

An Empirical Study on Using Visual Embellishments in Visualization

Rita Borgo, Alfie Abdul-Rahman, Farhan Mohamed, Philip W. Grant, Irene Reppa, Luciano Floridi, and Min Chen

Abstract—In written and spoken communications, figures of speech (e.g., metaphors and synecdoche) are often used as an aid to help convey abstract or less tangible concepts. However, the benefits of using rhetorical illustrations or embellishments in visualization have so far been inconclusive. In this work, we report an empirical study to evaluate hypotheses that visual embellishments may aid memorization, visual search and concept comprehension. One major departure from 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 lectures). The secondary task introduces “divided attention”, and makes the effects of visual embellishments more observable. In addition, it also serves as additional masking in memory-based trials. The results of this study show that visual embellishments can help participants better remember the information depicted in visualization. On the other hand, visual embellishments can have a negative impact on the speed of visual search. The results show a complex pattern as to the benefits of visual embellishments in helping participants grasp key concepts from visualization.

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

1 INTRODUCTION

In written and spoken communications, figures of speech (e.g., metaphor, synecdoche) may aid communication and thought processes through *compactness*, *vividness* and *inexpressibility* [24]. Visual embellishments are a form of non-linguistic rhetorical figures that can be seen frequently in the visual arts, performing arts, advertisements, icons and signs, culture symbols, color symbolism, graphical user interfaces, and so forth. Similar to figures of speech, 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 [33]. Naturally, one cannot help but wonder whether these three features of visual embellishments 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.

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 [11], 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 [19, 12], and use of “tree” and “container” metaphors in treemaps [39]. On the other hand, there have been overwhelming criticisms about inaccurate depiction of visual metaphors (e.g., [35]). There have been empirical studies showing that both 2D and 3D visual embellishments do not necessarily bring the desired advantages to information visualization (e.g., [6, 32]). In this work, we re-examine the question about the effects of visual embellishments in visualization. The investigation is partly motivated by a new observation about typical situations where visualization often features visual embellishments, such as visualiza-

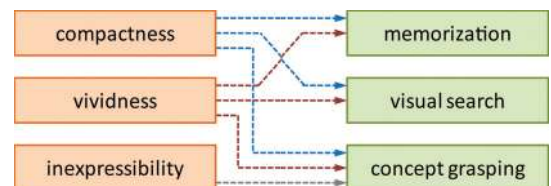


Fig. 1. Hypothesized transition from features of visual embellishments to effects on visualization.

tions 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 divide their attention between visualization and other tasks (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 embellishments 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 [26, 8]. We found that the *subsidiary task paradigm* [15] 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 embellishments. The secondary task was designed to introduce “divided attention” in order to reduce slightly the participants’ cognitive capacity available for the primary task [26]. It is necessary to emphasize that the focus of this work was on the effects of visual embellishments rather than attention or workload assessment. The dual task methodology was adopted to make the effects of visual embellishments more apparent.

As illustrated in Fig. 1, we hypothesized that the three features of rhetorical figures 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 embellishments may help participants remember better in visualization (in terms of working memory).
- H2. Visual embellishments may help participants remember better in visualization (in terms of long-term memory).
- H3. Visual embellishments may help participants perform visual search more efficiently during visualization.
- H4. Visual embellishments 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.grant}@swansea.ac.uk.
- Alfie Abdul-Rahman and Min Chen are with Oxford e-Research Centre, University of Oxford. E-mail: {alfie.abdulrahman,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.

Manuscript received 31 March 2012; accepted 1 August 2012; posted online 14 October 2012; mailed on 5 October 2012.

For information on obtaining reprints of this article, please send e-mail to: ivcg@computer.org.

2 RELATED WORK

Terminology. In linguistics, there are many forms of figures of speech, ranging from *antithesis* to *irony*. Among these forms are *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 use of the 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 hyperbole suggest some form of exaggeration.

There is no similar detailed categorization for visual communication where some linguistic subtleties would not normally be encoded in visual representations. A possible encompassing term could be “figures of visualization”, but it would be very confusing. An alternative term is “visual metaphor” which is commonly defined to as a visual image that is used to convey a primary concept in visualization by making use of a secondary concept. A number of approaches have been proposed to analyze visual metaphors as conceptual phenomena; Conceptual Metaphor Theory (CMT) [16] proposes that verbal metaphors are merely surface manifestations of metaphorical thought and that “[a] metaphor is fundamentally conceptual, not linguistic in nature”. It is highly attractive to adopt this term in order to encapsulate all such associations between the two concepts with the goal being of assisting the expression of a primary concept. A drawback is that the word “metaphor” has a narrower implication in linguistics. Confusion could arise from its use as a subcategory of rhetorical figures of speech in written and spoken communication, and its use as an encompassing term for all rhetorical figures in visual communication. Nevertheless, we used the term *visual metaphor* throughout the design and experimental stages of this study in 2010 and 2011 until we were made aware of the phrase “visual embellishment” used by Bateman *et al.* to describe *chart junk* that is not essential to the understanding of the data depicted in visualization [4]. For this paper, we make use of *embellishment* as the encompassing term, while maintaining the term of *visual metaphor* as an encompassing term in places (e.g., stimuli set and results charts) where we need to provide a faithful record of the study. Similar to figures of speech, we refer the primary concept as the *tenor*, and the secondary as the *vehicle* (or *target* and *source* in some literature). There are sub-categories of visual embellishments. For example, a vehicle of a light bulb may be used to imply a tensor of electricity (cf. metaphor). A picture of a sheep may be used to represent a drove of sheep (cf. synecdoche). A skull may be used to suggest risk even if it is not death-threatening (cf. hyperbole). A national flag may be used to substitute a country name (cf. metonymy).

Visual Embellishments in Visualization. The uses and benefits of visual embellishments have always been a debatable subject in visualization. Some have advocated strongly the removal of visual embellishments in visualization (e.g., [35]). In the literature, some academic researchers have ventured into the territory of visual embellishments. Pang and Clifton proposed the use everyday objects to create intuitive 3D interfaces [25]. 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 [39]. Healey studied colour symbolism [11], and Huggins and Entwisle on iconic communication [13]. Aley wrote a short book on metaphorical visualization [1], which presents many applications where “vivid” metaphors are used in visualization. Bateman *et al.* reported an empirical study showing that visual embellishments could improve long-term recall without noticeable impact on comprehension [4]. Hullman *et al.* articulated the benefits of visual embellishments by building on existing research findings [14].

Ware [36] 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 visual design [35], and it was studied in the context of visualization by [22]. McDougall *et al.* [20] and

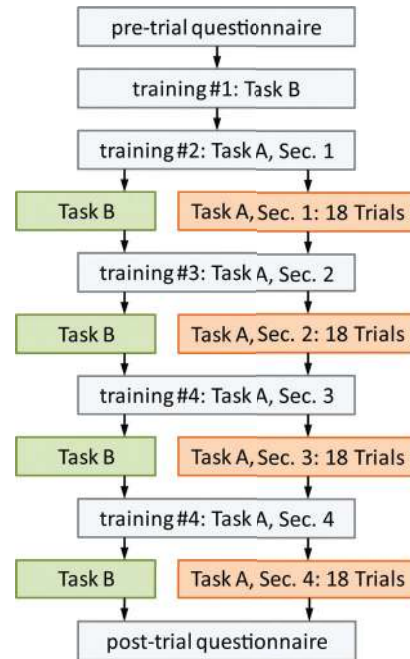


Fig. 2. Activity sequence of the study.

Blackwell [5] provide detail guidelines on icon design, while Reppa *et al.* [29] analyzed the impact of icon design on performances of both familiarity and aesthetic appeal. Familiarity is a concept probed by [40] which is related to metaphor usability [38].

Our work is intended to study the impact of visual embellishments on the user performance in visualization, building on the theory of conceptual structure by [23]. The main differences between our study and that of [4] are: (i) we make use of 36 pairs of stimuli (cf. 14 pairs in [4]); (ii) we use “milder” and less artistic visual embellishments than [4]; (iii) we study impact on working memory, long-term memory, visual search and concept grasping (cf. long-term and comprehension in [4]); and (iv) we design a dual-task study to reflect typical situations where visual embellishments may be used in visualization.

We draw connections with the literature in perception and cognition in individual sections where the previous works are the most relevant.

3 EXPERIMENT OVERVIEW

This study has three control variables. The most important variable is *embellished visualization* vs. plain visualization. This variable is controlled by having equal number of stimuli in each category, and pair them in the design stage to ensure a similar visual representation (except embellishments), and a similar amount of information, complexity and cognitive load. In the experimental stage, stimuli display order is randomized.

The second variable is the four hypotheses. This variable is controlled by the four separate sections of the study. Each section is composed of 18 stimuli, all designed for evaluating the same hypothesis. In a pilot study, we found noticeable variations in performance of tasks in different sections, which likely reflect the different levels of perceptual and cognitive loads. We thus decided not to mix stimuli from different sections. All tests followed the same order of sections for hypotheses H1, H2, H3 and H4.

The third variable is the variation of visual designs. This is controlled by the three sets in each section. Each set has 6 stimuli, organized into three pairs of embellished and plain visualizations (i.e., the first variable). It is necessary to ensure that there is an adequate number of stimuli for the each visual design. This alleviates the confounding effect due to unexpected anomalies in an individual design.

In order to provide our study with an intuitive scenario that all 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

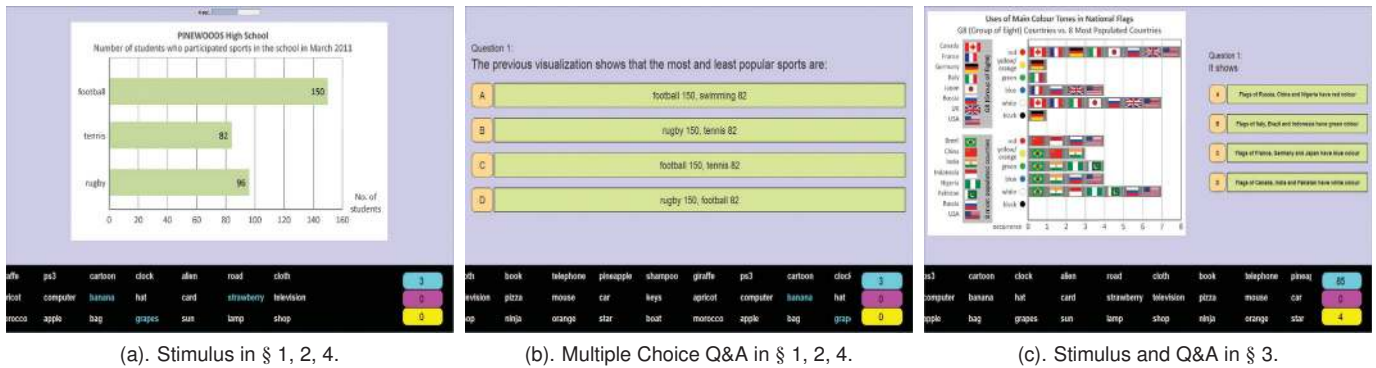


Fig. 3. Screenshots from the experiment interface: a)-b) example of the layout for stimulus and multiple choice question for Sections 1, 2 and 4; c) example of the layout for stimulus and multiple choice question for Section 3.

stimuli were selected from visual representations common in everyday life. To increase stimuli reliability, the chosen visual embellishments underwent scrutiny by a small but ethnically diverse group of people. In our experiment we maintained a one-to-one correspondence between our embellished and plain stimuli, e.g., every plain stimulus had a corresponding (and equivalent) embellished 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 [20], 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 be 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.

Visual Complexity. Complexity is defined as the amount of detail in the visual representation. Complexity is a direct function of (i) the degree of perceivable structure, (ii) variety of parts and (iii) 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 [29].

Meaningfulness. Meaningfulness indicates the relationship between what is depicted in the visual representation and the function it refers. Research suggests that particularly meaningful icons in displays can capture human attention and drive visual search [18]. If the meaning can be extracted at the same time as 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 [27].

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 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.

These factors are not control variables as it is not appropriate to have too many variables in such a study. Nevertheless, it is necessary to consider them in the design to minimize the confounding effects due to the variations in familiarity, meaningfulness, complexity or other above-mentioned factors.

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 embellished and plain data representation. 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 visualization. Task B, running throughout the entire duration of the study, was also performed. This concurrent task acted as a distracting factor and helped us mimic the effects of divided attention which are unavoidable when performing every day tasks.

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 men-

tioned 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 embellishments 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 in order to avoid the confounding effect that may have 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 (referred to as Set1, Set2 and Set3) of 6 stimuli each (3 with visual embellishments and 3 without). For sections 1-3, stimuli in each group (or set) 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, with one member of the pair featuring visual embellishments and the other member featuring a plain representation of the data.

Some examples of the stimuli are shown in Fig. 4, where four columns correspond to 4 sections respectively (referred to as WM, LM, VS and CG in short). Each column shows three pairs of example stimuli, chosen from the three different sets in the corresponding section. For example, Fig. 4a and Fig. 4b are the first two stimuli (as a pair) in Set 1 of Section 1 (WM). Fig. 4w and Fig. 4x are the last two stimuli (as a pair) in Set 3 of Section 4 (CG).

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 embellishments), 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 embellishments, 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 embellishments have exactly 50% chance to be shown before the corresponding plain stimulus, and vice versa. The scheme also ensures that two stimuli in the same pair are separated 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, some emulate real world concepts and data distribution, others are obtained from public domain sources (e.g., Wikipedia). Questions and optional answers were carefully designed to ensure that a correct answer could not be easily inferred 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 embellishments 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 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 [28]. In working memory written text competes with visually encoded information for storage. If multiple data attributes are integrated into a 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 embellishments. We use clipart pictures, icons and photographic images for visual embellishments, all of which are additional to the text labeling. In bubble charts, the data values correspond to the areas of circles, and all embellishments are contained within the circles.

In each trial, a stimulus is first displayed for 9 seconds (Fig. 3a). 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 previous stimulus is presented with four optional answers (in the format of Fig. 3b). For example, the question and multiple choices for stimulus WM1.1 (Fig. 4a) 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 embellishments 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 [28], and there are also suggestions that it can last for up to 30 seconds [10]. 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 embellishments 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.

Before each stimulus is shown, the corresponding question and optional answers are presented to the participant for 9 seconds. For example, for stimulus VS 1.5 (Fig. 4o), 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 (Fig. 3c). The participant is then allowed to answer the question. Each trial requires the participants to perform multiple visual searches, which removes some of knowledge bias e.g., 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 embellishments 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* places 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 into 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



Fig. 5. User study secondary task.

- 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 focuses 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”.

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 [30]:

- **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

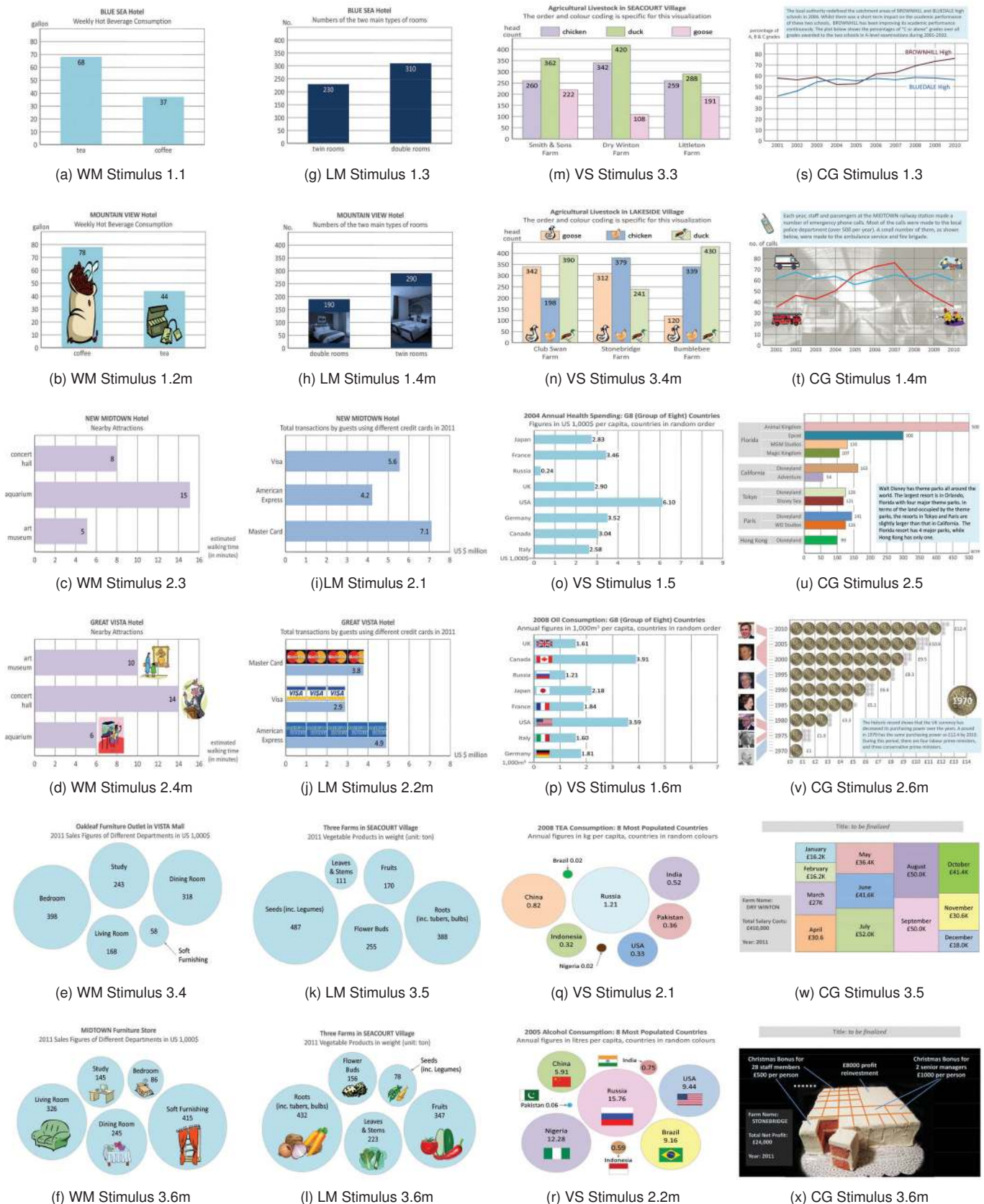


Fig. 4. Examples of stimuli used in the study, where those tagged with a letter "m" are stimuli with visual embellishments.

right hand corner of the screen kept the count of how many fruit words had been correctly selected, missed and how many words had been wrongly selected. 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 written in Java. Stimuli were saved as static images and presented to participants using a custom made interface. Experiments were run using an Intel Dual-Core PCs, 2.13 GHz, 2 GB of RAM and Windows 7 Professional. The display was 19" LCD at 1280x1024 resolution and 32bit sRGB color mode. Each monitor was adjusted to the same brightness and 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 sections of 18 trials. The 4 sections were always completed in sequential order. Given the nature of the experiment each section assessed a different aspect of the cognitive process. Maintaining the same section 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 section, trials were randomized to avoid learning effects. A secondary task was completed by all participants and ran for the entire duration of the study. Specific instructions were given onscreen before each section and a total of 7 practice trials were also completed (1 for each of the 4 sections in the primary task plus three to familiarize with the secondary task alone). At the end of each section of 18 trials, participants took a short break. When all tasks had been completed each participant completed a short debriefing questionnaire.

6 RESULTS AND ANALYSIS

Primary Task A was the subject of our performance analysis. For each hypothesis we analyzed performances as a function of plain vs. embellished 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) \times 2 (embellished vs. plain) repeated measure analysis of variance (ANOVA) was used to examine the accuracy and the response time data. Fig. 6a-b summarize performances as a function of plain vs. embellished visual representations, Fig. 6c summarizes performances as a function of plain vs. embellished 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 the session ($F(2, 68) = 28.42, p=.0001$). 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 embellished vs. plain ($F(1,34)=23.10, p=.0001$) and significant interaction between sets and embellished vs. plain ($F(2,68)=12.21, p=.001$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was no significant effect of embellished vs. plain

($t(34)=1.43, p=.16$), (ii) within Set2 there was no significant effect of embellished vs. plain ($t(34)=.30, p>.05$), (iii) within Set3 there was significant effect of embellished vs. plain ($t(34)=6.91, p=.0001$).

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 embellished vs. plain ($F(1,34)=13.16, p=.001$) and significant interaction between sets and embellished vs. plain ($F(2,68)=12.84, p=.0001$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was a significant effect of embellished vs. plain ($t(34)=5.55, p=.0001$), (ii) within Set2 there was significant effect of embellished vs. plain ($t(34)=-2.01, p=.04$), (iii) within Set3 there was significant effect of embellished vs. plain ($t(34)=3.71, p=.0001$).

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 embellished vs. plain ($F(1,34)=2.78, p=.10$) and no significant interaction between sets and embellished vs. plain ($F(2,68)<1, p>.05$).

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 embellished vs. plain ($F(1,34)=16.84, p=.001$) with embellished visualizations yielding a shorter response time than plain visualizations; and no significant interaction between sets and embellished vs. plain ($F(2,68)=12.84, p=.0001$).

Hypothesis 3: Visual Search. Fig. 6c summarizes performance as a function of plain vs. embellished 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 embellished vs. plain ($F(1,34)=32.07, p=.0001$) with embellished visualizations yielding significantly slower response time than plain visualizations; significant interaction between sets and embellished vs. plain 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 embellished vs. plain ($t(34)=1.23, p>.05$), (ii) within Set2 there was significant effect of embellished vs. plain ($t(34)=3.58, p=.001$), (iii) within Set3 there was significant effect of embellished vs. plain ($t(34)=5.04, p=.0001$).

Hypothesis 4: Concept Grasping. For the accuracy data (see Fig. 6a) the ANOVA analysis showed a significant main effect of set ($F(2,68)=5.04, p=.009$). Further paired-sample t-tests 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 embellished vs. plain ($F(1,34)=2.46, p>.05$) and significant interaction between sets and embellished vs. plain ($F(2,68)=4.20, p=.02$). Further paired-sample t-tests to examine the interaction showed that: (i) within Set1 there was no significant effect of embellished vs. plain ($t(34)=1.38, p=.17$), (ii) within Set2 there was

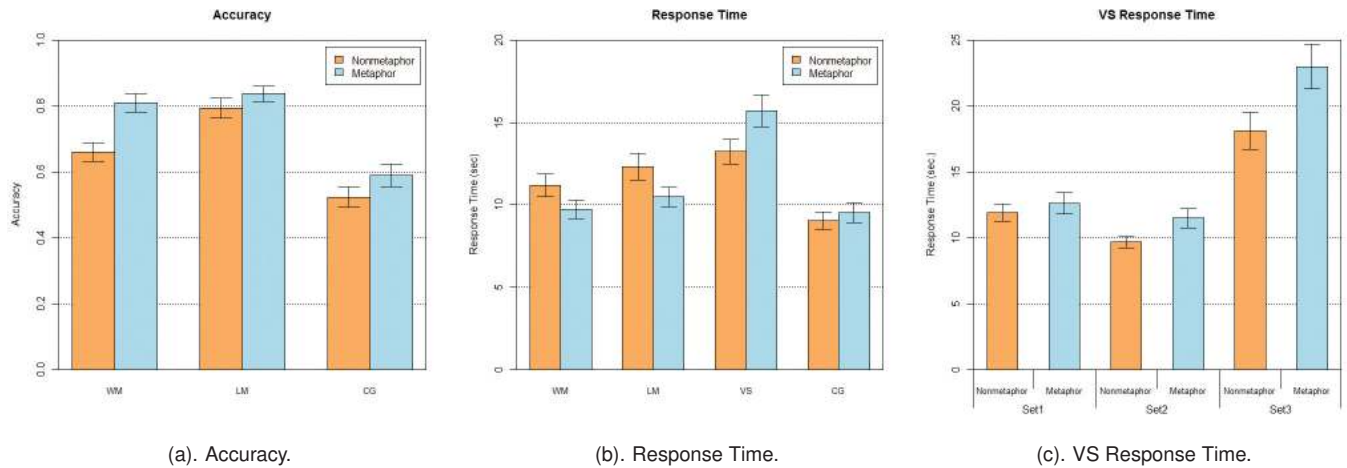


Fig. 6. Performance analysis for all four hypotheses: a) accuracy for Working Memory (WM), Long-term Memory (LM) and Concept Grasping (CG) hypotheses; b) response time for Working Memory (WM), Long-term Memory (LM), Visual Search (VS) and Concept Grasping (CG) hypotheses; c) example of set based analysis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

significant effect of embellished vs. plain ($t(34)=2.89$, $p=.007$), (iii) within Set3 there was no significant effect of embellished vs. plain ($t(34)=1.05$, $p>.05$).

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 embellished vs. plain ($F(1,34)=2.422$, $p>.05$) and no significant interaction between sets and embellished vs. plain ($F(2,68)=1.23$, $p>.05$).

7 FINDINGS AND DISCUSSION

In general, the effects of the secondary task upon the primary task are as much as expected. As shown in Fig. 7, participants did perform a fair amount of the secondary task during the masking phases in the two memory sections and the Q&A preview phase in Section 3 (visual search). Hence the secondary task has fulfilled the additional masking role, while the Q&A preview phrase did alleviate the potential bias in terms of reading speed. Other phases where participants performed a fair amount of the secondary task were during stimuli display, between trials, and during training, viewing instructions and breaks (labeled as others). In general, the performance of the secondary task were very good. On average, 174.8 fruit words were correctly identified, whereas 22.7 were missed. There was a very small number of errors, as on average 1.4 non-fruit words were selected by mistake. Given that the secondary task did fulfill its function, we can focus on the primary task designed to evaluate the four hypotheses, which are discussed in turn.

Hypothesis 1: Working Memory. Visual embellishments aided performance in the working memory task, in terms of both accuracy and response time. This benefit was conditional on *grouping* between the embellishments and the to-be-remembered information. Accuracy benefits for embellishments 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 embellishments (e.g., a picture of a curtain) appeared grouped with the “bubble” chart, the benefit for embellishments was significant relative to a lack of embellishments. In contrast, in stimulus Set2, where the embellishments appeared outside the graphical representation (bar) there was no difference between the embellished and plain conditions. In terms of response time (RT), the presence of the embellishments had a positive effect relative to the plain conditions but only when the embellishments appeared grouped within the charts (as in stimulus Set1 and Set3). In

contrast, when the embellishments appeared outside the graphical representation, it had a significant negative effect on RT (stimulus Set2).

The finding that grouping of visual embellishments 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 [21, 17]. 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 embellishments) 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 embellishments 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 embellishments. Regardless of the presence of embellishments, 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 embellished compared to plain visualizations, but this difference did not reach significance in the accuracy results. This trend was mirrored significantly in the response time. Response time was faster when the task involved remembering visualizations containing 2 items in comparison with either 3 (stimulus Set2) or 5 items (stimulus Set3). More importantly, supporting our Hypothesis 2 that embellishments would facilitate long-term memory, the presence of embellishments facilitated response latencies, with significantly shorter response times for embellished compared to plain visualizations. This was true for all three sets of stimuli. Given that in this task all visual embellishments were grouped within the to-be-remembered information, the advantage of embellished visualizations over plain ones, lends further support to the importance of grouping in enhancing the advantage of embellished 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” [37]. In guided search, top-down information (e.g., the instruction to look for the oil consumption in India, Indonesia, and Pakistan amongst 5 other countries) 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 plain visualization, participants use both colors and positions of the bars for their search. With embellished visualization, participants seem to have used the visual embellishments in addition to or instead of colors and positions, resulting in slower response time. In a way, the presence of visual embellishments has directed participants to a more complex feature for visual search. This is known as “conjunction search” [34], which typically takes longer than single feature searches [37, 34]. The second, and most important finding is that embellished visualizations were significantly slower than plain ones. Indeed, although in both memory tasks the extra information conveyed by the embellishments speeded response time, presumably by providing extra encoding and retrieval cues, the presence of embellishments significantly impaired performance by adding extra “distracting” visual information. The vividness of visual embellishments led viewers to use less effective cues (i.e., icons and pictograms) for visual search, in addition to or instead of 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 advantage over national flags in Set1 and Set2.

Hypothesis 4: Concept Grasping. The findings from the concept-grasping task confirmed the age-old saying “a picture is worth a thousand words”. The experimental results show that visual embellishments bring more benefits with stimulus Set2 than Set1 and Set3. We can observe easily that visualizations in Set2 are in general more complex than those in Set1 and Set3. It is likely that participants did not have enough time to read all the texts in those stimuli in Set2, and relied on visual embellishments to gain an impression of the key concepts when they were available. For Set1 and Set3, there is no significant effect of embellished vs. plain visualizations. This finding is similar to that of Bateman *et al.* [4].

Findings about concept grasping are mixed, with positive confirmation of hypothesis H4 by Set2 stimuli, but not Set1 and Set3 stimuli. They suggest that at least there was no negative impact. Using appropriate visual embellishments in presenting complex information may help viewers grasp key concepts more effectively. This effect may also be indirectly caused by the impact of visual embellishments upon memory, which has been confirmed by the results of Set1 and Set2.

Discussions on Minimalistic Designs. One should not naïvely conclude that the findings of this study contradict the minimalistic design principles promoted by Tufte [35], Few [9] and many others. First, the finding about the negative impact on visual search tasks provides scientific evidence to indicate some disadvantages of using visual embellishments. Second, it is important to note that all visual tasks in this study were performed with the condition that stimuli were designed by authors and viewed by participants. Hence the findings about memory and concept grasping should not be generalized to situations where visualizations are created by data analysts for their own use.

From an information theoretical perspective [7], visual communication between a visualization creator and a group of viewers is “noisier” and less reliable than that between a data analyst and himself/herself. In many ways, the minimalistic design principle echoes Shannon’s source coding theorem [31] that places emphasis on efficient use of the communication bandwidth. In general, there is much less need for a data analyst to help himself/herself to grasp the key concepts or to remember a visualization by using embellishments, though data analysts can benefit from other forms of redundancy in visualization (e.g., lines that join dots in time series, and domain-specific metaphors).

On the other hand, the use of embellishments in visualization echoes Shannon’s noisy-channel coding theorem [31] that places emphasis on making use of redundancy to facilitate automatic error detection and correction at the receiver’s end. Visual embellishments can thus be viewed as redundancy, which strengthens “signals” in a visualization, hence helping memorization and concept grasping.

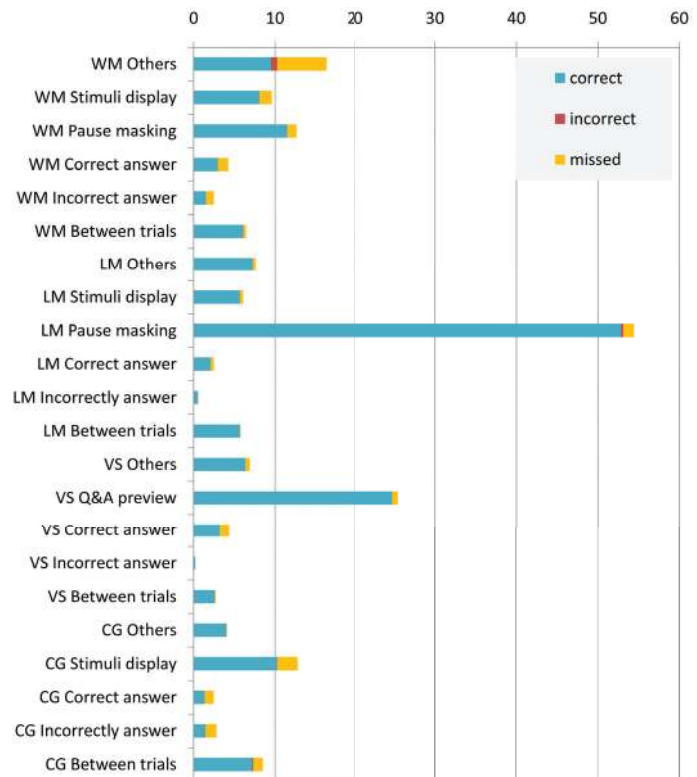


Fig. 7. Statistical correspondence between the primary and secondary tasks in different phases of the experiment. Blue indicates the average number of fruit words correctly identified, yellow for unidentified fruit words, and red for non-fruit words mistakenly chosen.

8 CONCLUSION

In this work we conducted a user study to examine the effects of visual embellishments in visualization processes in relation to several fundamental aspects of perception and cognition like: working memory, long-term memory, visual search and concept-grasping. In addition, the study provided the basis for a quantitative analysis of the effects in relation to task difficulty and representational variations. Our results are most relevant to the users and developers of visualizations involving embellished visual representation of data, particularly those working with visualization for the masses and time series analysis. Visual embellishments 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 embellished visualization. Based on our results, we can conclude that information retention is improved by the use of visual embellishments 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 embellishments 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: embellishments usability is indeed directly linked to user experience however embellishments 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 characteristics of visual embellishments on user performance. We also hope to further explore our findings as to how the visual embellishments 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 embellishments, such as metaphors and symbolisms, to make connections between visual and mental representations.

REFERENCES

- [1] E. Aley. *Metaphorical Visualization: Using Multidimensional Metaphors to Visualize Data*. VDM Verlag, 2007.
- [2] V. L. Averbukh. Visualization metaphors. *Programming and Computer Software*, 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] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proc. International Conference on Human Factors in Computing Systems (CHI 10)*, pages 2573–2582. ACM, 2010.
- [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 *Proc. IEEE Symposium on Visual Languages (VL 99)*, pages 246–254, 1999.
- [7] M. Chen and H. Jänicke. An information-theoretic framework for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1206–1215, 2010.
- [8] W. D. Chiles. Workload, task and situational factors as modifiers of complex human performance. In E. A. Alluisi and E. A. Fleishman, editors, *Human Performance and Productivity*, pages 11–52. Lawrence Erlbaum Associates, 1982.
- [9] S. Few. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, 2009.
- [10] M. Glanzner and A. R. Cunitz. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5(4):351–360, 1966.
- [11] C. Healey. Choosing effective colours for data visualization. In *Proc. IEEE Visualization*, pages 263–270, 1996.
- [12] 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.
- [13] W. H. Huggins and D. R. Entwisle. *Iconic Communication; An Annotated Bibliography*. Johns Hopkins University Press Baltimore, 1974.
- [14] J. Hullman, E. Adar, and P. Shah. Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2213–2222, 2011.
- [15] W. B. Knowles. Operator loading tasks. *The Journal of the Human Factors and Ergonomics Society*, 5:155–161, Apr. 1963.
- [16] G. Lakoff. *The Contemporary Theory of Metaphor*, volume 2, pages 202–251. Cambridge University Press, 1993.
- [17] S. J. Luck and E. K. Vogel. The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657):279–281, 1997.
- [18] 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.
- [19] R. Mander, G. Solomon, and Y. Y. Wong. A “pile” metaphor for supporting casual organization of information. In *Proc. International Conference on Human Factors in Computing Systems (CHI '92)*, pages 627–634, 1992.
- [20] S. McDougall and M. Curry. More than just a picture: Icon interpretation in context. In *Proc. Coping with Complexity Workshop*, pages 16–17, 2004.
- [21] 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.
- [22] C. J. Morris, D. S. Ebert, and P. Rheingans. An experimental analysis of the pre-attentiveness of features in Chernoff faces. In *Proc. SPIE Applied Imagery Pattern Recognition: 3D Visualization for Data Exploration and Decision Making*, pages 12–17. 1999.
- [23] G. L. Murphy. On metaphoric representation. *Cognition*, 60:173–204, 1996.
- [24] A. Ortony. Why metaphors are necessary and not just nice. *Education Theory*, 25:45–53, 1975.
- [25] A. Pang and M. Clifton. Metaphors for visualization. In *Proc. Eurographics Workshop on Visualization in Scientific Computing*, 1995.
- [26] H. E. Pashler. *Attention*. Psychology Press, 1998.
- [27] C. Peirce. *The New Elements of Mathematics*. 1902.
- [28] L. R. Peterson and M. J. Peterson. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58(3):193–198, 1959.
- [29] I. Reppa, D. Playfoot, and S. J. P. McDougall. Visual aesthetic appeal speeds processing of complex but not simple icons. *Proc. Human Factors and Ergonomics Society Annual Meeting*, pages 1155–1159, 2008.
- [30] M. S. Sanders and E. J. McCormick. *Human Factors in Engineering and Design*. McGraw-Hill Book Company, 6th edition, 1987.
- [31] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, Jan. 2001.
- [32] 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.
- [33] T. G. Sticht. Educational uses of metaphor. In A. Ortony, editor, *Metaphor and Thought*, pages 621–632. Cambridge University Press, 1993.
- [34] A. Treisman. Properties, parts, and objects. In K. R. Boff, L. Kaufman, and J. P. Thomas, editors, *Handbook of Perception and Human Performance*, volume II, page 35:135:70. Wiley, 1986.
- [35] E. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 2001.
- [36] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers, second edition, 2004.
- [37] J. M. Wolfe. Guided search 2.0 - a revised model of visual-search. *Psychonomic Bulletin & Review*, 1(2):202–238, 1994.
- [38] M. K. Yousef. Assessment of metaphor efficacy in user interfaces for the elderly: A tentative model for enhancing accessibility. In *Proc. EC/NSF Workshop on Universal Accessibility of Ubiquitous Computing*, pages 120–124, 2001.
- [39] C. Ziemkiewicz and R. Kosara. The shaping of information by visual metaphors. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1269–1276, 2008.
- [40] C. Ziemkiewicz and R. Kosara. Preconceptions and individual differences in understanding visual metaphors. *Computer Graphics Forum*, 28(3):911–918, 2009.