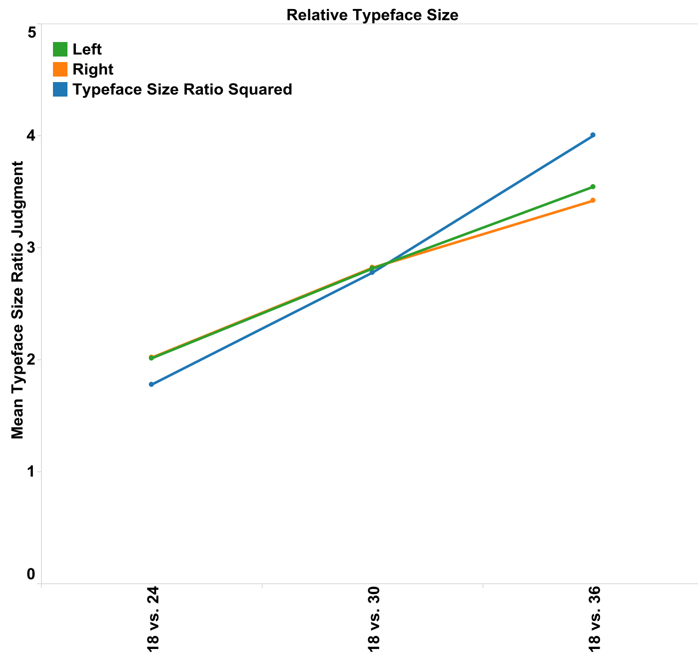
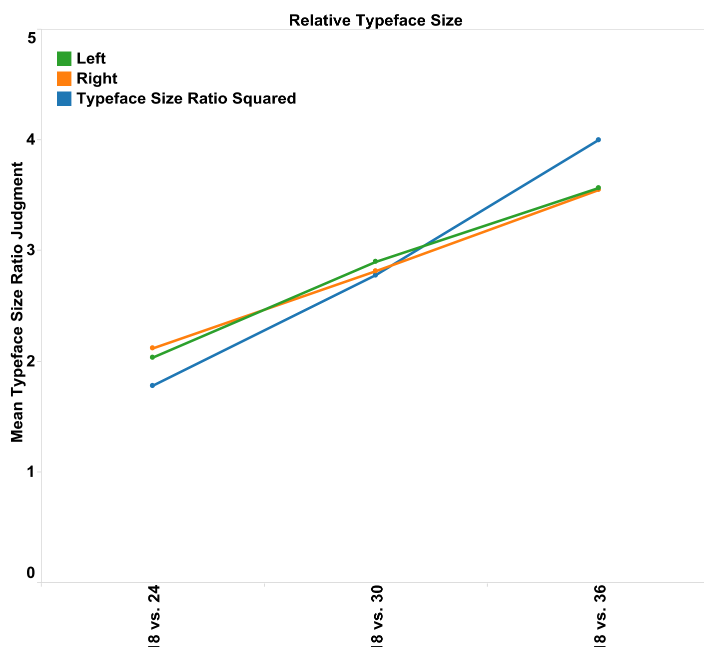


Figure 53: The interaction between vertical and relative size. The two levels of vertical IV are represented by two lines: green (upper part) and orange (lower part). The blue line represents the typeface ratio squared at each relative size. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36



(A) Regular typeface style



(B) Bold typeface style

Figure 54: The interaction between horizontal and size IVs for regular and bold typeface styles. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36. The mean size ratio judgments of the relative typeface sizes are in green and orange colors for left and right portions of the tag cloud, respectively

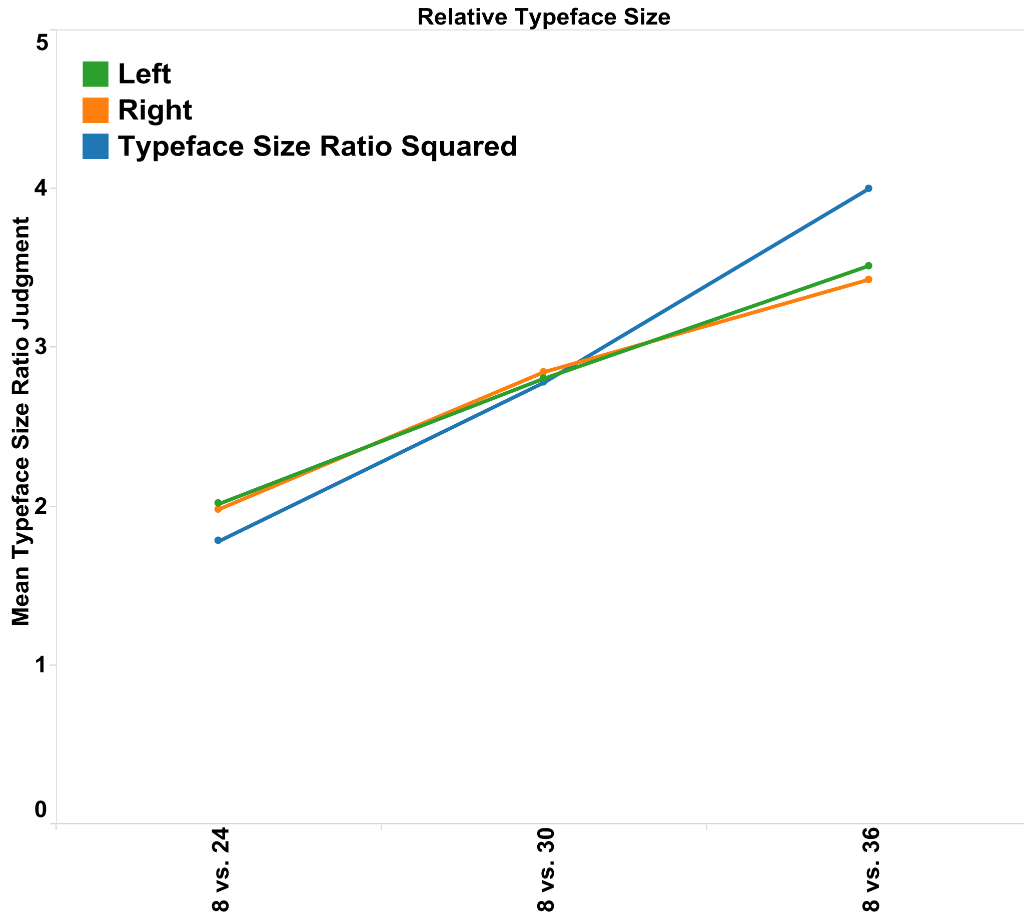


Figure 55: The interaction between horizontal and size IVs in the upper part of the tag cloud for regular typeface. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36. The mean size ratio judgments of the relative typeface sizes are in both green and orange colors for left and right portions of the tag cloud, respectively

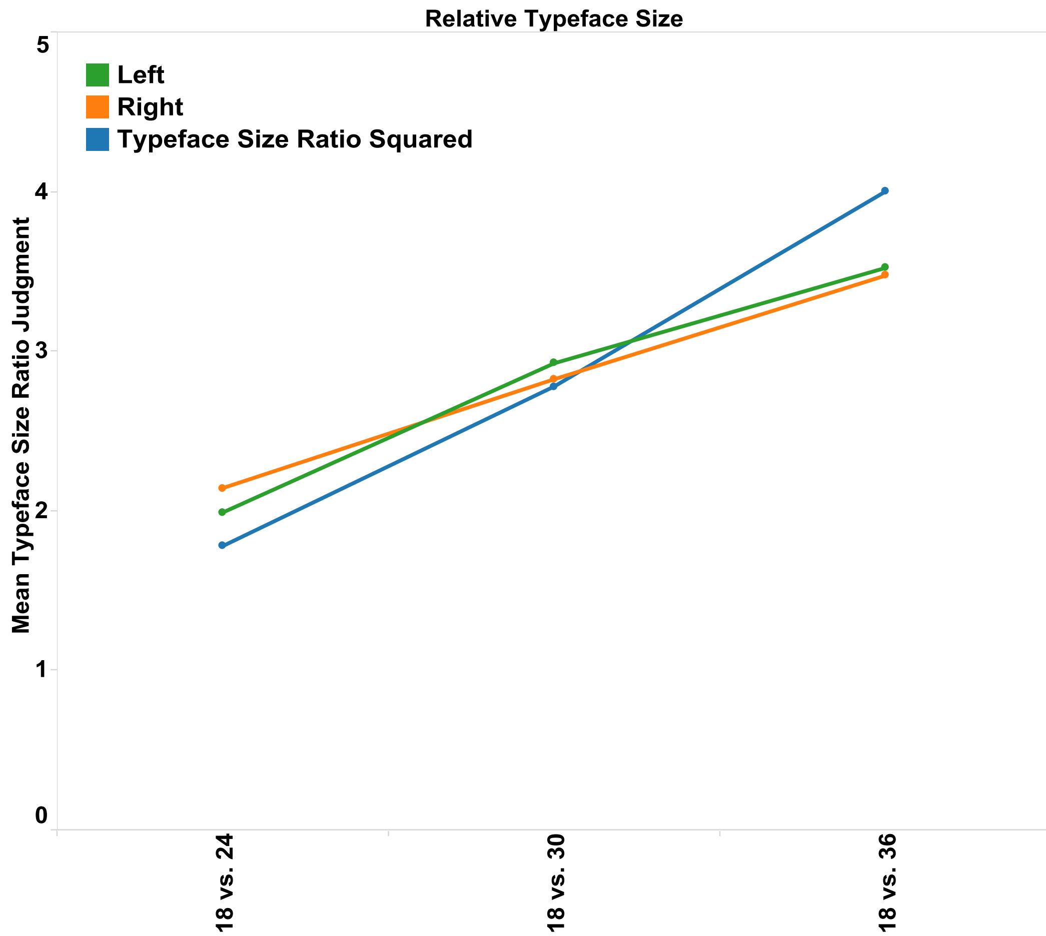


Figure 56: The interaction between horizontal and size IVs in the upper part of the tag cloud for bold typeface. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36. The mean size ratio judgments of the relative typeface sizes are in both green and orange colors for the left and right portions of the tag cloud, respectively

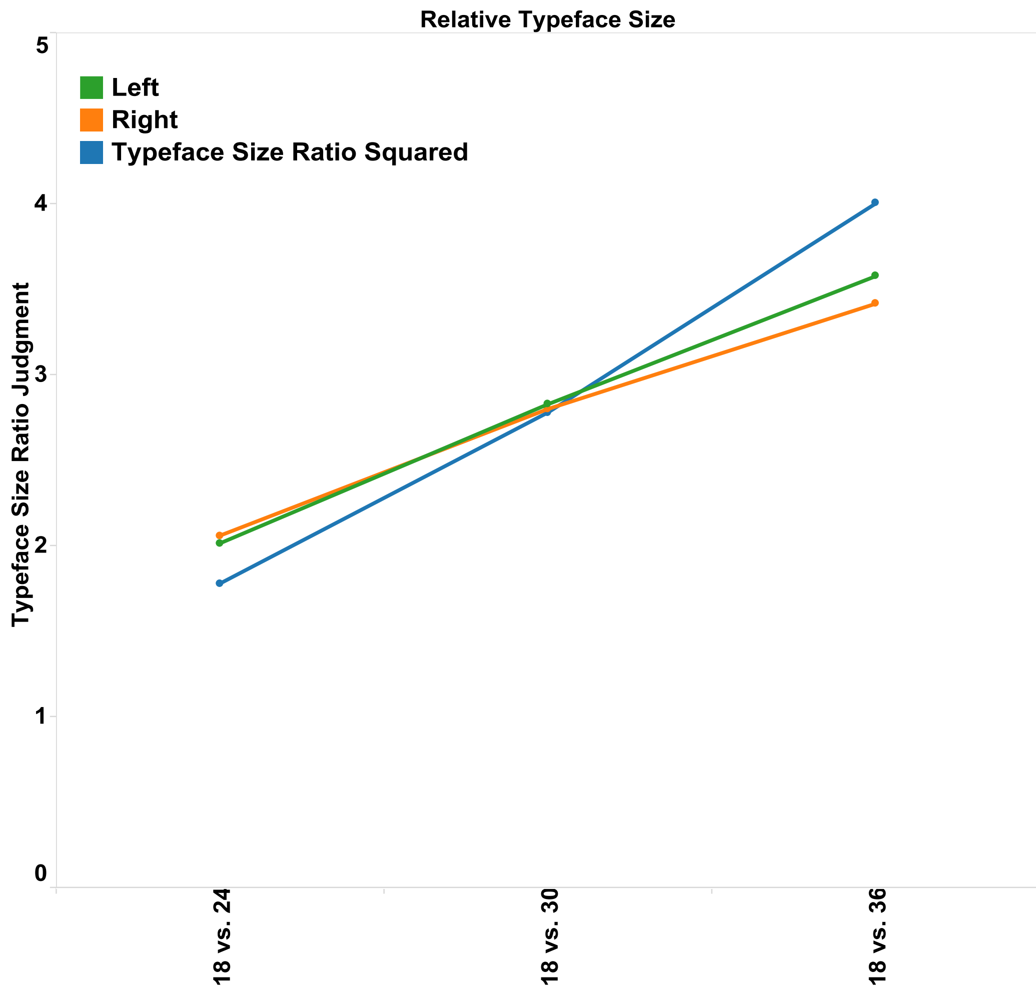


Figure 57: The interaction between horizontal and size IVs in the lower part of the tag cloud for regular typeface. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36. The mean size ratio judgments of the relative typeface sizes are in both green and orange colors for the left and right portions of the tag cloud, respectively

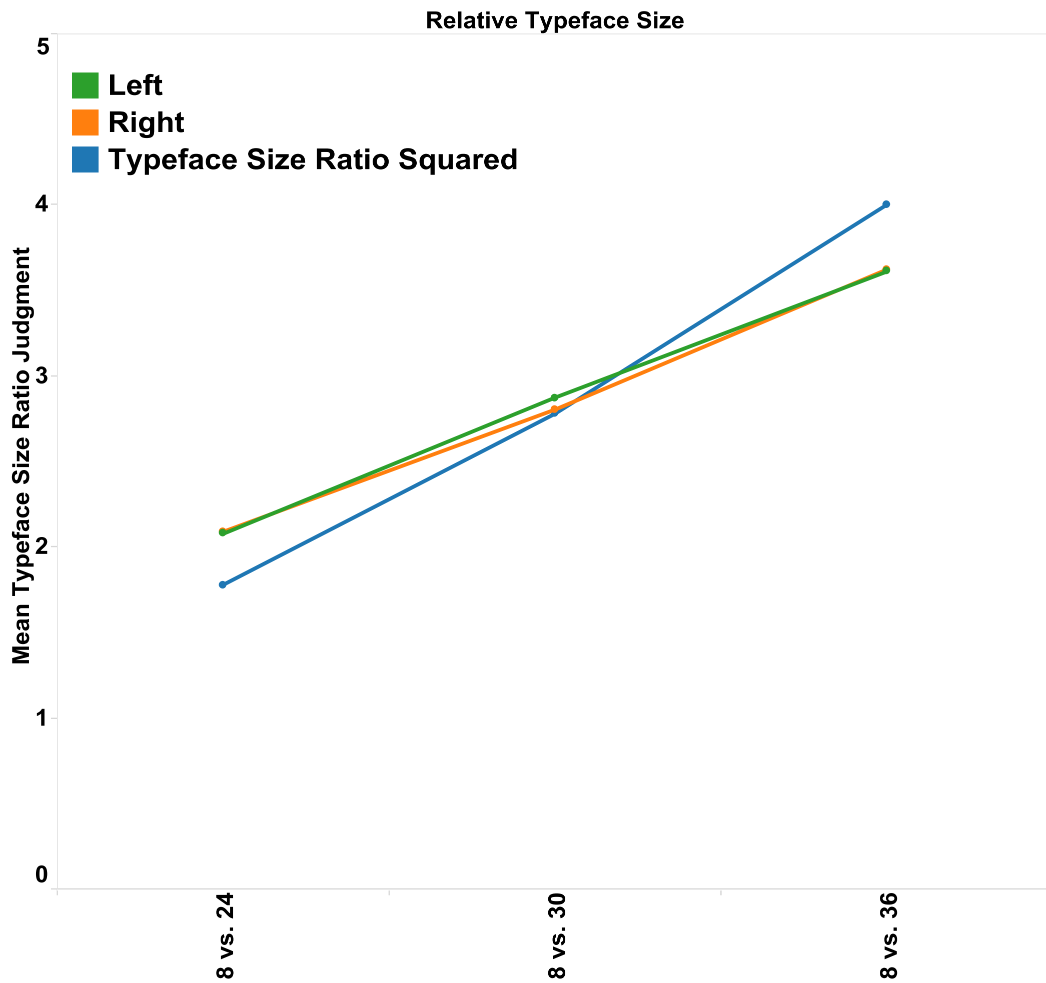


Figure 58: The interaction between horizontal and size IVs in the lower part of the tag cloud for bold typeface. The blue line is the typeface ratio squared: $(24/18)^2 = 1.778$ for relative size 18 vs. 24, $(30/18)^2 = 2.778$ for relative size 18 vs. 30 and $(36/18)^2 = 4$ for relative size 18 vs. 36. The mean size ratio judgments of the relative typeface sizes are in both green and orange colors for the left and right portions of the tag cloud, respectively

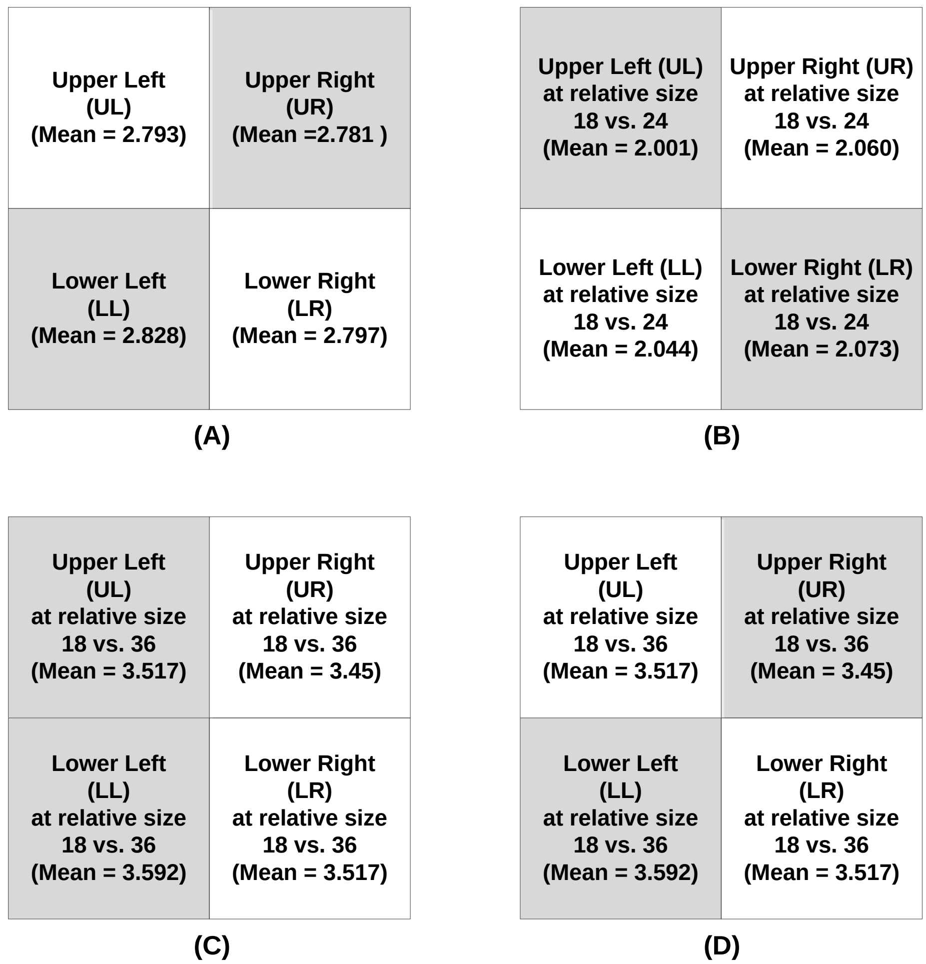


Figure 59: Significant difference between mean values of typeface size ratio judgment in different quadrants. The two highlighted quadrants in (A), (B), (C) and (D) indicate that the means of typeface size ratio judgment of target pairs in these quadrants are statistically different

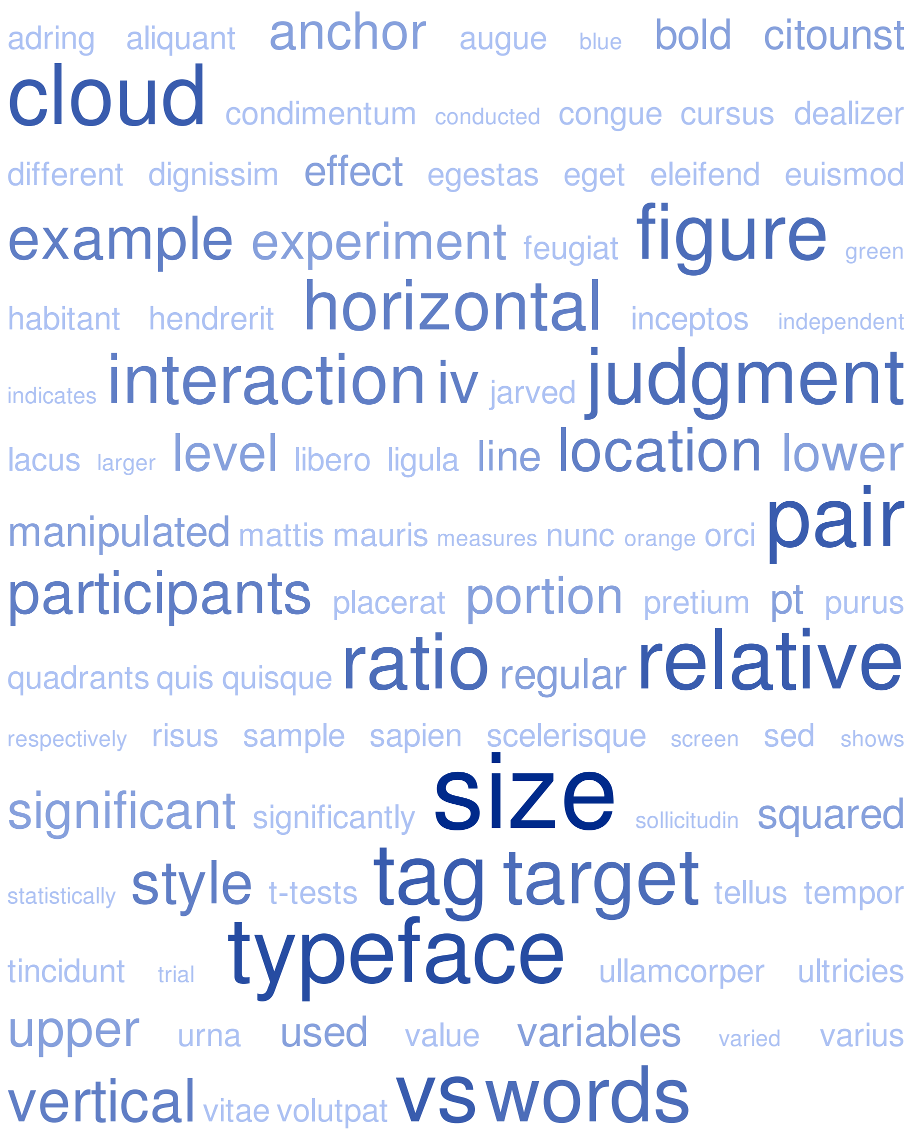


Figure 60: Tag cloud of chapter 7 generated by `www.tagcrowd.com`

CHAPTER 8: GENERAL DISCUSSION

Freedom is hammered out on the anvil of

discussion, dissent, and debate

— Hubert H. Humphrey

The empirical results of experiment 1 showed that there was a small main effect of display type where the target pair was judged a bit larger when the tag cloud was present, thus ameliorating the general under-judgment of the size ratio to a small extent. This confirms H2. There was only a single small interaction effect that included display type, so the pattern of size ratio judgment across the cells of the experimental design was qualitatively similar with tag clouds present and absent display types. The overall under-judgment in experiment 1 was strongly modulated by target word size (Figures 31 and 32). As the target word size was increased the under-judgment of the size ratio became larger. In addition, it was noticeable that the mean size ratio judgment was nearly veridical for the 18 pt. target words. The substantial under-judgment effect may have been due, at least in part to an anchoring effect [47] due to using a relatively small anchor word (12 pt.) in this experiment. To adjust for this possibility, a larger anchor word size for experiment 2 was chosen.

In addition, in experiment 1, most participants started out having nearly correct judgment and then they began under-judging the ratio size and that under-judgment increased as the target word got bigger in size (Figure 31 and 32). The fan-out from left to right on the graph, where the data of overall size ratio judgment and actual size ratio lines diverge indicates increasing perceptual bias as the size of the target word is increased. The linearity of the data in comparison to the upward curve of the typeface squared ratio (in blue) curve indicates that participants are relying on a single dimension, probably width, over the other dimension (height) or the combined product (2D size).

Based on the empirical results in experiment 1, size proved to be a powerful factor on the typeface size ratio judgment of words. This is an indication of the strength of the size factor in capturing the attention of the viewers. This suggests that designers should pay attention to the size factor and how it interacts with other design elements such as the number of pixels, location, color and saturation of words in a tag cloud.

The findings in experiment 2 also confirm the findings of experiment 1 in that the effect of the size of the typeface is robust. This finding is confirmed by the finding of Halvey and Keane [42] who emphasized the importance of typeface size and Bateman, *et al* [15] who found a strong visual influence of the typeface size in tag clouds and determined that users could identify even small variations in the font size. In addition, experiment 2 further explored the typeface size ratio judgment and how that could be influenced by varying the relative size. In experiment 1, the overall mean size ratio judgment was under-judged for large target words, while the judgment was nearly the same as the typeface size ratio when a small target word was used. What is new in experiment 2 is that the mean size ratio was over-judged for the smaller target words, with a point where the mean size ratio was judged near to the correct value (i.e. for the medium target word size). From an applied perspective, this gives tag cloud designers information about the relative typeface sizes to use in their tag clouds. From a theoretical perspective, this is an interesting finding that might point towards a better understanding of how viewers perceive the size of irregular complex figures with a clear dominant horizontal axis. This finding refuted H1.

Experiment 1 did, however, reveal some findings on the appearance of words in tag clouds. Viewers seem to be slightly influenced with ascending and descending letters when the distractors are absent while these effects are dissipated in a tag cloud environment. This finding confirms H3 and suggests that tag cloud designers should not put a lot of emphasis on the existence of ascending and descending letters in tag clouds as they don't seem to bias the perception of the viewer. In contrast, this study clearly showed that the size ratio judgment was greatly influenced by varying the width between the two words in the target pair, which is an indication that the width factor needs to receive more attention in the perception of tag clouds research. This finding is a contrary to H4.

Although mean typeface size ratio judgments between different decorations seem to be close to each other (See Figure 33), empirical results showed a minor influence of boxes and filled areas in the typeface ratio size judgment in a tag cloud present display type when a small target word was used. This correlates with what has actually been applied in journalism where decorations around the words were used and in small typeface size [5] (Figure 61). Based on the empirical results in this dissertation, decorations should only be used for smaller target words. In addition, as showed by statistical tests, the effect of the decoration on typeface size ratio judgment was stronger in the tag cloud absent display type. A further study on different sizes and decorations will likely help understand this effect better.

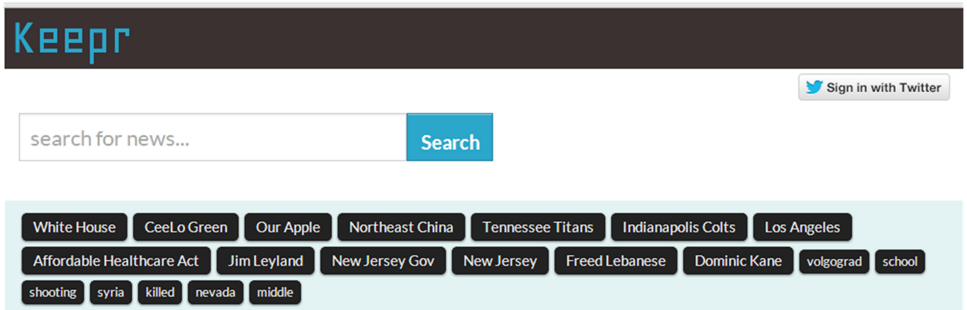


Figure 61: Example of a tag cloud used by a journalist adapted from [5]

The typeface style also had a significant effect and participants tended to over-judge the typeface size ratio judgment of bold typeface. This clearly shows the influence of bold typeface and this is also supported by Bateman, *et al* [15] who found that the font weight in a tag cloud had a strong visual effect. This confirms H5. Bold text also seemed to be a noticeable factor as some comments from participants about that were reported. There were 39 comments of users indicating that the typeface style (i.e. bold vs. regular) influenced their perception of relative size such as: "I based my decisions on the boldness of each of the font of each of the words and tried to visualize blowing them up from the small word to match the large", "The bold face letters led me to believe the word was bigger than it actually was" and "the boldness of the word creates the image of the word being bigger than it really is". Based on the empirical results, it was also clear that for regular font, the left-right difference emerges at larger sizes for the target word, whereas, for bold font the left-right differences emerge for the medium and smaller target word sizes (See Figure 54). These effects are small and may indicate different spatial attention biases for regular and bold font. Overall, bold and regular fonts were not appreciably differentially biased to a degree that tag cloud creators should worry about this variable. Spatial attention may have small effects where relative word size ratio judgment is more distorted on the right than the left for both bold and regular typefaces, but in different size ranges.

The main effects of both horizontal and vertical independent variables were not significant. However, an examination of the quadrants reveals, the means of typeface size ratio judgment of target pairs in some quadrants were statistically different from each other. In addition, the empirical results showed that the mean values of typeface size ratio judgment of target pairs in each two quadrants at the relative size 18 vs. 30 were not statistically different from each other. The minor influence of the location on typeface size ratio judgment was only notable between some quadrants at relative sizes 18 vs. 24 and 18 vs. 36. Empirical results also showed that the typeface size ratio judgment of pairs in the upper-left portion was not significantly higher than that in the other quadrants in tag cloud. Interestingly, participants were slightly over-judging target pairs in the lower quadrants as opposed to target pairs in the upper quadrants. This is a contrary to H6.

Figure 62: Tag cloud of chapter 8 generated by `www.tagcrowd.com`

CHAPTER 9: SUMMARIES AND CONCLUSIONS

> *Most action is based on redemption and revenge, and that's a formula. Moby Dick was formula. It's how you get to the conclusion that makes it interesting.*
>
> — Sylvester Stallone

This dissertation studied the effect of different tag cloud visualization parameters on perceiving the relative size of words consisting of dissimilar characteristics in tag clouds. In this research, two experiments were conducted and both of them measured typeface size ratio judgment of a target pair of words. In experiment 1, the display type (tag cloud absent vs. tag cloud present), appearance of words (e.g. words with and with no ascending letters and wide versus narrow words), the decorations around words (e.g. boxes, shadows) and relative typeface size were manipulated. In experiment 2, the typeface style (regular vs. bold), the horizontal location (left vs. right), the vertical location (upper vs. lower) and the relative size were manipulated. The size of the anchor word was also varied, using a smaller anchor word in Experiment 1 and a larger anchor word in Experiment 2.

The main goals of this dissertation were to understand how relative typeface size was perceived in a tag cloud and how size interacted with other design elements. Size is one the main characteristics varied by tag cloud designers to indicate the relative frequency of words in a search text. Because size has a strong influence in capturing the attention of the viewers, it was chosen to be studied in this research. An extensive literature review did not reveal any study that observed how users quantitatively judge relative size of two words in a tag cloud and it was studied for the first time in this dissertation.

The goals of this dissertation were achieved. First, the outcomes demonstrated that there was a generally increasing under-judgment of the relative size of larger words in a tag cloud that escalates with increasing size. However, if the larger and smaller words differ by less than a factor of 3 in size ratio squared, relative size judgments were approximately veridical. In addition, there was an increasing over-judgment of the relative size that increased with decreasing the size of the target word. Moreover, the fact that the relative size judgments scaled up linearly with increasing target word typeface, rather than increasing upward in a nonlinear trend (i.e., in a concave quadratic pattern) suggests that viewers were weighting a single dimension, perhaps the more salient dimension of word width, more heavily than the area (i.e., product of height and width). This hypothesis is supported by the observation, in Experiment 1 that displays where the larger word had a greater proportion of wide letters and resulted in an increase in size ratio judgments in the absence of a typeface size change.

Beyond experiment 1, The second goal of this dissertation was understanding how size interacts with other design characteristics. Empirical results clearly showed the influence of the appearance of words. Overall, ascenders and descenders did not seem to have an influence as opposed to words with wider letters that did have an influence. In addition, results clearly showed the minor influence of boxes and filled areas on typeface size ratio judgment of pair of words in tag cloud displays. Furthermore, the results showed the influence of the typeface style (e.g. bold typeface) and how it also interacted with the location of the target pairs.

The significance of this research is that it answers many questions about typeface size ratio judgment in tag clouds and how different characteristics of tag clouds can interact with each other to influence the typeface size ratio judgment. It provides many findings for tag cloud designers to take into consideration when designing tag clouds, given the importance of the relative size in communicating the frequency of the words. The empirical results of this dissertation pave the road for new experiments that look at typeface size ratio judgments in tag clouds. The overall design used in this dissertation can be replicated to study the influence of other features such as color, saturation, number of pixels and number of characters on relative typeface size judgment.

Figure 63: Tag cloud of chapter 9 generated by `www.tagcrowd.com`

REFERENCES

[1] Amazon mechanical turk. http://mturk.com/, November 2011.

[2] Lorem ipsum. http://www.lipsum.com/, May 2011.

[3] Non-sense words. http://www.soybomb.com/tricks/words/, May 2011.

[4] Nonsense words for use in training phonemic awareness. http://www.speech-language-development.com/nonsense-words.html, May 2011.

[5] Keepr. Online, October 2013. Tag cloud retrieved from www.keepr.com on Oct 21, 2013.

[6] Tagcrowd. http://www.tagcrowd.com, cited Aug 2013.

[7] Wordle. http://www.wordle.net/, retrieved Dec 2013.

[8] Amazon. http://www.amazon.com/gp/tagging/cloud, retrieved Oct 2013.

[9] Chrome. www.google.com/chrome, retrieved Oct 2013.

[10] Flikr. http://www.flickr.com/photos/tags/, retrieved Oct 2013.

[11] J. Alsakran. *Supporting Interactive Visual Analytics for High Dimensional Data Exploration*. PhD thesis, Kent State University, 2012.

[12] C. M. Angel. *A Comparison of Descriptive Tagging Practices by Library, Archive and Museum Professionals using an Inter-Indexing Consistency Approach*. PhD thesis, University of South Carolina, 2012.

[13] N. Z. Aydinoglu, A. Krishna, and B. Wansink. *Sensory marketing : research on the sensuality of products*, chapter Do size labels have a common meaning among consumers?, pages 343–355. 2010.

[14] S. Bateman, October 2013. Personal correspondence by email.

[15] S. Bateman, C. Gutwin, and M. Nacenta. Seeing things in the clouds: the effect of visual features on tag cloud selections. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, HT '08, pages 193–202, New York, NY, USA, 2008. ACM.

[16] K. Börner. Visualization framework & workflow design. available online from http://ivmooc.cns.iu.edu retrieved on August 2, 2013, January 2013. Lecture titled "Visualization Framework & Workflow Design" from Information Visualization MOOC.

[17] R. Bringhurst. *The Elements of Typographic Style*. Hartley & Marks Publishers, 2nd edition, 2002.

[18] J. Bumgardner. *Building Tag Clouds in Perl and PHP.* O'Reilly, 2006.

[19] C. M. Carmean. *e-Learning Design 2.0: Emergence, Connected Networks and the Creation of Shared Knowledge.* PhD thesis, Capella University, 2008.

[20] C. Chase. Using design patterns in user interface design. Master's thesis, University of Cincinnati, 2012.

[21] W. S. Cleveland, C. S. Harris, and R. McGill. Judgments of circle sizes on statistical maps. *Journal of the American Statistical Association*, 77(379):541–547, September 1982.

[22] W. S. Cleveland, C. S. Harris, and R. McGill. Experiments on quantitative judgments of graphs and maps. *Bell System Technical Journal*, 62(6):1659–1674, August 1983.

[23] C. Collins, F. Viegas, and M. Wattenberg. Parallel tag clouds to explore and analyze faceted text corpora. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*, pages 91–98, 2009.

[24] C. M. Collins. *Interactive Visualizations of Natural Language.* PhD thesis, University of Toronto, 2010.

[25] A. W. R. Cortez. *Tag clouds: How format and categorical structure affect categorization judgment.* PhD thesis, University of Maryland, College Park, 2008.

[26] C. L. Crawford. *Using a Collaborative Center for Integrative Reviews and Evidence Summaries to Narrow the Education-Practice-Research Gap.* PhD thesis, College of Graduate Nursing, 2012.

[27] G. Critel. *Investigating The Rhetoric of Student Participation: Uncovering and Historicizing Commonplaces in Composition Studies.* PhD thesis, Ohio State University, 2012.

[28] R. Das. *Software System Selection based on First-Order Security Risk Assessment.* PhD thesis, George Washington University, 2013.

[29] J. de Fockert, G. Rees, C. Frith, and N. Lavie. Neural correlates of attentional capture in visual search. *Journal of Cognitive Neuroscience*, 16(5):751–759, 2004.

[30] S. Desai. System for interpretation of reviewer feedback on social web. Master's thesis, University of Texas at Arlington, 2012.

[31] A. K. Dhone. A video library management software toolkit for the nevada climate change portal. Master's thesis, University of Nevada, Reno, 2012.

[32] J. N. Druckman. Evaluating framing effects. *Journal of Economic Psychology*, 22(1):91 – 101, 2001.

[33] E. du Plessis. *The Advertised Mind: Groundbreaking Insights into How Our Brains Respond to Advertising*. Kogan Page, 2005.

[34] L. S. Elting, C. G. Martin, S. B. Cantor, and E. B. Rubenstein. Influence of data display formats on physician investigators' decisions to stop clinical trials: prospective trial with repeated measures. *BMJ*, 318(7197):1527–1531, 6 1999.

[35] J. Emanuel. *The Millennials: Assessing the next generation of academic librarians*. PhD thesis, University of Missouri-Columbia, 2012.

[36] J. Felici. *The Complete Manual of Typography: A Guide to Setting Perfect Type*. Adobe Press, 2nd edition, 2011.

[37] B. Furr. *Single and Dual-Directional Adult Caregiver Compliance to Medical Instructions*. PhD thesis, Walden University, 2013.

[38] T. Garg. Topic network: a semantic model for effective learning. Master's thesis, University of Missouri Kansas City, 2009.

[39] S. Golder and B. A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32:198–208, 2005.

[40] R. Gurney. Native american food security, traditional foods, and obesity. Master's thesis, Oklahoma State University, 2012.

[41] K. L. Guzak. Engineering design motivations: A qualitative and quantitative analysis of university students. Master's thesis, Michigan Technological University, 2012.

[42] M. J. Halvey and M. T. Keane. An assessment of tag presentation techniques. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 1313–1314, New York, NY, USA, 2007. ACM.

[43] G. W. Hayman. *Turning Prisons into Learning Communities: A New Vision for Corrections Education*. PhD thesis, Wilmington University, 2012.

[44] M. A. Hearst and D. Rosner. Tag clouds: Data analysis tool or social signaller? In *Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, HICSS '08, pages 160–, Washington, DC, USA, 2008. IEEE Computer Society.

[45] K. Hoyt. *Changing the climate while reproducing power?: Examining the social construction of "renewable" frames in mass print news media*. PhD thesis, University of Colorado, 2012.

[46] B. Hultn, N. Broweus, and M. van Dijk. *Sensory Marketing*. Palgrave Macmillan, 2009.

[47] K. Jacowitz and D. Kahneman. Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21:1161–1166, 1995.

[48] Y. A. Kang. *Informing design of visual analytics systems for intelligence analysis: understanding users, user tasks, and tool usage*. PhD thesis, Georgia Institute of Technology, 2012.

[49] G. D. Kantarakias. *Discourse, Evaluation, and Learning in Communications: A Content Analysis of Presidential Incidents with Leadership Theory*. PhD thesis, University of Phoenix, 2011.

[50] G. D. Kantarakias, October 2013. Personal correspondence by email.

[51] A. Kirk. The 8 hats of data visualization design, June 2012. Presentation on Meeting of The Chicago Data Visualization Group (CDVG) by Andy Kirk from visualisingdata.com.

[52] R. Klanten, M. Mischler, and S. Bilz. *The Little Know-It-All: Common Sense for Designers*. Gestalten Verlag, 2006.

[53] S. Klemmer. Visual design. available online from `https://www.coursera.org/` retrieved on October 3, 2013, October 2013. Lecture from Human Computer Interaction Course in Coursera.

[54] B. Y.-L. Kuo, T. Hentrich, B. M. . Good, and M. D. Wilkinson. Tag clouds for summarizing web search results. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 1203–1204, New York, NY, USA, 2007. ACM.

[55] D. A. Lauer. *Design Basics*. Cengage Learning, 2007.

[56] B. Lee, N. H. Riche, A. K. Karlson, and M. S. T. Carpendale. Sparkclouds: Visualizing trends in tag clouds. *IEEE Trans. Vis. Comput. Graph.*, 16(6):1182–1189, 2010.

[57] M. D. LeNoue. *Educational Social Software: The Use of Social Network Sites for Teaching and Learning*. PhD thesis, North Dakota State University, 2012.

[58] A. G. Levy. *Territory and Sovereignty in a Borderland: the Case of Moldova*. PhD thesis, University of Colorado, 2012.

[59] M. Lindstrom. *Brand Sense Build Power Brands through Touch, Taste, Smell, Sight and Sound*. Free Press, 2005.

[60] S. L. Liu. *U.S. Women Bullying Women in the Pharmaceutical/Biotechnology/Medical Device Industry*. PhD thesis, Pepperdine University, 2012.

[61] S. Magazine. *Getting the Hang of Web Typography*. Smashing eBook Series. Smashing Media GmbH, 2011.

[62] C. A. Marriott. *Assessment Methods and Tools for Architectural Curricula*. PhD thesis, Illinois Institute of Technology, 2012.

[63] J. Mates, V. Di Maio, and P. Lansky. A model of the perception of area. *Spatial Vision*, 6:101–116, 1992.

[64] M. Matsunaga. Electronic form design. Master's thesis, Iowa State University, 2012.

[65] K. McKnight. *The Teacher's Big Book of Graphic Organizers: 100 Reproducible Organizers that Help Kids with Reading, Writing, and the Content Areas.* Jossey-Bass teacher. Wiley, 2010.

[66] B. Medler. *Play with Data - An Exploration of Play Analytics and its Effect on Player Experiences.* PhD thesis, Georgia Institute of Technology, 2012.

[67] D. N. Mueller. *Clouds, Graphs, and Maps: Distant Reading and Disciplinary Imagination.* PhD thesis, Syracuse University, 2009.

[68] S. D. Nowell. *Stories of Digital Lives: Teacher-Student Relationships in Secondary Classrooms and Digital Spaces.* PhD thesis, Oklahoma State University, 2012.

[69] J. G. Olund. *Qualitative Research in Understanding Effects of Mobile Technology Use in Worker-Leader Relationships.* PhD thesis, University of Phoenix, 2012.

[70] J. M. Plumert and P. Nichols-Whitehead. Developmental differences in preferences for using color, size, and location information to disambiguate hiding places. *Journal of Cognition and Development*, 8(4):427–454, 2007.

[71] D. Poole. *A Qualitative Analysis of the Effectiveness of Professional Ethics Training Standards in Law Enforcement.* PhD thesis, Walden University, 2012.

[72] J. Provost. Improved document summarization and tag clouds via singular value decomposition. Master's thesis, Queens University, 2008.

[73] A. W. Rivadeneira, D. M. Gruen, M. J. Muller, and D. R. Millen. Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 995–998, New York, NY, USA, 2007. ACM.

[74] M. Sandow. *An appreciative inquiry review of effective performance management within the Federal government environment.* PhD thesis, University of Phoenix, 2012.

[75] T. Sanocki and M. C. Dyson. Letter processing and font information during reading: Beyond distinctiveness, where vision meets design. *Attention, Perception, & Psychophysics*, 74(1):132–145, 2011.

[76] M. M. Schapira, A. B. Nattinger, and C. A. McHorney. Frequency or probability? a qualitative study of risk communication formats used in health care. *Med Decis Making*, 21(6):459–67, 2001.

[77] M. Schonlau and E. Peters. Graph comprehension: An experiment in displaying data as bar charts, pie charts and tables with and without the gratuitous 3rd dimension. *RAND Working Paper*, September 2008.

[78] J. Schrammel, S. Deutsch, and M. Tscheligi. Visual search strategies of tag clouds - results from an eyetracking study. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part II*, INTERACT '09, pages 819–831, Berlin, Heidelberg, 2009. Springer-Verlag.

[79] P. Shankar. *Improving performance, privacy and relevance of location-based services for mobile users*. PhD thesis, Rutgers, The State University of New Jersey, 2011.

[80] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *J. Inf. Sci.*, 34:15–29, February 2008.

[81] I. Spence. Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Performance and Perception*, 16:683–692, 1990.

[82] J. Steele and N. Iliinsky. *Designing Data Visualizations: Representing Informational Relationships*. Oreilly and Associate Series. O'Reilly Media, 2011.

[83] S. S. Stevens. On the psychophysical law. *Psychology Revue*, 64(3):153–181, 1957.

[84] B. M. Stiles. *Tinkering at the Edges of Public EducationTeacher and Parent Perceptions of West Virginia Universal Pre-kindergarten*. PhD thesis, West Virginia University, 2012.

[85] A. A. Stirratt. Social media use in march 2011 japanese crisis: Impact on emergency preparedness advocacy. Master's thesis, Purdue University, 2011.

[86] I. Strizver. *Type Rules!: The Designer's Guide to Professional Typography*. Wiley, 2nd edition, 2006.

[87] K. E. Sulak. *What Makes a Font Persuasive?: An Eye-Tracking Study of Perception in American and Chinese Assessment of Fonts*. PhD thesis, Clemson University, 2012.

[88] K. Thompson. Scalable methods for analyzing human movement from location-based social media. Master's thesis, Binghamton University State University of New York, 2012.

[89] X. Wang. *Connecting Users with Similar Interests for Group Understanding*. PhD thesis, Arizona State University, 2013.

[90] H. C. White. *Organizing scientific data sets: Studying similarities and differences in metadata and subject term creation*. PhD thesis, University of North Carolina at Chapel Hill, 2012.

[91] K. S. White. *Uncovering Crowd Behavioral Variables to Understand Crowd Violence*. PhD thesis, Walden University, 2013.

[92] Y. Wu. Statistical analyses of the perceptions and trust of adaptive cruise control owners. Master's thesis, University of Washington, 2012.

[93] N. C.-Y. Yau. *An Online Tool for Personal Data Collection and Exploration*. PhD thesis, University of California Los Angeles, 2013.

[94] Q. Zhang, W. Qu, and L. Wang. How font size and tag location influence chinese perception of tag cloud? In *Proceedings of the 9th international conference on Engineering psychology and cognitive ergonomics*, EPCE'11, pages 273–282, Berlin, Heidelberg, 2011. Springer-Verlag.

[95] J. Zwislocki. *Sensory Neuroscience: Four Laws of Psychophysics*. SpringerLink: Springer e-Books. Springer London, Limited, 2009.

APPENDIX A: EXPERIMENT 1 INITIAL APPROVAL

UNC CHARLOTTE

Office of Research Services

9201 University City Boulevard, Charlotte, NC 28223-0001
t/ 704.687.2291 f/ 704.687.2292 www.research.uncc.edu

Institutional Review Board (IRB) for Research with Human Subjects

Approval of Exemption

Protocol #	**11-03-21**			
Title:	**Testing the Effect of Font Scale in Tag Clouds**			
Date:	**3/22/2011**			
Responsible Faculty	**Dr.**	**Robert**	**Kosara**	**Computer Science**
Investigator	**Mr.**	**Khaldoon**	**Dhou**	**Computer Science**

The Institutional Review Board (IRB) certifies that the protocol listed above is exempt under category 2 (45 CFR 46.101 2.b.4).

Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:

a) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and

b) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

This approval will expire one year from the date of this letter. In order to continue conducting research under this protocol after one year, the "Annual Protocol Renewal Form" must be submitted to the IRB. Please note that it is the investigator's responsibility to promptly inform the committee of any changes in the proposed research, as well as any unanticipated problems that may arise involving risks to subjects. Amendment and Event Reporting forms are available on our web site: http://www.research.uncc.edu/comp/human.cfm

M. L. 3/30/11

Dr. M. Lyn Exum, IRB Chair Date

The UNIVERSITY of NORTH CAROLINA at CHARLOTTE

APPENDIX B: EXPERIMENT 2 INITIAL APPROVAL

UNC CHARLOTTE

Office of Research Services

9201 University City Boulevard, Charlotte, NC 28223-0001
t/ 704.687.2291 f/ 704.687.2292 www.research.uncc.edu

Institutional Review Board (IRB) for Research with Human Subjects

Approval of Exemption

Protocol #	**13-01-34**			
Title:	**Size Judgment and Comparison in Tag Clouds**			
Date:	**2/11/2013**			
Responsible Faculty	**Dr.**	**Mirsad**	**Hadzikadic**	**Software & Information Systems**
Investigator	**Mr.**	**Khaldoon**	**Dhou**	**Software & Information Systems**

The Institutional Review Board (IRB) certifies that the protocol listed above is exempt under category 2 (45 CFR 46.101).

Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
a) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and
b) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

This approval will expire one year from the date of this letter. In order to continue conducting research under this protocol after one year, the "Annual Protocol Renewal Form" must be submitted to the IRB. Please note that it is the investigator's responsibility to promptly inform the committee of any changes in the proposed research, as well as any unanticipated problems that may arise involving risks to subjects. Amendment and Event Reporting forms are available on our web site: http://research.uncc.edu/compliance-ethics/human-subjects/amending-your-protocol or http://research.uncc.edu/compliance-ethics/human-subjects/reporting-adverse-events

M. L. Exum (signature) 2/12/13
Dr. M. Lyn Exum, IRB Chair Date

The UNIVERSITY *of* NORTH CAROLINA *at* CHARLOTTE

APPENDIX C: EXPERIMENT 2 AMENDMENT APPROVAL

UNC CHARLOTTE

Office of Research Compliance

9201 University City Boulevard, Charlotte, NC 28223-0001
t/ 704.687.1888 f/ 704.687.0980 www.research.uncc.edu/comp/complian.cfm

Institutional Review Board (IRB) for Research with Human Subjects

University of North Carolina at Charlotte

Approval of Amendment

Protocol #	**13-01-34**		
Title:	**Size Judgment and Comparison in Tag Clouds**		
Date:	**3/14/2013**		
Investigator	**Mr. Khaldoon**	**Dhou**	**Software & Information Systems**
Responsible Faculty	**Dr. Mirsad**	**Hadzikadic**	**Software & Information Systems**

The Institutional Review Board (IRB) has approved the amendment of the protocol listed above for Research with Human Subjects.

Please note that it is the investigator's responsibility to promptly inform the committee of any changes in the proposed research, as well as any unanticipated problems that may arise involving risks to subjects.

Amendment Details: The number of questions is increased from 64 to 96 to test participant judgment of size of words in a tag cloud. Words will be shown in regular and bold font rather than colored fonts. Recruitment may include word of mouth.

M. L. E. 3/15/13

Dr. M. Lyn Exum, IRB Chair Date

The **UNIVERSITY** ***of*** **NORTH CAROLINA** ***at*** **CHARLOTTE**
An Equal Opportunity/Affirmative Action Employer

APPENDIX D: EXPERIMENT 1 CONSENT FORM

UNC CHARLOTTE

Visualization Use Study – Consent Form

Project Title and Purpose. This study investigates how users perceive words in tag clouds. The results of this study will help improve visualizations in the future.

Investigators. This study is performed by Khaldoon Dhou and Dr. Robert Kosara of the Computer Science Department at UNC Charlotte. They can be reached at kdhou@uncc.edu and rkosara@uncc.edu, respectively.

Eligibility. You may participate in this study if you can comfortably communicate in written English and have 20/20 full color-vision or vision corrected to 20/20. You must also be at least 18 years old.

Task Description. You will be asked about your personal information: age and gender. After that, there are two questionnaires: the first one asks to compare two words and the second one is similar except that the two words appear in a tag cloud. You will receive $0.50 as a base rate and for each correct answer, you will receive $0.04. So, the maximum will be $4.34. An answer is considered correct if it is within the 15% of the actual value.

Anonymity, Risks. No personally-identifiable data is collected in the course of this study. There are no known risks from taking part.

Volunteer Statement You are a volunteer. The decision to participate in this study is completely up to you. If you decide to be in the study, you may stop at any time.

Statement of Fair Treatment and Respect UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the University's Research Compliance Office (704-687-3309) if you have any questions about how you are treated as a study participant.

This form was approved on 03/30/2011 for use for one year.

☐ **I have read and understood the above**

Continue

APPENDIX E: EXPERIMENT 2 CONSENT FORM

UNC CHARLOTTE

Visualization Use Study – Consent Form

Project Title and Purpose. This study investigates how users perceive words in tag clouds. The results of this study will help improve visualizations in the future.

Investigators. This study is performed by Khaldoon Dhou and Dr. Mirsad Hadzikadic of the Software and Information Systems Department at UNC Charlotte. They can be reached at kdhou@uncc.edu and mirsad@uncc.edu, respectively.

Eligibility. You may participate in this study if you can comfortably communicate in written English and have 20/20 full color-vision or vision corrected to 20/20. You must also be at least 18 years old.

Task Description. You will be asked about your personal information: age, gender and email (email is optional). After that, there will be a questionnaire that asks to compare two words that appear in a tag cloud. The participant who performs the most accurately in the user study will win a $25 Starbucks gift card. If you are a winner, we can only approach you by email. So, if you choose not to provide an email we cannot consider you for the prize in case you were a winner.

Anonymity, Risks. No personally-identifiable data is collected in the course of this study other than your email to be able to reach you in case you win. There are no known risks from taking part.

Volunteer Statement You are a volunteer. The decision to participate in this study is completely up to you. If you decide to be in the study, you may stop at any time.

Statement of Fair Treatment and Respect UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the University's Research Compliance Office (704-687-1871) if you have any questions about how you are treated as a study participant.

This form was approved on 03/14/2013 for use for one year.

☐ **I have read and understood the above**

Continue

APPENDIX F: MEAN VALUES IN EXPERIMENT 1

Table 8: Mean values of typeface size ratio judgment for the interaction between display type, appearance and relative size for the tag cloud absent display type in experiment 1

Decoration	Relative Size	Appearance		
		No Ascenders & Descenders	Narrow vs. Wide	Ascenders & Descenders
No Decoration	12 vs. 18	2.092	2.382	2.037
	12 vs. 24	2.538	3.092	2.511
	12 vs. 30	3.478	3.715	3.422
	12 vs. 36	4.105	4.712	4.160
Boxes	12 vs. 18	2.115	2.251	2.175
	12 vs. 24	2.702	3.038	2.772
	12 vs. 30	3.455	4.002	3.638
	12 vs. 36	4.063	4.674	4.232
Filled Areas	12 vs. 18	2.132	2.314	2.115
	12 vs. 24	2.685	3.017	2.751
	12 vs. 30	3.38	3.832	3.542
	12 vs. 36	3.929	4.574	4.366
Shadow	12 vs. 18	2.042	2.143	1.995
	12 vs. 24	2.514	3.049	2.749
	12 vs. 30	3.358	3.823	3.354
	12 vs. 36	4.028	4.691	4.14

Table 9: Mean values of typeface size ratio judgment for the interaction between display type, appearance and relative size for the tag cloud present display type in experiment 1

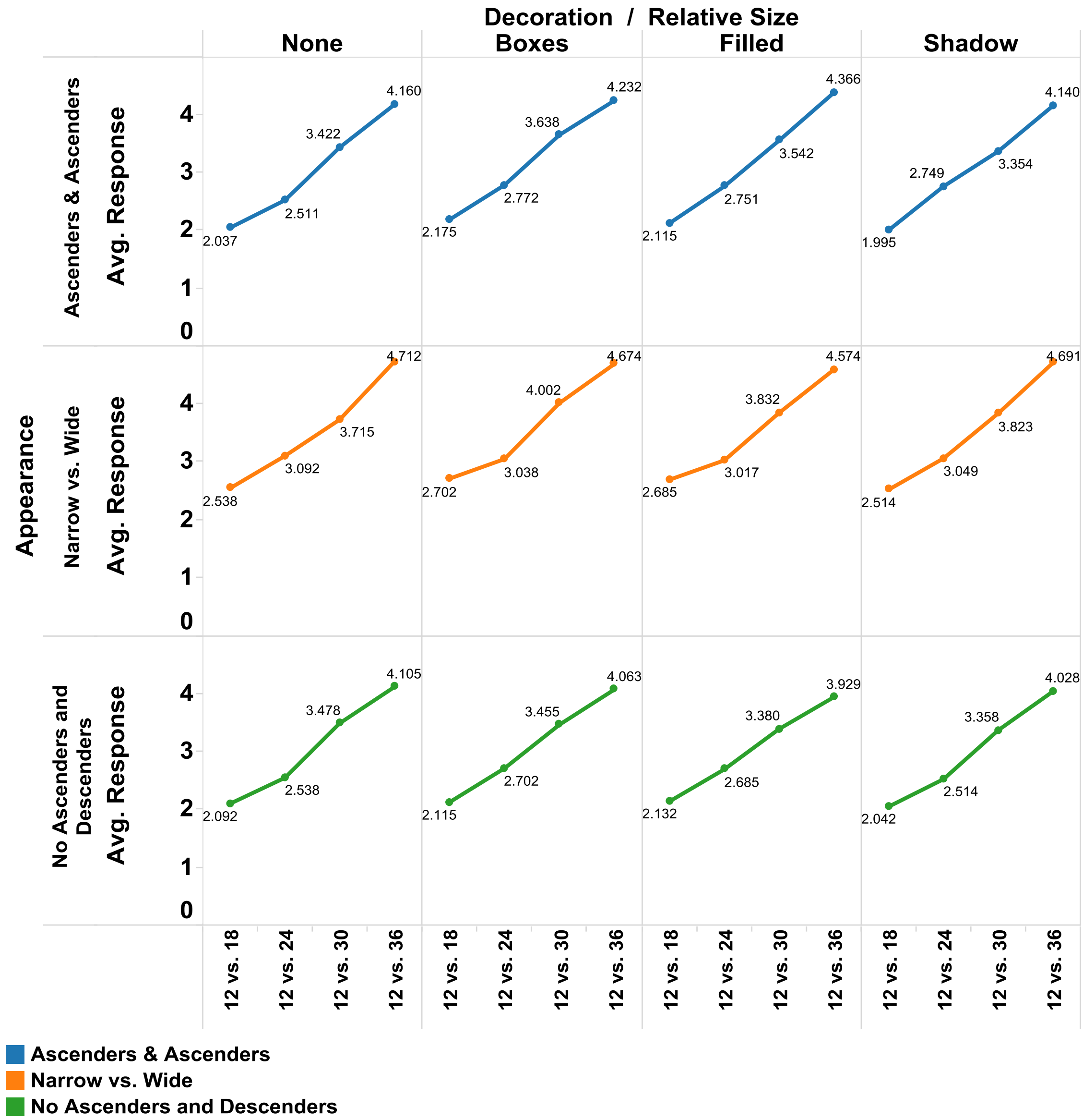
Decoration	Relative Size	Appearance		
		No Ascenders & Descenders	Narrow vs. Wide	Ascenders & Descenders
No Decoration	12 vs. 18	2.109	2.372	2.138
	12 vs. 24	2.558	3.065	2.834
	12 vs. 30	3.534	4.114	3.465
	12 vs. 36	4.415	4.654	4.325
Boxes	12 vs. 18	2.192	2.483	2.220
	12 vs. 24	2.931	3.280	2.815
	12 vs. 30	3.528	3.908	3.614
	12 vs. 36	4.343	4.748	4.372
Filled Areas	12 vs. 18	2.260	2.465	2.131
	12 vs. 24	2.897	3.162	2.957
	12 vs. 30	3.472	4.008	3.62
	12 vs. 36	4.248	4.672	4.357
Shadow	12 vs. 18	2.075	2.342	2.128
	12 vs. 24	2.798	3.078	2.897
	12 vs. 30	3.666	4.006	3.494
	12 vs. 36	4.351	4.803	4.366

APPENDIX G: MEAN VALUES IN EXPERIMENT 2

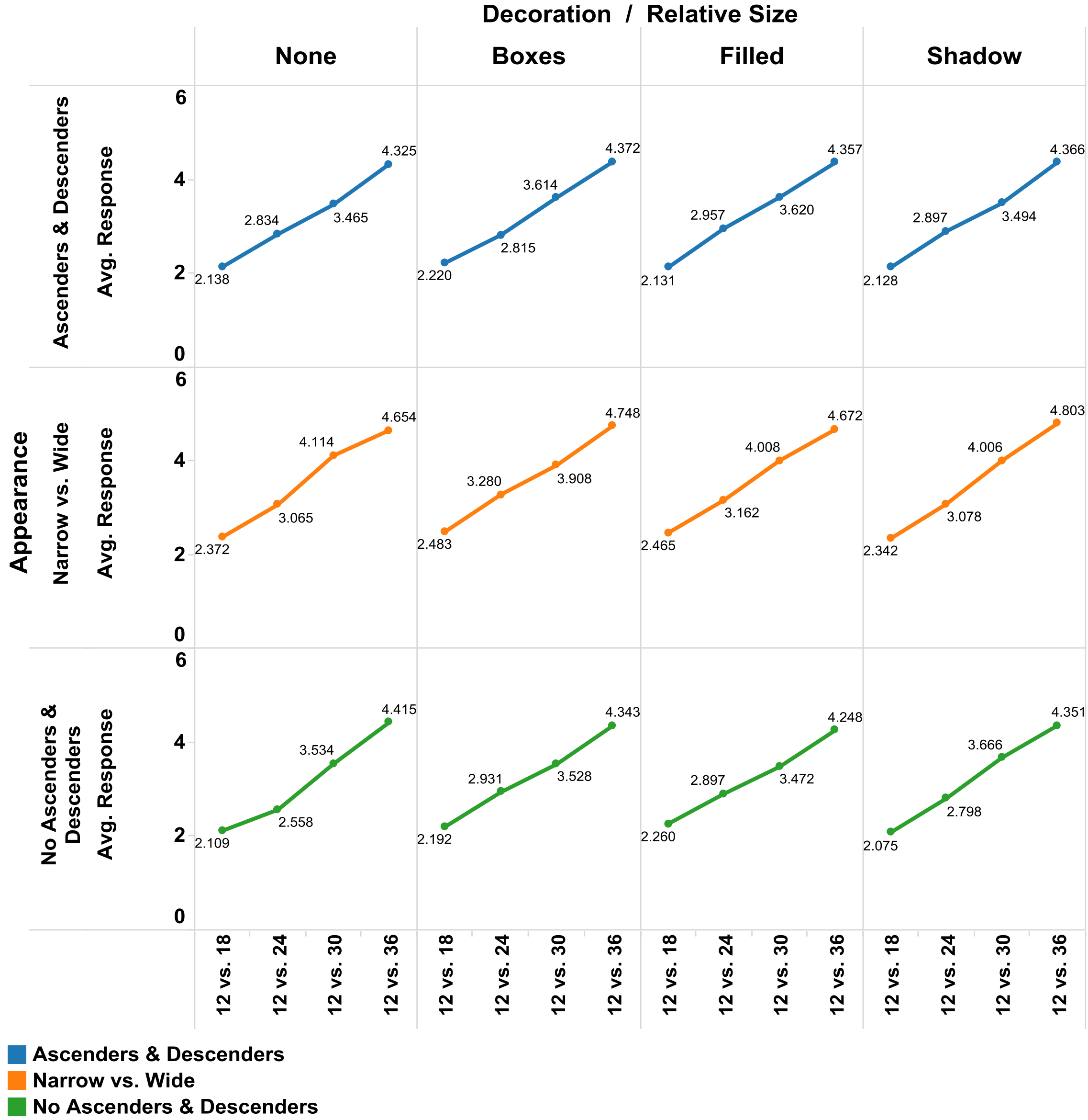
Table 10: Mean values of size ratio judgment for the interaction effect typeface style x horizontal x vertical x relative size in experiment 2

Typeface Style	Horizontal	Vertical	Relative Size	Mean
Regular	Left	Upper	18 vs. 24	2.013
			18 vs. 30	2.803
			18 vs. 36	3.514
		Lower	18 vs. 24	2.013
			18 vs. 30	2.826
			18 vs. 36	3.575
	Right	Upper	18 vs. 24	1.979
			18 vs. 30	2.844
			18 vs. 36	3.427
		Lower	18 vs. 24	2.058
			18 vs. 30	2.798
			18 vs. 36	3.414
Bold	Left	Upper	18 vs. 24	1.988
			18 vs. 30	2.923
			18 vs. 36	3.521
		Lower	18 vs. 24	2.076
			18 vs. 30	2.872
			18 vs. 36	3.609
	Right	Upper	18 vs. 24	2.141
			18 vs. 30	2.823
			18 vs. 36	3.473
		Lower	18 vs. 24	2.088
			18 vs. 30	2.802
			18 vs. 36	3.621

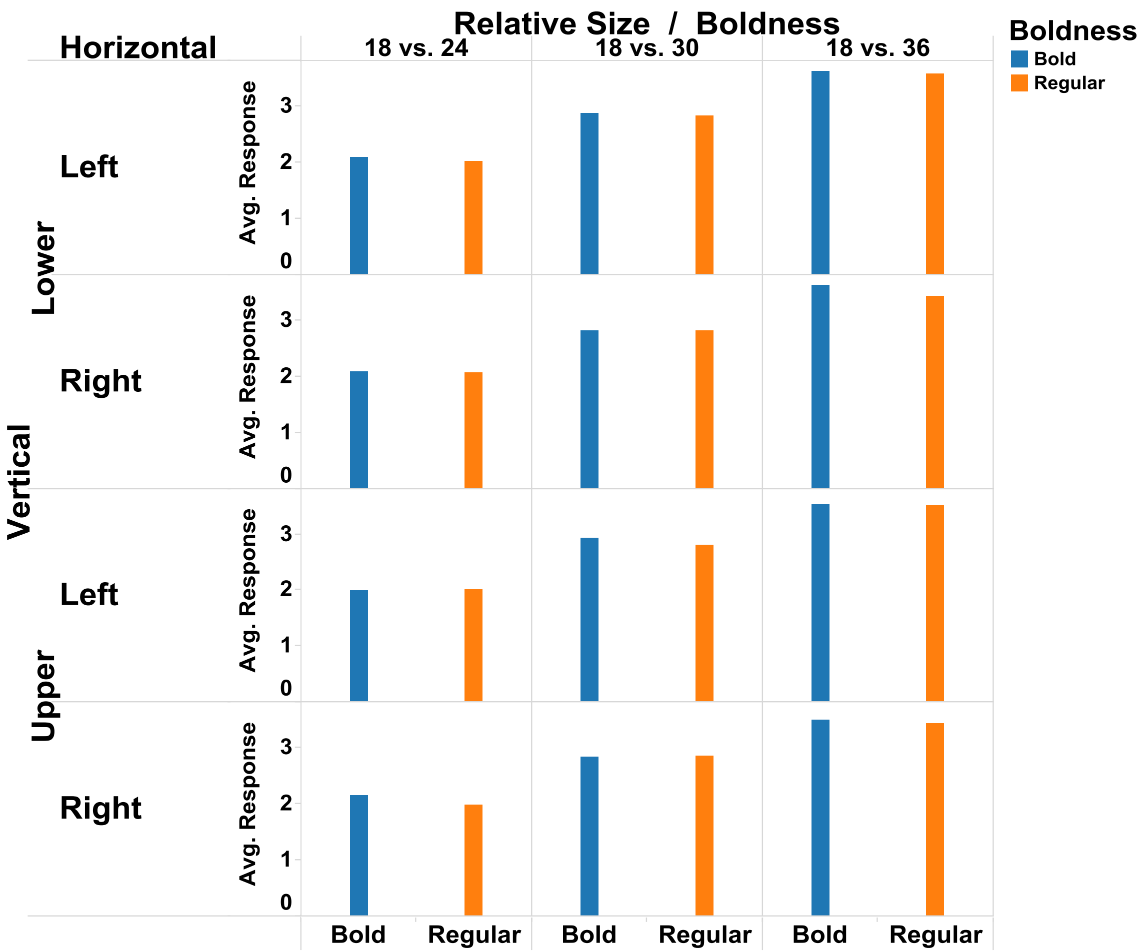
APPENDIX H: EXPERIMENT 1 VISUALIZATION (TAG CLOUD ABSENT)



APPENDIX I: EXPERIMENT 1 VISUALIZATION (TAG CLOUD PRESENT)



APPENDIX J: EXPERIMENT 2 VISUALIZATION



APPENDIX K: A WORDLE OF THE WHOLE DISSERTATION [7]