An Empirical Study on Using Visual Metaphors in Visualization

R. Borgo, A. Abdul-Rahman, F. Mohamed, P. W. Grant, I. Reppa, L. Floridi, M. Chen

Abstract—In written and spoken communications, metaphors are often used as an aid to help convey abstract or less tangible concepts. However, the benefits of using visual metaphors in visualization have so far been inconclusive. In this work, we report an empirical study to evaluate hypotheses that visual metaphors may aid memorization, visual search and concept comprehension. One major departure from previous metaphor-related experiments in the literature is that we make use of a dual-task methodology in our experiment. This design offers an abstraction of typical situations where viewers do not have their full attention focused on visualization (e.g., in meetings and classes). The use of the secondary task introduces "divided attention", and makes the effects of visual metaphors more observable. In addition, it also serves as additional masking in memory-based trials. The results of this study show that visual metaphors can help participants better remember the information depicted in visualization. On the other hand, visual metaphors can have a negative impact on the speed of visual search. The results also show a complex pattern as to the benefits of visual metaphors in helping participants grasp key concepts from visualization.

Index Terms—Visual metaphors, icons, cognition, working memory, long-term memory, visual search, evaluation.

1 Introduction

Visual metaphors are a form of non-linguistic metaphors and can be seen frequently in the visual arts, performing arts, advertisements, icons and signs, culture symbols, color symbolism, graphical user interface, and so forth. Ortony suggests that metaphors may aid communication and thought processes through *compactness*, *vividness* and *inexpressibility* [30]. In terms of visual metaphors, compactness facilitates the transfer of human "experience from well-known to less well-known contexts", vividness "impresses a more memorable learning" and understanding, and inexpressibility enables conveying "extra meanings" that are difficult to encode in a language [40]. Naturally, one cannot help but wonder whether these three features of visual metaphor can be transferred to positive effects in visualization, for example, to improve cognitive processes for memorization, visual search and concept grasping as illustrated in Fig. 1.

In written and spoken communications, metaphors are often used as an aid to help convey abstract or less tangible concepts. However, the benefits of using visual metaphors in visualization have so far been inconclusive. On one hand, there is a collection of common use of linguistic and visual metaphors, such as color symbolism [15], the semantic notions encoded in names of various visual representations such as "bubble", "bar", "flow", and "stream", the use of "pile" and "room" metaphors in document visualization [24, 16], and use of "tree" and "container" metaphors in treemap [46]. On the other hand, there have been overwhelming criticisms about inaccurate depiction of visual metaphors (e.g., Tufte [42]). There have been empirical studies showing that both 2D and 3D visual metaphors do not necessarily bring the desired advantages to information visualisation (e.g., [6, 39]). In this work, we re-examine the question about the effects of visual metaphors in visualization. The investigation is partly motivated by a new observation about typical situations where visualization often features visual metaphors, such as visualizations published in newspapers and magazines, and visualizations presented in meetings or lectures. In these situations, users often have a limited amount of time to view visualization images (e.g., glancing at a figure in a newspaper), or to not paying full attention (e.g., viewing a presentation slide in a meeting while reading emails on a mobile device). Most previous empirical studies in the field of visualization were conducted using the single task methodology, where participants usually focus their full attention on the task. It is possible that the effects of visual metaphors are more difficult to observe under such conditions. The reorganization of the common factor of "divided attention" enthused us to design and conduct an empirical study that reflects more closely these real-life situations. *Dual task methodology*, which requires concurrent performance of two tasks by a participant, has been used for evaluating human perception and performance in psychology [33, 9]. We found that the *subsidiary task paradigm* [18] offers a better abstraction of the real-life situations to be encapsulated. The emphasis was placed on the primary task that consists of 72 trials designed to record participants' performance in relation to visualization with and without visual metaphors. The secondary task was designed to introduce "divided attention" in order to reduce slightly the participants' cognitive capacity available for the primary task [33]. It is necessary to emphasize that the focus of this work was on the effects of visual metaphors rather than attention or workload assessment. The dual task methodology was adopted to make effects of visual metaphors more apparent.

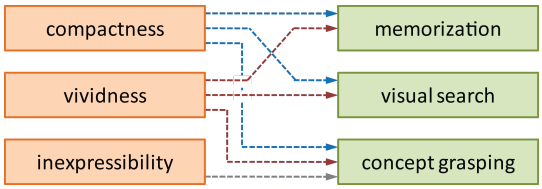


Fig. 1: Hypothesized transition from features of visual metaphors to effects on visualization.

As illustrated in Fig. 1, we hypothesized that the three features of metaphors might be transformed to three effects in cognitive processes. Based on these three possible effects, we proposed the following hypotheses to be evaluated in this work:

H1. Visual metaphors may help participants remember better in visualization (in terms of working memory).

H2. Visual metaphors may help participants remember better in visualization (in terms of long-term memory).

H3. Visual metaphors may help participants perform visual search more efficiently during visualization.

H4. Visual metaphors may help participants grasp the concepts shown in visualization more effectively.

- *Rita Borgo and Philip W. Grant are with Computer Science, Swansea University, E-mail: {r.borgo,p.w.gran}@swansea.ac.uk.*
- *Alfie Abdul-Rahman and Min Chen are with Oxford e-Research Centre, University of Oxford, E-mail: alfie.abdulrahman@gmail.com, min.chen@oerc.ox.ac.uk.*
- *Farhan Mohamed is with Universiti Teknologi Malaysia, E-mail: farhan@utm.my.*
- *Irene Reppa is with Psychology Department, Swansea University, E-mail: I.Reppa@swansea.ac.uk.*
- *Luciano Floridi is with St. Cross College, University of Oxford, E-mail: luciano.floridi@philosophy.ox.ac.uk.*

In the remainder of the paper, we describe the related work, the design of the stimuli, the process of the study, the analysis of the study results and our findings.

2 Related Work

Metaphor and Terminology. *Metaphor* is an important subject in linguistics (e.g., [14, 7]), philosophy (e.g., [4, 8]), and psychology (e.g., [12, 10, 31, 11]). In poetry for example metaphors are classified as rhetorical figures and considered forms of expression capable of reflecting a primeval knowledge typical of the human nature, a knowledge of the external world built upon similarities and analogies between inanimate and living objects. In artistic representations the use of metaphors correspond to what we linguistically call "abstraction".

The English word *metaphor* is derived from *meta* in Greek meaning "between" and *phero* meaning "to bear" or "to carry". Metaphor is a form of *figure of speech* that enables rhetorical use of words in some distinguish ways [25, 36]. In linguistics, there are many forms of figures of speech, ranging from *antithesis* to *irony*. Among these forms, *metaphor*, *simile*, *metonymy*, *hyperbole*, *personification*, and *synecdoche*, that are often considered together as a subset of figures of speech, where a primary concept is expressed by making use of a secondary concept. Each of these represents different stylistic emphasis. For example, similes typically make explicit uses of words "like" and "as", while metonymies exhibit implicit association through substitution. Synecdoche stresses the part and whole relationship between the two concepts, and personification focuses on association between objects and people. Metaphors emphasize the familiarity and tangibility properties, while hyperboles suggest some exaggerations.

There is no similar detailed categorization for visual communication where some linguistic subtleties would not normally be encoded in visual representations. The only term commonly used is *visual metaphor*, where a visual image is used to convey a primary concept by making use of a secondary concept. We hence adapted the term *visual metaphor* to encapsulate all such association between the two concepts with the goal to assist the expression of a primary concept. A possible alternative would be "figures of visualization", but it would be very confusing. We refer the primary concept as the *tenor* and the secondary as the *vehicle* (or *target* and *source* in some literatures).

Researchers in cognitive science has argued that mental representation is, at least in part, metaphorical. Rather than seeing metaphors as solely a primarily linguistic phenomenon it has been proposed as a mode of representation and thought. The reasoning behind it being that the complexity in the representation of certain aspects of our knowledge, and therefore their representation in terms of easier-to-understand domains, is metaphorical [19, 20].

Metaphor in Visualization. Data visualization focuses on the graphical representation of abstract information with two main purposes: data exploration and analysis, and knowledge sharing and communication. However, visual metaphor has drawn limited attention in the visualization literature. Pang and Clifton proposed to use everyday objects to create intuitive 3D interface [32]. Averbukh analyzed the role of metaphors in the theory of computer-human interface [2]. Ziemkiewicz and Kosara examined how the structure of a visual metaphor can influence the processing of the information [46]. Healey studied colour symbolism [15], and Huggins and Entwisle on iconic communication [17]. Aley wrote a short book on metaphorical visualization [1], which presents many applications where "vivid" metaphors are used in visualization.

Ware [43] provides a comprehensive review of the principles behind visual thinking and cognition and how to apply such knowledge to data visualization. Many visualization books included Chernoff faces as positive examples of metaphoric visualization [42], and it was studied in the context of visualization by [28]. McDougall et al. [26] and Blackwell [5] provide detail guidelines on icon design, while Reppa et al. [37] analyzed the impact of icon design on performances of both familiarity and aesthetic appeal. Familiarity is a concept probed by [47] which is related to metaphor usability [45].

Our work is intended to study the impact of visual metaphors on the user performance in visualization, building on the theory of conceptual structure by [29]. We will draw connections with the relevant literature in perception and cognition in individual sections where the previous works are the most relevant.

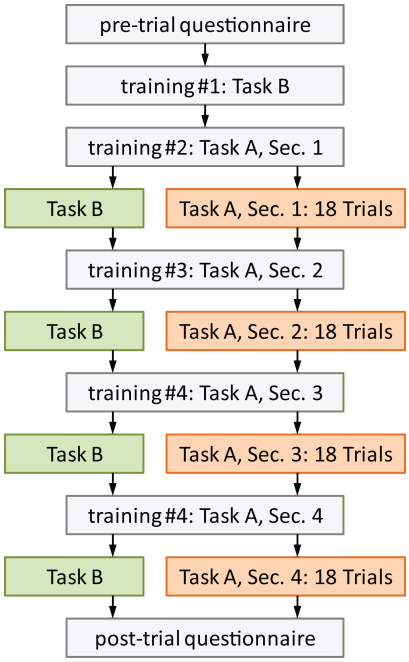


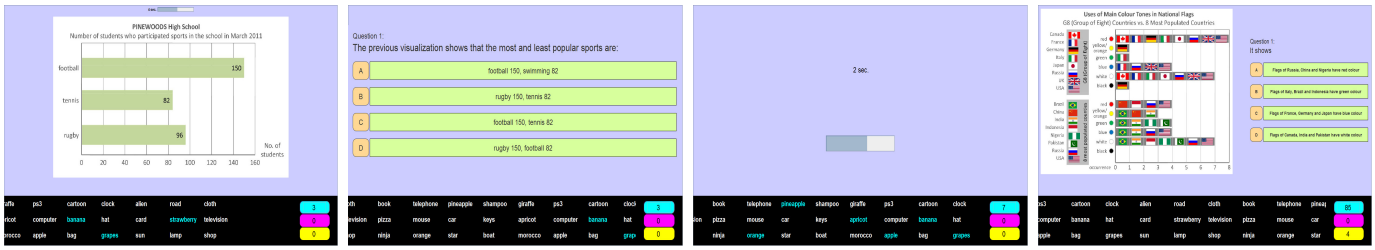
Fig. 2: Activity sequence of the study.

3 Experiment Overview

Over the last few years, there have been a number of approaches proposed to analyse visual metaphors as conceptual phenomena. One of the more well known of these approaches is that of Conceptual Metaphor Theory (CMT). CMT proposes that verbal metaphors are merely surface manifestations of metaphorical thought and that "[a] metaphor is fundamentally conceptual, not linguistic in nature" [21]. Although the nature of visual metaphors has been investigated, much of this work remains largely theoretical, with few empirical studies, therefore with this current work we have devised an experiment to determine the role of visual metaphors in the comprehension and knowledge acquisition process. The experiment was structured into four main sections each probing a perception or cognitive related aspect.

In order to provide our study with an intuitive scenario that all the participants could easily understand, we decided to focus on simple statistical representation of data, and chose the common graphical representations of 2D histograms/bar charts and bubble charts. The visual stimuli were selected from visual representations common in everyday life. To increase stimuli reliability the chosen visual metaphors underwent scrutiny by a small but ethnically diverse group of people. In our experiment we maintained a one-to-one correspondence between our metaphorical and non-metaphorical stimuli, e.g. every non-metaphorical stimuli had a corresponding (and equivalent) metaphorical one, and vice-versa.

In designing our experiment, we considered the following factors:
Data Focus/Concreteness. Concreteness (as opposed to abstraction) which indicates the degree of pictorial resemblance that a visual representation bears to its counterpart [26], is somehow in opposition to visual complexity; concrete symbols tend to be more visually obvious because they depict objects, places, and people that are already familiar in the real world, abstract symbols, in contrast, represent information using graphic features such as shapes, arrows, and so on. One of the reasons why concrete symbols are more visually obvious may simply because the extra detail provided makes them easier to use. In contrast, however, design guidelines typically suggest that the design of symbols or icons should be kept as simple as possible. Other researchers have focused on the fact that concrete symbols are more meaningful than abstract symbols.



(a) Stimulus in § 1, 2, 4. (b) Multiple Choice Q&A in § 1, 2, 4. (c) Pause Masking in § 1, 2. (d) Stimulus and Q&A in § 3.

Fig. 3: Screenshots from the experiment interface: a)-b) example of stimulus and multiple choice question for Section 1, 2 and 4; c) example of gray masking screen; d) example of stimulus and multiple choice question for Section 3.

Visual Complexity. Complexity is defined as the amount of detail in the visual representation. Complexity is a direct function of: the degree of perceivable structure, variety of parts and separation of parts vs. their conceptualization as a whole. Visual search can be considerably influenced by the complexity of a visual representation. Visual complexity is a stimulus characteristic that has been shown to influence not only performance in perception tasks such as visual search, but also subjective appraisals of appeal [37].

Meaningfulness. Meaningfulness indicates the relationship between what is depicted in the visual representation and the function it refers to. Research suggests that particularly meaningful icons in displays can capture human attention and drive visual search [23], if the meaning can be extracted at the same time the visual search is performed then usability can be enhanced by ensuring that key visual representations, or features, stand out during search.

Semantic Distance. Semantic, or articulatory, distance is a measure of the closeness of the relationship between the symbol and what it is intended to represent. A number of classification systems have been developed in order to attempt to characterize the different relationships that occur between symbols and their functions [34].

Familiarity. Familiarity reflects the frequency with which symbols are encountered. This property is thought to be an important determinant of usability. It is evident that user performance improves dramatically as a result of learning symbols and signs. The effects of some symbol characteristics on performance, such as color and concreteness, diminish as symbols become more familiar but others, such as complexity, do not.

Icon-based metric. Behind the design of every icon there is a visual syntax. It is therefore possible to numerically measure the complexity of an icon by summing up its "syntactical" components such as letters, lines, arrows and so on.

All of these factors could influence both the perceptual and cognitive load of a task. We thus designed our user study to capture such influence by measuring the variation in task performance. Our focus was on visual perception and cognitive speed-focused tasks that leverage cognitive abilities common to tasks where multiple streams of information are analyzed. Such analytical tasks are commonly found in situations of divided attention due to interruptions of a primary task by either unforeseeable events or by the requirement of engaging in dual- or multi-tasking, making it hard to predict when information can be attended to. This study therefore followed a divided attention design. An interruptible context was created by enforcing attention to switch from one task to another, with the interruption being either relevant or a simple distraction. We devised two main tasks, referred to as Task A and Task B in Fig. 2, which were run in parallel and had to be executed simultaneously by the participants. Both tasks engaged simultaneous visual signals, therefore the same information channels were used. The structure of Task B, acting as the distractor, remained constant throughout the entire experiment. Task A was the primary task and therefore changed its structure depending on the main focus of the trial. As it is not feasible to explore the effects of all combinations of the aforementioned different factors (e.g., concreteness, complexity etc.), we divided our primary task, Task A, into four main sections. Each section reflected typical tasks performed when analyzing both metaphorical and non-metaphorical data representation. We found noticeable performance variations between the tasks performed in each section, which likely reflect the different levels of perceptual and cognitive loads. Within each section a task was composed of 18 trials. To be able to perform a more detailed investigation and comparison between trials in different sections each group of 18 trials was designed in such a way that the first 3 stimuli of each section had similar visual features, so had the following two groups of 3 stimuli. Performance was assessed by analyzing both accuracy and response time (RT). However, for section 3 (visual search) RT results were collected as the primary factor because participants were encouraged to focus on accuracy and were allowed to take as long as they wished to perform the trials. Aspects like *subjective rating* were collected in a separate questionnaire. Aesthetic appeal was a secondary factor not covered within the scope of our study. The study and tasks are described in detail in the following sections. Fig. 2 outlines the workflow of our study.

4 Tasks

Participants performed two main tasks: a primary task (Task A in Fig. 2) and a secondary task (Task B in Fig. 2). Task A was subdivided into four main sections each probing a specific aspect of the exploratory process typically conducted by a generic user browsing information organized and represented via both spatial and non-spatial visualizations. Task B, running across the entire duration of the study, was also performed. This parallel task acted as a distracting factor and helped us mimic interference effects which are unavoidable when performing every day tasks. Fig. 8 shows the temporal correspondence between the primary and secondary tasks for the best recorded performance. Only a few actions for Task B were performed during a question and answer phase (light or dark green). This pattern is common among all participants, indicating that Task B does not affect metaphoric or non-metaphoric visualization differently.

4.1 Primary Task and Stimuli

The primary task, which is referred to as Task A in communication with participants, is designed to evaluate the four hypotheses mentioned in Section 1. The stimuli for this task are organized into four experimental sections corresponding to the four hypotheses. Each section has 18 stimuli including 9 with visual metaphors and 9 without. There are thus 72 stimuli in total for the trials and 4 additional stimuli for training.

In order to evaluate all four hypotheses with a common experimental setting in the context of visualization, we adopt the basic format of multiple choice question-and-answer to collect participants' performance in terms of both accuracy and response time. This basic format is familiar to all participants, hence requiring little learning effort. To address the needs of different hypotheses, we design a specific temporal format for each section.

The design of the stimuli needs to avoid a number of confounding effects, including

- *Knowledge bias* – In order to capture the effects of visual metaphors on memorization, visual search and concept comprehension, stimuli used in the study have to feature a variety of concepts and data associated with such concepts. Familiarity

about the concepts and data used in the stimuli can affect the performance of the participants.

- *Ordering bias* – Stimuli presented in the earlier part of the study may have positive and negative effects on the stimuli presented later (e.g., learning). The order of multiple choices may certainly affect the time required to reach the correct answer as participants may choose not to read all optional answers.

- *Attention bias* – It is unavoidable that some participants may experience tiredness or attention lapses, which could affect the participants' performance in different phases of the study.

Like most empirical studies, such confounding effects cannot be completely eliminated, but should be reduced to the level such that they will not have noticeably impact on the performance of participants. For example, as illustrated in Fig. 2, we ran the four sections separately in a fixed sequence order in order to avoid the confounding effect that may be resulted from mixing tasks of different temporal formats associated with individual sections. We also placed the training for each section at the beginning of the section to make familiarization immediately relevant, and introduced a short break between each section to ease the tiredness and reduce attention bias.

To minimize biases in the design, we organized the stimuli design in a structured manner, though the structure was removed during the actual trials through pseudo-randomization. In each section, the 18 stimuli are organized into three groups of 6 stimuli, 3 with visual metaphors and 3 without. For sections 1-3, stimuli in each group use the same visualization styles, i.e., vertical bar chart, horizontal bar chart and bubble chart. For section 4, the stimuli are grouped by the estimated levels of cognitive load. All stimuli were designed in pairs. Some examples of the stimuli are shown in Figs. 4, where stimuli are juxtaposed in pairs, one pair from each group, and three pairs in each section. All stimuli, together with questions and optional answers, can be found in the supplementary materials. As we cannot use the same data for different stimuli, we designed each pair carefully to ensure that they showed similar concepts, had similar visual design (except the use of metaphors), and represented a similar level of cognitive load. In addition, we ensured that correct answers for the stimuli in each pair were placed at the same position to avoid order bias. All 18 stimuli in each section were pseudo-randomized, so participants would not be able to reason about the order of stimuli with or without metaphors, or to guess the likely position of a correct answer. The reason for devising a specific pseudo-randomization scheme for each section is to ensure that in each pair, a stimulus with visual metaphors have the exactly 50% chance to be shown before the stimulus without visual metaphors, and vice versa. The scheme also ensures that two stimuli in the same pair are separately by at least 5 other stimuli (i.e., show distance > 5).

Each stimulus is a 1360×878 visualization image, corresponding to a unique dataset. Most datasets are synthetic, but some emulate real world concepts and data distribution. Some datasets are obtained from public domain sources (e.g., Wikipedia), and questions and optional answers were carefully designed to ensure that a correct answer could not easily be derived simply from a priori knowledge without viewing the stimuli.

4.1.1 Stimuli for Hypothesis 1: Working Memory

The 18 stimuli in this experimental section were designed to evaluate a hypothesis (H1) that visual metaphors may help participants remember better in visualization (in terms of working memory). Atkinson and Shiffrin's multi-store model [1] is commonly accepted as an explanation of how human memory works. It suggests that human memory is composed of three main stages whose structural features can be summarized as: sensory, working (or short-term) and long-term memory. In terms of human vision, sensory memory is also referred to as iconic store. In this work, we followed the widely accepted notion that information about a visual stimulus remains in the sensory memory for less than 1 second [3], and in working memory between 15-30 seconds unless one attentively rehearses the information to increase its retention [35]. In working memory written text competes with visually encoded information for storage. If multiple data attributes are integrated into an integrated visual representation, working memory may hold more information [21].

In this experimental section, there are 6 vertical bar charts (WM 1.1, 1.2m, 1.3, 1.4m, 1.5, 1.6m), 6 horizontal bar charts (WM 2.1, 2.2m, 2.3, 2.4m, 2.5, 2.6m) and 6 bubble charts (WM 3.1, 3.2m, 3.3, 3.4m, 3.5, 3.6m). All those tagged with a letter "m" are stimuli with visual metaphors. We use clipart pictures, icons and photographic images for visual metaphors. All visual metaphors are additional to the text labelling. In bubble charts, the data values correspond to the areas of circles, and all metaphors are contained within the circles.

In each trial, a stimulus is first displayed for 9 seconds. It is then replaced by a gray masking screen for 5 seconds. This is a common masking technique in perception studies for cleaning up the sensory memory. The masking effect is further enhanced by the concurrent Task B (see Section 5). After the gray screen, a question regarding the previously stimulus is presented with four optional answers. For example, the question and multiple choices for stimulus WM1.1 are:

The two numbers shown in the previous visualization are:
A. hot chocolate 68 and tea 37
B. coffee 68 and tea 37
C. tea 68 and coffee 37
D. coffee 68 and hot chocolate 37

The correct answer in this trial is C. The response time between when the question is shown and an answer is selected is recorded.

4.1.2 Stimuli for Hypothesis 2: Long-term Memory

It is not known by which mechanisms novel visual representations are stored in long-term memory. Nevertheless, there is evidence that working memory plays an important role in the formation of long-term memory. Information may be gradually transferred from working memory into long-term memory. The more frequent the information is repeated or used, the more likely it will eventually end up in long-term memory, or be "retained".

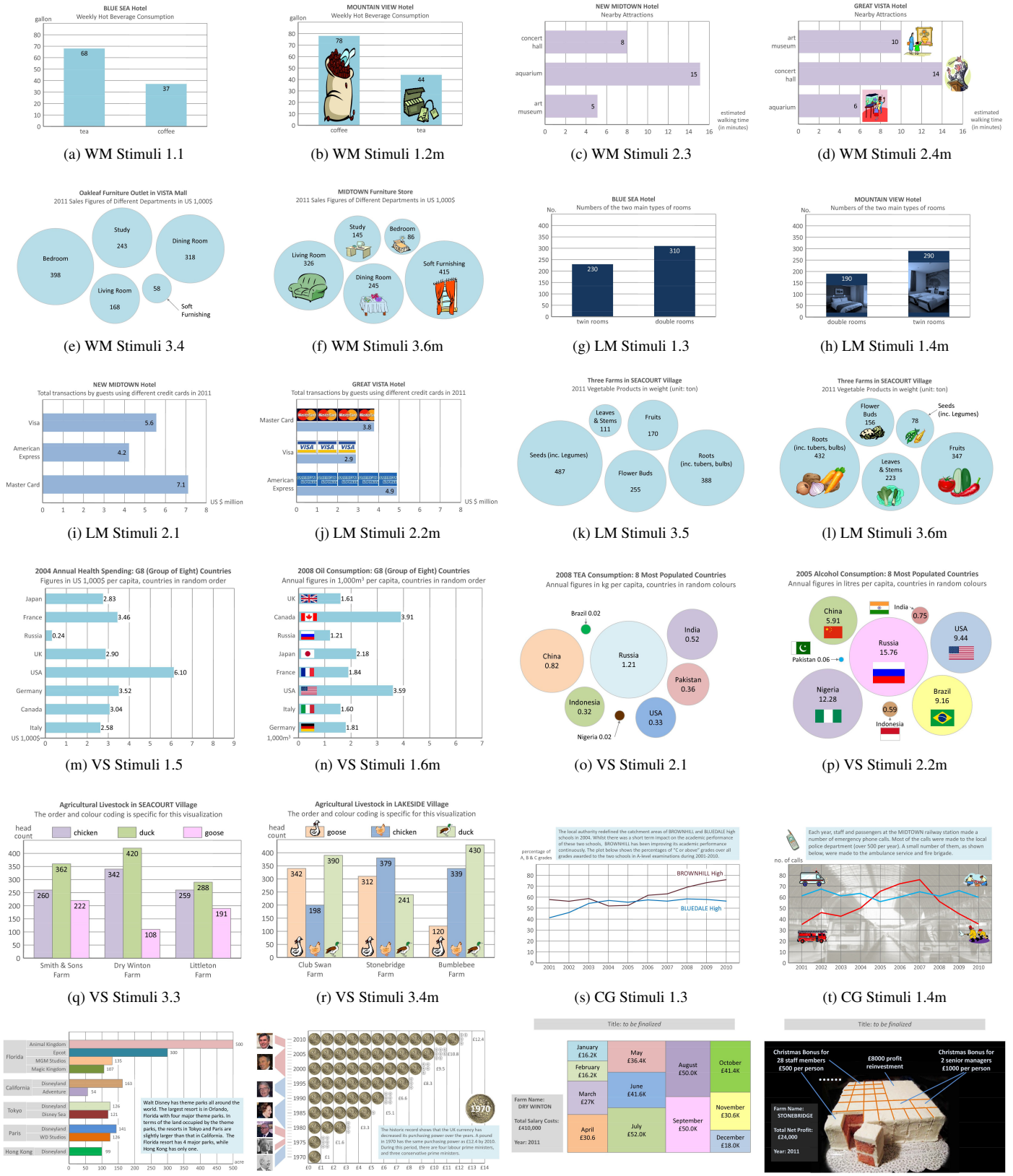
Because of the temporal and functional difference between working memory and long-term memory, it is necessary to examine the hypothesis that visual metaphors may help participants remember better in visualization under different conditions in order to separate the effects of these two types of memory. Most studies in the literature show that working memory lasts for 10-20 seconds [35], and there are also suggestions that it can last for up to 30 seconds [13]. We thus selected a safe threshold of 30 seconds for masking the effects of working memory. In a format similar to the section for working memory, each stimulus is displayed for 9 seconds. This is followed by a gray masking screen which is then shown for 30 seconds. During this period, Task B provides further masking effect for removing the information about the stimulus from the working memory. A multiple choice question is then presented to the participant.

We made every effort to ensure that trials for working memory and long-term memory are comparable across these two experimental sections. Stimuli in the corresponding groups (i.e., vertical and horizontal bar charts and bubble charts) have similar visual design and cognitive load. Fig. 4 shows 6 of the 18 stimuli used in this section.

4.1.3 Stimuli for Hypothesis 3: Visual Search

Visual search is an integral part of visualization, and has an important role in the cognitive process. Visual search occurs as a sequence of active visual queries operating through a focusing of attention while relying on perceptual cues. Stimuli in this experimental section were designed to evaluate a hypothesis (H3) that visual metaphors may help participants perform visual search more efficiently during visualization.

While the visual designs used in this section are horizontal and vertical bar charts and bubble charts, the amount of data depicted in the stimuli is significantly increased. The temporal format is also changed to address the need for focusing the performance evaluation on visual search rather than other facts such as the speed of reading.



(a) WM Stimuli 1.1 (b) WM Stimuli 1.2m (c) WM Stimuli 2.3 (d) WM Stimuli 2.4m

(e) WM Stimuli 3.4 (f) WM Stimuli 3.6m (g) LM Stimuli 1.3 (h) LM Stimuli 1.4m

(i) LM Stimuli 2.1 (j) LM Stimuli 2.2m (k) LM Stimuli 3.5 (l) LM Stimuli 3.6m

(m) VS Stimuli 1.5 (n) VS Stimuli 1.6m (o) VS Stimuli 2.1 (p) VS Stimuli 2.2m

(q) VS Stimuli 3.3 (r) VS Stimuli 3.4m (s) CG Stimuli 1.3 (t) CG Stimuli 1.4m

(u) CG Stimuli 2.5 (v) CG Stimuli 2.6m (w) CG Stimuli 3.5 (x) CG Stimuli 3.6m

Fig. 4: Examples of stimuli used in the study, where those tagged with a letter "m" are stimuli with visual metaphors.

Before each stimuli is shown, the corresponding question and optional answers are presented to the participant for 9 seconds. For example, for stimuli VS 1.5 (Fig. 4m), the preview screen shows: Please read the question first. The visualization will appear soon. It shows:
A. UK 6.10, Germany 3.52, France 2.83
B. UK 2.90, Germany 3.52, France 3.46
C. UK 2.90, Germany 3.04, France 2.83
D. UK 6.10, Germany 3.04, France 3.46

After 9 seconds, the stimulus is then presented, together with the question and optional answers. The participant is allowed to answer the question only in this phase. Each trial requires the participants to perform multiple visual searches, which removes some of knowledge bias in terms of familiarity of certain national flags or certain images.

4.1.4 Stimuli for Hypothesis 4: Concept Grasping

The stimuli in this experimental section were designed for evaluating a hypothesis (H4) that visual metaphors may help participants grasp the concepts shown in visualization more effectively. Here we focus on *concept grasping* as an element of the more complex cognitive processes of information gathering, concept understanding and semantic reasoning. The term *concept grasping* gives an emphasis on noticing key or important concepts depicted in visualization.

Unlike stimuli in experimental sections 1-3, all stimuli in this section do not have titles. The questions presented to the participants are more or less in the form of asking for the identification of some key concepts that would otherwise be in the title. Each stimulus is displayed for 9 seconds. This is immediately followed by a question and four optional answers. The participant is allowed to answer the question as soon as it is presented.

The 18 stimuli are divided in 3 groups with 6 stimuli each. The first group has three pairs of stimuli, 2 pairs show vertical bar charts and 1 pair shows line graphs. Each stimulus is accompanied by a textbox that provides some commentary about the visualization, including some relevant remarks as a distraction. For example, for CG 1.4m (Fig. 4t), the textbox reads as: "*Each year, staff and passengers at the MIDTOWN railway station made a number of emergency phone calls. Most of the calls were made to the local police department (over 500 per year). A small number of them, as shown below, were made to the ambulance service and fire brigade.*"

As the visualization shows only the data for ambulance service and fire brigade, the mention of police department in the text is the distraction. The question and optional answers for this stimulus is:

The previous visualization compares the calls for:
A. police department and ambulance in a shopping centre
B. ambulance service and fire brigade in a railway station
C. ambulance service and fire brigade in a shopping centre
D. police department and ambulance in a railway station

The second group of six stimuli has a similar visual appearance as the first group. However they contain more information and are slightly more cluttered than the first group. This is because a participant's performance of concept grasping depends on the amount of information in a stimulus and how long the participant is allowed to view the stimulus. If the information/time ratio were too low, the trial would not be able to differentiate different performance as most would perform well. If the information/time ratio were too high, the trial would not be effective either as most would perform badly. Hence by having two different levels of difficulties, we prevent the experiment from swinging too much either side.

The third group of six stimuli focus on the concept of *part and whole* that is implicitly encoded in some visualization such as a bubble chart. The three metaphoric stimuli feature cakes as the metaphor of cake sharing. The three non-metaphoric stimuli feature similar datasets presented in the corresponding geometric shapes. For example, CG3.6m shows a square cake being divided, while CG3.5 shows a one-level treemap (Fig. 4x and Fig. 4w respectively). All questions for this group of stimuli are in the form of asking for "the most suitable short title for the previous visualization", chosen from 4 optional answers. All correct answers feature either the word "proportion" or "distribution". All incorrect answers feature words such as "correlation", "trend", "imbalance", "variation", "asymmetry", "ordering".

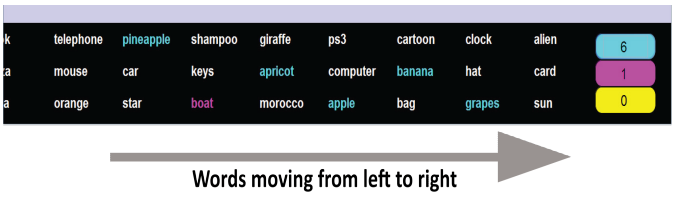


Fig. 5: User study secondary task.

4.2 Secondary Task and Stimuli

Only when the tasks are very basic (e.g., responding to a simple signal as soon as it occurs in either of the two modalities) is performance unimpaired under divided attention conditions. In devising our secondary task we followed the following guidelines for designing divided attention tasks [38]:

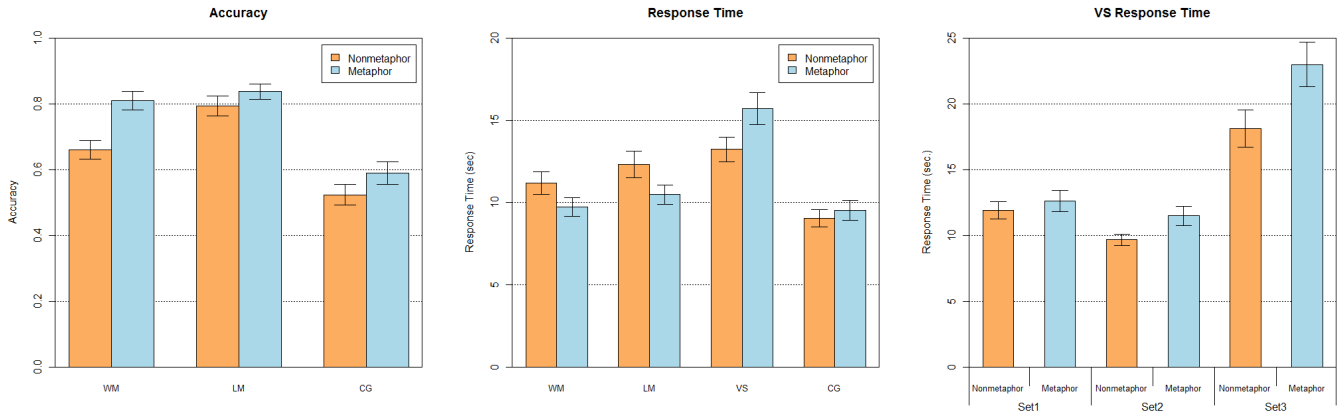
- *Task Difficulty Recommendation*: When more than one task must be done at once, efforts should be made to keep the difficulty level of the tasks as low as possible.
- *Task Sensory Channels Recommendation*: Where possible, the number of potential sources of information should be minimized.
- *Task Priority Recommendation*: Where time-sharing is likely to stress a person's capacity, the person should be provided with information about the relative priorities of the tasks so that an optimum strategy of dividing attention can be formulated.
- *Task Similarity Recommendation*: Tasks to be performed simultaneously should be made as dissimilar as possible in terms of demands on processing stages, input and output modalities, and memory codes.
- *Task Memory Recommendation*: When manual tasks are time-shared with sensory or memory tasks, the greater the learning of the manual task, the less will be its effect on the sensory or memory tasks.

Our secondary task (Task B) consisted of a sequence of words appearing at the bottom of the screen moving horizontally, like crawling (vs. rolling text) in films, across the screen from left to right. Participants were required to point and click at any fruit word that appeared on the screen. The total set of words which would appear in the list was composed by instances of simple commonly used English terms. A correctly selected word changed its color from white to cyan, a wrongly selected word from white to magenta. Three counters at the bottom right hand corner of the screen kept the count of how many fruit words had been correctly selected, how many had been missed and how many words had been wrongly selected. The counters were respectively colored in cyan, yellow and magenta. Fig. 5 shows a close up of the secondary task layout.

5 User Study Design

Participants. A total of 35 participants (16 females, 19 males) took part in this experiment in return for partial course credit or a £10 book voucher. Participants belonged to both the student and working communities and were recruited from Swansea University and related communities, with a very large variety of disciplines including Psychology, Humanities, Engineering and Economics. Ages ranged from 18 to 42 (Mean=23.7, SD=4.8). All participants had normal or corrected to normal vision and were not informed about the purpose of the study at the beginning of the session.

Apparatus. Visual stimuli were created using custom software that was written in Java. Stimuli were saved as static images and presented to participants using a custom made interface. Experiments were run using Intel Dual-Core PCs running at 2.13 GHz, with 2 GB of RAM and Windows 7 Professional. The display was 19" LCD at 1280x1024



(a) Accuracy. (b) Response Time. (c) VS Response Time.

Fig. 6: Performance analysis for all four hypotheses: a) accuracy for Working Memory (WM), Long-term Memory (LM) and Concept Grasping (CG) hypotheses; b) response time for Working Memory (WM), Long-term Memory (LM), Visual Search (VS) and Concept Grasping (CG) hypotheses; c) example of set based analysis for the response time for Visual Search (VS) hypothesis.

resolution with a 32bit sRGB color mode. Each monitor was adjusted to have same brightness and same level of contrasts. Participants interacted with the software using a standard mouse at a desk in a dimmed experimental room.

Procedure. The experiment began with a brief overview read by the experimenter using a predefined script. Detailed instructions were then given through a self-paced slide presentation. Brief descriptions of the requirements of each task were also provided. The experiment was divided into a primary task (Task A in Fig. 2) and a secondary task (Task B in Fig. 2). Within the primary task each participant completed a total of 72 trials, separated into 4 blocks of 18 trials. The 4 blocks were always completed in sequential order. Given the nature of the experiment each block assessed a different aspect of the cognitive process. Maintaining the same block order for each participants meant that each participants experienced similar experimental conditions. This allowed for a more sound analysis of the responses. Randomness was introduced at trial level. Within a given block, trials were randomized to avoid learning effects. A secondary task was completed by all the participants and ran for the entire duration of the study. Specific instructions were given onscreen before each block and a total of 7 practice trials were also completed (1 for each of the 4 blocks in the primary task plus three to familiarize with the secondary task alone). At the end of each block of 18 trials, participants took a short break. When all tasks had been completed each participant completed a short debriefing questionnaire and was provided with information about our experimental goals.

6 Results and Analysis

Primary Task A was the subject of our performance analysis. For each hypothesis we analyzed performances as a function of non-metaphoric vs. metaphoric visual representations and categories. Categories represented the 3 groupings, of 6 stimuli, within each experiment section and were named: Set1, Set2 and Set3. To analyze the patterns a 3 (sets) x 2 (metaphor vs. non-metaphor) repeated measure analysis of variance (ANOVA) was used to examine the accuracy and the response time data. Fig. 7 shows the overall performance of metaphoric visualizations vs. non-metaphoric visualization for all our participants. Fig. 6 summarizes performances as a function of non-metaphoric vs. metaphoric visual representations, Fig. 6c summarizes performances as a function of non-metaphoric vs. metaphoric visual representations and sets for the visual search hypothesis.

Hypothesis 1: Working Memory. For the accuracy data (see Fig. 6a) the ANOVA analysis showed a significant main effect of set ($F(2, 68) = 28.42$, $p=.0001$). This prompted us to perform further paired-sample t-tests to examine the source of the main effect. The t-test analysis revealed that all sets were significantly different from each other with Set1 yielding overall the highest accuracy, followed by Set2 and finally by Set3 (with all t-values$>$2.8 and $p<.005$). The ANOVA analysis also showed an overall main effect of metaphor vs. non-metaphor ($F(1,34)=23.10$, $p=.0001$) and significant interaction between sets and metaphor vs. non-metaphor ($F(2,68)=12.21$, $p=.001$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was no significant effect of metaphor vs. non-metaphor ($t(34)=1.43$, $p=.16$), (ii) within Set2 there was no significant effect of metaphor vs. non-metaphor ($t(34)=.30$, $p>.05$), (iii) within Set3 there was significant effect of metaphor vs. non-metaphor ($t(34)=6.91$, $p=.0001$).

For the response time data (see Fig. 6b) the ANOVA analysis showed a significant main effect of set ($F(2,68)=48.96$, $p=.0001$). Further paired-sample t-tests to examine the source of the main effect established that Set1 yielded overall the fastest response time relative to Set2 ($t(34)=-8.18$, $p=.0001$) and relative to Set3 ($t(34)=-9.11$, $p=.0001$), while there were no differences in response time between Set2 and Set3 ($t(34)=-1.45$,$p>.05$). The ANOVA analysis also showed an overall main effect of metaphor vs. non-metaphor ($F(1,34)=13.16$, $p=.001$) and significant interaction between sets and metaphor vs. non-metaphor ($F(2,68)=12.84$, $p=.0001$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was a significant effect of metaphor vs. non-metaphor ($t(34)=5.55$, $p=.0001$), (ii) within Set2 there was significant effect of metaphor vs. non-metaphor ($t(34)=-2.01$, $p=.04$), (iii) within Set3 there was significant effect of metaphor vs. non-metaphor ($t(34)=3.71$, $p=.0001$).

Hypothesis 2: Long-term Memory. For the accuracy data (Fig. 6a) the ANOVA analysis showed a significant main effect of set ($F(2,68)=9.39$, $p=.0001$). Further paired-sample t-tests confirmed Set1 as the source of the main effect. Accuracy in Set1 was significantly greater than accuracy in Set2 ($t(34)=3.93$, $p<.0001$) and than accuracy in Set3 ($t(34)=3.58$, $p<.001$). No difference in accuracy was detected between Set2 and Set3 ($t(34)<1$, $p>.05$). The ANOVA analysis showed no significant main effect of metaphor vs. non-metaphor ($F(1,34)=2.78$, $p=.10$) and no significant interaction between sets and metaphor vs. non-metaphor ($F(2,68)<1$, $p>.05$).

For the response time data (see Fig. 6b) the ANOVA analysis showed a significant main effect of set ($F(2,68)=25.32$, $p=.0001$). Further paired-sample t-tests to examine the source of the main effect established that Set1 yielded overall the fastest response time relative to both Set2 ($t(34)=-5.35$, $p=.0001$) and Set3 ($t(34)=-7.33$, $p=.0001$), while there were no differences in response time between Set2 and Set3 ($t(34)=-1.27$, $p>.05$). The ANOVA analysis also showed a significant main effect of metaphor vs. non-metaphor ($F(1,34)=16.84$, $p=.001$) with metaphoric visualizations yielding a shorter response time than non-metaphoric visualizations; and no significant interaction between sets and metaphor vs. non-metaphor ($F(2,68)=12.84$, $p=.0001$).

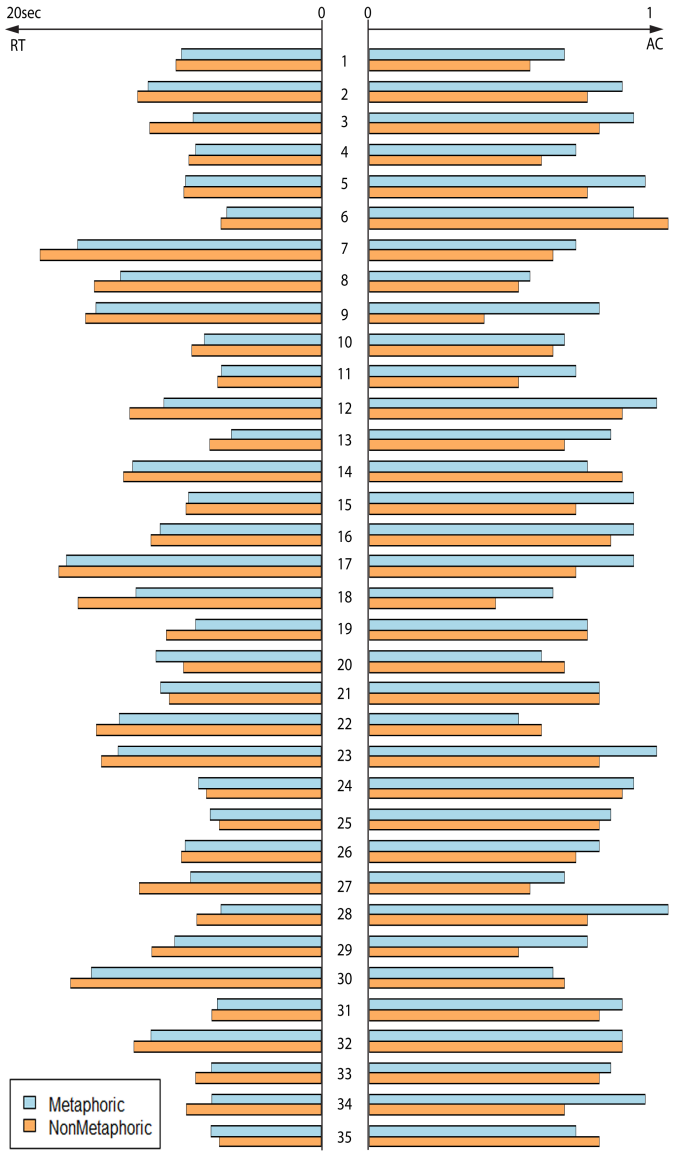


Fig. 7: Summary of Accuracy and Response Time for each participant.

Hypothesis 3: Visual Search. Fig. 6c summarizes performance as a function of non-metaphoric vs. metaphoric visual representations and sets for the visual search hypothesis.

For the response time data (Fig. 6b) the ANOVA analysis showed a significant main effect of set ($F(2,68)=70.31$, $p=.0001$). Further paired-sample t-tests to examine the source of the main effect established that Set2 yielded the fastest response time followed by Set1 and then Set3. All comparisons were significant (with all $t>4.00$ and all $p<.001$). The ANOVA analysis also showed a significant main effect of metaphor vs. non-metaphor ($F(1,34)=32.07$, $p=.0001$) with metaphoric visualizations yielding significantly slower response time than non-metaphoric visualizations; significant interaction between sets and metaphor vs. non-metaphor was found ($F(2,68)=9.41$, $p=.0001$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was no significant effect of metaphor vs. non-metaphor ($t(34)=1.23$, $p>.05$), (ii) within Set2 there was significant effect of metaphor vs. non-metaphor ($t(34)=3.58$, $p=.001$), (iii) within Set3 there was significant effect of metaphor vs. non-metaphor ($t(34)=5.04$, $p=.0001$).

Hypothesis 4: Concept Grasping. For the accuracy data (see Fig. 6a) the ANOVA analysis showed a significant main effect of set ($F(2,68)=5.04$, $p=.009$). This prompted us to perform further paired-sample t-tests to examine the source of the main effect. The t-test analysis revealed that there was significantly more accuracy in concept grasping within Set3 compared to Set1 ($t(34)=-2.93$, $p=.006$) and to Set2 ($t(34)=-2.01$, $p=.05$), while there was no difference between Set1 and Set2 ($t(34)=1.18$, $p>.05$). The ANOVA analysis also showed no significant main effect of metaphor vs. non-metaphor ($F(1,34)=2.46$, $p>.05$) and significant interaction between sets and metaphor vs. non-metaphor ($F(2,68)=4.20$, $p=.02$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was no significant effect of metaphor vs. non-metaphor ($t(34)=1.38$, $p=.17$), (ii) within Set2 there was significant effect of metaphor vs. non-metaphor ($t(34)=2.89$, $p=.007$), (iii) within Set3 there was no significant effect of metaphor vs. non-metaphor ($t(34)=1.05$, $p>.05$).

For the response time data (see Fig. 6b) the ANOVA analysis showed a significant main effect of set ($F(2,68)=11.20$, $p=.0001$). Further paired-sample t-tests to examine the source of the main effect revealed that there was a significantly higher response time in Set1 compared to both Set2 ($t(34)=3.93$, $p=.0001$) and Set3 ($t(34)=3.81$, $p=.001$), while there were differences in response time between Set2 and Set3 ($t(34)=1.24$, $p>.05$). The ANOVA analysis also showed no significant main effect of metaphor vs. non-metaphor ($F(1,34)=2.422$, $p>.05$) and no significant interaction between sets and metaphor vs. non-metaphor ($F(2,68)=1.23$, $p>.05$).

7 Findings and Discussion

The effect of metaphors on each of the fours tasks is discussed in turn.

Hypothesis 1: Working Memory. Visual metaphors aided performance in the working memory task, in terms of both accuracy and response time. This benefit was conditional on *grouping* between the metaphor and the to-be-remembered information. Accuracy benefits for metaphors were most prominent when the to-be-remembered information appeared grouped with the visual representation (e.g., within the bar or "bubble"). In particular in Set3, where both the labels (e.g., soft furnishings) and the visual metaphor (e.g., a picture of a curtain) appeared grouped with the "bubble" chart, the benefit for metaphors was significant relative to a lack of metaphors. In contrast, in stimulus Set2, where the metaphors appeared outside the graphical representation (bar) there was no difference between the metaphor and non-metaphor conditions. In terms of response time (RT), the presence of the metaphor had a positive effect relative to the non-metaphor conditions but only when the metaphor appeared grouped within the charts (as in stimulus Set1 and Set3). In contrast, when the metaphor appeared outside the graphical representation, it had a significant negative effect on RT (stimulus Set2).

The finding that grouping of visual metaphors with the to-be-remembered information is compatible with previous findings in psychology that visual working memory can be determined by the number of "objects" or "chunks" to be memorized [27, 22]. Grouping via proximity and closure (as is the case in Set3, and somewhat less so in Set1) is a powerful Gestalt cue to "objecthood". Such grouping of information was lacking in Set2, where the "objects" to be memorized were almost double in number (3 bars, 3 category labels, and 3 visual metaphors) than in Set3, where all the to-be-remembered information was grouped within a single object (e.g., a single "bubble"). The data suggest that it is when metaphors are grouped with the numerical representation, that they have the most beneficial influence on working memory tasks.

Hypothesis 2: Long-term Memory. Long-term memory was influenced both by the amount of the to-be-remembered information and by the presence of metaphors. Regardless of the presence of metaphors, the result of the long-term memory was best when only two data points (stimulus Set1) had to be remembered in contrast to three (stimulus Set2) or five (stimulus Set3) data points. There was a trend for a benefit of metaphoric compared to non-metaphoric visualizations, but this difference did not reach significance in the accuracy data. This trend was mirrored significantly in the response time data. Response time was faster when the task involved to remember visualizations containing two items to either 3 (stimulus Set1) or 5 items (stimulus Set3). More importantly, supporting our Hypothesis 2 that metaphors would facilitate long-term memory, the pres-

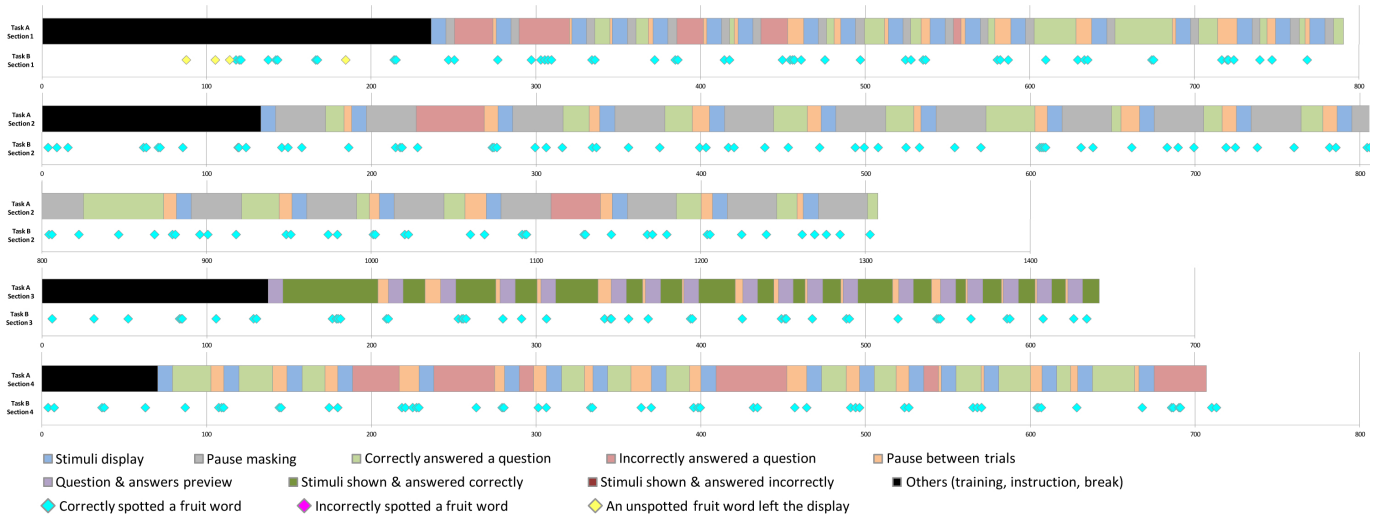


Fig. 8: Temporal correspondence between the primary and secondary tasks.

ence of metaphors facilitated response latencies, with significantly shorter response times for metaphoric compared to non-metaphoric visualizations. This was true for all three sets of stimuli. Given that in this task all the visual metaphors were grouped within the to-be-remembered information, the advantage of metaphoric visualizations over non-metaphoric ones, lends further support to the importance of grouping in enhancing the advantage of metaphoric visualizations.

Hypothesis 3: Visual Search. Two interesting results emerged in the visual search task. First, visual search response times were fastest in stimulus Set2 (bubble charts) followed by response times for stimulus Set1 (bars), and with the slowest response times for stimulus Set3 (bars). Critically, the difference in the pattern of results lies in the type of search participants had to engage in for each stimulus set. In Set1 and Set2 participants could have performed what is known as "guided search" [44]. In guided search, top-down information (e.g., the instruction to look for the oil consumption in India, Indonesia, and Pakistan amongst 5 other counties) can influence search times by rejecting any items that do not share any of the target characteristics (e.g., any countries other than India, Indonesia and Pakistan). Assuming that the distractors do not share any properties of the target (as was the case in stimulus Set1 and Set2), then search times are very fast. In Set3, the task is to search for the numbers in bars based on the semantic relationship of the bars. The names of farms do not affect the search in principle. For non-metaphoric visualization, participants use both colors and positions of the bars for their search. With metaphoric visualization, participants seem to have used the visual metaphors in addition to or instead of colors and positions, resulting in slower response time. In a way, the presence of visual metaphor has directed participants to a more complex feature for visual search. This is known as "conjunction search" [41], which typically take longer than single feature searches [44, 41]. The second, and most important finding is that metaphoric visualizations were significantly slower than non-metaphoric ones. Indeed, although in both memory tasks the extra information conveyed by the metaphors speeded response time, presumably by providing extra encoding and retrieval cues, the presence of metaphors significantly impaired performance by adding extra "distracting" visual information. The vividness of visual metaphors led viewers to use less effective cues (i.e., icons and pictograms) for visual search, in addition to or instead simple features such as colors and positions. In some cases, the cues such as national flags may not be familiar to the participants. Text labels were shown to have a small advantages than national flags in Set1 and Set2.

Hypothesis 4: Concept Grasping. The findings from the concept-grasping task confirmed the age-old saying of "a picture is worth a thousand words". In stimulus Set2, where the use of metaphors replaced a large amount of words in the visualization the difference was highly significant. However, where metaphors were potentially ambiguous or unfamiliar, as in some stimulus in Set1 visualizations, the presence of metaphors adds no benefit. The pattern of accuracy in the concept-grasping task suggests that metaphors can increase the accuracy of concept grasping when they are familiar and can convey information with less amount of visual information. The performance on Set3 suggests that participants might unconsciously take in information from visual metaphors. This usually works more effective when there is less visual metaphors. The metaphoric stimuli in Set3 have only one visual metaphor, and these in Set1 have smaller number of visual metaphors than those in Set2. The benefits of visual metaphor appear to decrease when the number of visual metaphor increases.

8 Conclusion

In this work we conducted a thorough user study inspecting the fundamental aspects of the data analysis process when visual metaphors are involved in the representation of the data. Results provided a quantitative analysis of both task difficulty and representational variations when using such a technique. Our results are most relevant to the users and developers of visualizations involving metaphorical visual representation of data, particularly those working with visualization for the masses and time series analysis. Visual metaphors are a powerful communication tool and are widely used. We believe this work represents a significant step towards an understanding of the relationship between the two dimensions of visual complexity and task requirements in metaphor based visualization. Based on our results, we can conclude that information retention is improved by the use of visual metaphors at the expenses however of an increase in processing time. Memory consolidation is indeed a complex phenomena and long-term potentiation relies heavily on the amount of repetitions, our results have shown how visual metaphors have a significant, and positive, impact on the speed of memory recalling.

The perceptual load associated with an increase in visual details and features impacts upon the performance of the users when target search is the primary task. This suggests that aspects like familiarity and context are not the only factors involved: metaphor usability is indeed directly linked to user experience however metaphor cognitive characteristics remain the dominant factor in user performances.

Based on our result it clearly emerges that the design process should pay particular attention to the effects of cognitive metaphor characteristics on user performance. We also hope to further explore our findings as to how the can alleviate tasks like visual search in more complex practical situations. A final challenging direction to consider would be the development of a systematic way of ensuring that visualization designs make optimal use of the power of expression of metaphors.

REFERENCES

[1] E. Aley. *Metaphorical Visualization: Using Multidimensional Metaphors to Visualize Data*. VDM Verlag, 2007.

[2] V. L. Averbukh. Visualization metaphors. *Program. Comput. Softw.*, 27(5):227–237, Sep. 2001.

[3] A. D. Baddeley. Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal Of Experimental Psychology*, 18(4):362–365, 1966.

[4] M. Black. *Models and metaphors: Studies in language and philosophy*. G - Reference, Information and Interdisciplinary Subjects Series. Cornell University Press, 1962.

[5] A. Blackwell. *Visual Representation*. The Interaction-Design.org Foundation, Aarhus, Denmark, 2011.

[6] A. F. Blackwell and T. R. G. Green. Does metaphor increase visual language usability? In *Proceedings of the IEEE Symposium on Visual Languages*, VL '99, pages 246–254. IEEE Computer Society, 1999.

[7] L. Cameron and R. Maslen. *Metaphor analysis: Research practice in applied linguistics, social sciences and the humanities*. Studies in applied linguistics. Equinox Pub., 2010.

[8] C. Cazeaux. *Metaphor and continental philosophy: From Kant to Derrida*. Routledge studies in twentieth century philosophy. Routledge, 2007.

[9] W. D. Chiles. *Workload, task and situational factors as modifiers of complex human performance*, pages 11–52. Lawrence Erlbaum Associates, 1992.

[10] C. Forceville and E. Urios-Aparisi. *Multimodal metaphor*. Applications of cognitive linguistics. Mouton de Gruyter, 2009.

[11] D. Gentner, K. Holyoak, and B. Kokinov. *The Analogical Mind: Perspectives from Cognitive Science*. Bradford Books. Mit Press, 2001.

[12] R. Gibbs. *The Cambridge Handbook of Metaphor and Thought*. Cambridge Handbooks in Psychology. Cambridge University Press, 2008.

[13] M. Glanzer and A. R. Cunitz. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5(4):351–360, 1966.

[14] P. Hanks and R. Giora. *Metaphor and Figurative Language*. Critical Concepts in Linguistics. Taylor & Francis, 2011.

[15] C. Healey. Choosing effective colours for data visualization. In *Proc. IEEE Visualization*, pages 263–270, 1996.

[16] D. A. Henderson Jr. and S. K. Card. Rooms: The use of multiple workspaces to reduce space contention in a window-based graphical user interface. *ACM Transactions on Graphics*, 5(3):211–243, 1986.

[17] W. H. Huggins and D. R. Entwisle. *Iconic communication; An annotated bibliography*. Johns Hopkins University Press Baltimore,, 1974.

[18] W. B. Knowles. Operator loading tasks. *The Journal of the Human Factors and Ergonomics Society*, 5:155–161, Apr. 1963.

[19] Z. Kovecses. *Metaphors of anger, pride, and love: A lexical approach to the structure of concepts*. J. Benjamins Pub. Co., Amsterdam, Philadelphia, 1986.

[20] G. Lakoff. *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press, Chicago, 1987.

[21] G. Lakoff. *The contemporary theory of metaphor*, volume 2, pages 202–251. Cambridge University Press, 1993.

[22] S. J. Luck and E. K. Vogel. The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657):279–281, 1997.

[23] A. Mack, Z. Pappas, M. Silverman, and R. Gay. What we see: Inattention and the capture of attention by meaning. *Consciousness and Cognition*, 11(4):488–506, Dec. 2002.

[24] R. Mander, G. Solomon, and Y. Y. Wong. A "pile" metaphor for supporting casual organization of information. In *Proc. CHI'92*, pages 627–634. ACM Press, Monterey, 1992.

[25] T. McArthur. *The Oxford companion to the English language*. Oxford University Press, 1992.

[26] S. McDougall and M. Curry. More than just a picture: Icon interpretation in context. In *Coping with Complexity Workshop*, pages 16–17, 2004.

[27] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 101(2):343–52, 1956.

[28] C. J. Morris, D. S. Ebert, and P. Rheingans. *An Experimental Analysis of the Pre-Attentiveness of Features in Chernoff Faces*, pages 12–17. SPIE, 1999.

[29] G. L. Murphy. On metaphoric representation. *Cognition*, 60:173–204, 1996.

[30] A. Ortony. Why metaphors are necessary and not just nice. *Education Theory*, 25:45–53, 1975.

[31] A. Ortony, editor. *Metaphor and Thought*. Cambridge University Press, 2nd edition, 1993.

[32] A. Pang and M. Clifton. Metaphors for visualization. In *Proc. Eurographics Workshop on Visualization in Scientific Computing*, 1995.

[33] H. E. Pashler. *Attention*. Psychology Press, 1998.

[34] C. Peirce. *The New Elements of Mathematics*. 1902.

[35] L. R. Peterson and M. J. Peterson. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58(3):193–198, 1959.

[36] A. Quinn. *Figures of speech: 60 ways to turn a phrase*. Hermagoras Press, 1995.

[37] I. Reppa, D. Playfoot, and S. J. P. McDougall. Visual aesthetic appeal speeds processing of complex but not simple icons. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, pages 1155–1159, 2008.

[38] M. S. Sanders and E. J. McCormick. *Human Factors in Engineering and Design*. McGraw-Hill Book Company, 6th edition, 1987.

[39] H. S. Smallman, M. S. John, H. M. Oonk, and M. B. Cowen. Information availability in 2D and 3D displays. *IEEE Computer Graphics and Applications*, 21(5):51–57, 2001.

[40] T. G. Sticht. Educational uses of metaphor. In A. Ortony, editor, *Metaphor and Thought*, pages 621–632. Cambridge University Press, 1993.

[41] A. Treisman. *Properties, parts, and objects*, volume II, pages 35:1–35:70. Wiley, 1986.

[42] E. Tufte. *The visual display of quantitative information*. Graphics Press, 2001.

[43] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers, second edition, 2004.

[44] J. M. Wolfe. Guided search 2.0 - a revised model of visual-search. *Psychonomic Bulletin & Review*, 1(2):202–238, 1994.

[45] M. K. Yousef. Assessment of metaphor efficacy in user interfaces for the elderly: A tentative model for enhancing accessibility. In *Proceedings of the 2001 EC/NSF Workshop on Universal Accessibility of Ubiquitous Computing: Providing for the Elderly*, WUAUC'01, pages 120–124. ACM, 2001.

[46] C. Ziemkiewicz and R. Kosara. The shaping of information by visual metaphors. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1269–1276, 2008.

[47] C. Ziemkiewicz and R. Kosara. Preconceptions and individual differences in understanding visual metaphors. *Computer Graphics Forum*, 28(3):911–918, 2009.