Caren W. Campbell. Is Optical Optimal? Visualizing Document Sets to Facilitate Efficient Selection of Relevant Items. A Master's Paper for the M.S. in I.S. degree. May, 2012. 62 pages. Advisor: Ronald E. Bergquist

The visual representation of a cluster of similar text documents is typically accomplished with interactive information visualizations, yet some benefit can be derived even from static diagrams of the document collections. This study examines the comparative efficacy of text and diagram presentation formats for representing such collections, and attempts to determine the degree to which the format of an interface can affect users' ability to select relevant information being provided there. The hypothesis under examination holds that adults locate relevant documents more efficiently from a node-link diagram of the document set than from a textual list with the commonly used title-and-snippet format. Empirical data from a 40-person user study does not provide statistically significant support for the hypothesis; literature on similar studies helps to contextualize and explain this outcome, and directions for future study are suggested.

Headings:

Graphical User Interface

Information Visualization

Information Graphics

Human-Computer Interaction

TREC

Relevance

IS OPTICAL OPTIMAL? VISUALIZING DOCUMENT SETS TO FACILITATE EFFICIENT SELECTION OF RELEVANT ITEMS

By
Caren W. Campbell

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

May 2012

Approved by

_____

Ronald E. Bergquist

TABLE OF CONTENTS

I. INTRODUCTION

One of the fundamental roles of the information science professional is to enable precise, efficient retrieval of specific information upon demand. Behind the scenes this involves extended, collective, and carefully structured effort, such as the meticulous creation and application of metadata, the careful organization of relational databases for sharing data between institutions, and the compliance with established transfer protocols. Information retrieval (IR) researchers refine iteration after iteration of search algorithms, improving precision and recall to improve the systems' effectiveness. Thus optimized, the entire infrastructure of organized information is thereby harnessed for public use.

Between this enormous (indeed, worldwide) information infrastructure and the population of end users lies the search interface. Users enter their search queries in text form and receive a set of results (i.e., documents), at least some of which will probably be relevant to their information need. Even at this stage, however, the information need has not yet actually been met. Information seekers must examine the full text of the documents within a result set[1] in order to 1) understand the nature of each document and 2) assess the document's value. In other words, the first question a seeker asks about each result is "what is this?" and the second question is "what does it contain?" Only then can the document's relevance ("do the contents fulfill my information need?") be fully assessed.

---

[1] An exception would be a case of "good abandonment," where the information shown in the headline or text snippet satisfies the query (Chuklin & Serdyukov, 2012).

As a general rule, a standard text search will present users with a document set displayed in a textual format: a list of document titles, each followed by its Uniform Resource Locator (URL) and a snippet of the document's body text. Figure 1 shows the output of a Google text search, a typical example of this format. This presentation style embodies an expectation that users will assess the documents in a linear, sequential way, starting with the items located at or near the top of the list. Indeed, by design, search engine relevance algorithms encourage this approach by placing the "most relevant" documents at the beginning of the result list.

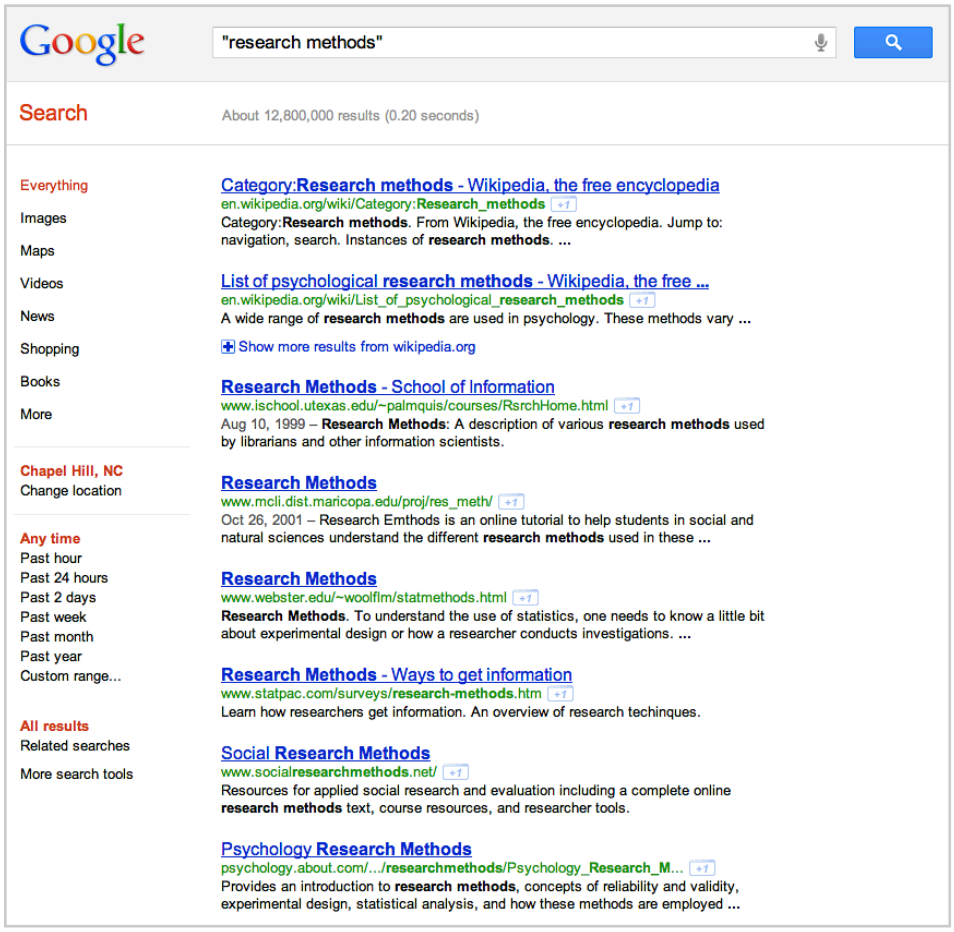


FIGURE 1 *Sample Output from a Google Text Search*

Because readers' behavior can be expected to conform to the principle of least effort (Zipf, 1949), they will usually begin by selecting the top item in a list (Joachims, Granka, Pan, Hembrooke, & Gay, 2005; O'Brien & Keane, 2006). Indeed, eye-tracking studies provide empirical evidence of users' habit of zeroing in on the top items first. Investigation by Nielsen (2006), among many others, consistently demonstrates users' gaze going first and longest at the horizontal zone at the top of the screen and then at a vertical zone along the left margin, as if tracing the capital letter *F*.

Such habituation may prevent users from noticing the flaws of the text format. As shown in recent studies of Google search results and "trust bias" (Joachims et al, 2005; O'Brien & Keane, 2006), users' acculturation to the top-to-bottom presentation mode may hinder their information-gathering mission (i.e., selecting the document with the greatest relevance). When users select documents on the basis of their list position, false starts or time-wasting detours — additional sources of potential frustration – may ensue.

The linear-sequential style also presents some significant usability drawbacks. To grasp how any given result document relates to the search query (the semantic connection) and compares to other search results (the ontological import), users must read the text snippet or follow the headline's hyperlink to the full article. In reading and assessing text, users expend a measure of cognitive effort, plus additional effort to remember and compare the contents (Bettman, Johnson & Payne, 1990). And given all this following of links, the speed with which users can fulfill an information need is necessarily affected by the nature and availability of their Web connections.

By contrast, a non-linear graphical display of a document set can provide cues to indicate each document's semantic content at the same time as the display positions that

content within an ontology, i.e., the universe of documents selected by the search engine's retrieval algorithm. That ontology will have a node-link relational structure, enabling users to determine the point of entry and order of access, with no predetermined path (Boechler, 2001).

Should such interfaces improve information seekers' ability to retrieve and quickly identify relevant documents from the Web, the effects on productivity and task satisfaction are likely to be significant. Visualized display of document sets would be a boon to librarians and educators as well, enabling concept-based presentations of journal article search results or of the holdings of the entire catalog. Thus not only users, but also information professionals, can benefit from the implementation of clearer and more intuitive displays of document sets.

II. LITERATURE REVIEW

The body of research on the graphical presentation of text documents is large and varied, and recent studies on visualizing text documents have rolled out interactive prototypes in a variety of formats. Because the visual ordering of intangibles, such as document content, calls for abstraction (Chen, 2010), all visualization necessitates an organizing metaphor — though the specific metaphorical model is less important than whether its schema effectively conveys the intended meaning (Chen, p. 388). A brief overview of these metaphorical models is warranted.

METAPHORICAL MODELS

Within the scholarly literature, radial representations are common. The influential Lyberworld project (Hemmje, Kunkel, & Willet, 1994) depicted "content space" with "relevance spheres." Other iterations using concentric circles or spheres include DocBall (Vegas, Crestani, & de la Fuente, 2007); DART (Amar, Day, Godfrey, & Plaue, 2004), which (naturally) uses a dartboard metaphor; DocuBurst (Collins, Carpendale, & Penn, 2009), which deploys radial graphs and coxcombs to encode semantic content; and Wivi (Lehmann, Schwanecke, & Dorner, 2010), which does the same for Wikipedia articles.

Also, maps remain a frequent and useful paradigm. Websom (Kaski, Honkela, Lagus, & Kohonen, 1998) uses latent semantic analysis to place documents on a "word category map." ResultMap (Clarkson, Desai, & Foley, 2009) works from hierarchical

metadata to create a squarified treemap representation of library holdings in context. Astronomical map metaphors inform InfoSky (Andrews et al., 2002), which creates nodes to represent documents, topics, and other semantic entities; and WebStar (Zhang & Nguyen, 2005), which uses hyperlinks as a basis for identifying nodes and the relationships among them.

Morphologically similar to starburst maps, node-link diagrams have become the format of choice recently for some well-known visualization researchers. Two high-profile instances from recent years are TopicNets, which colorfully depict semantic entities as connected nodes of different types (Gretarsson et al., 2012); and PhraseNet (van Ham, Wattenberg, & Viegas, 2009), which maps and links unstructured text units within documents, rather than across documents.

As noted, all of these prototypes are interactive. Interactivity is invaluable for full examination of a data set. It allows seekers to zoom in on useful information while filtering out the unhelpful or provide additional detail, as per Ben Shneiderman's famous mantra, "Overview first, zoom and filter, then details on demand" (1996). Its affordances enable users to explore data, thereby rendering search an iterative process (Koshman, 2004). The reliance on these functionalities implicitly asserts that multi-dimensional data sets necessitate interaction for sensemaking.[2]

THE ROLE AND CREATION OF STATIC DIAGRAMS

Despite the undeniable value of interactive functionality, this study sets aside the

[2] Note that interactivity alone does not suffice; an evaluation of visualization tools by Kobsa (2001) found representational accuracy of only 68–75% on simple user tasks.

topic of interaction[3] in order to assess the value of static graphical depictions, which might be considered subunits of an interactive visualization. Indeed, the ability to comprehend a two-dimensional static depiction must by definition precede interactive engagement: users must be able to make sense of simple diagrammatic representations before deploying interactive tools. On this principle, the current investigation tested a simple hypothesis — that information seekers locate relevant documents more efficiently within a diagrammatic depiction of a document set than within a list of the same items.

Not surprisingly, given their part-to-whole functional relationship, static and interactive renderings must resolve similar representational issues. One such issue is the multi-dimensionality of a textual data set. Even the most sophisticated and complex interactive visualizations cannot codify every single dimension of a text collection. When dimensions are so numerous that it is impossible to depict them all, or so redundant that it is unnecessary to keep them all, reduction of dimensionality becomes necessary. (See Ingram, Munzner, Irvine, Tory, Bergner, & Moller, 2010; and Cribbin, 2010.) Thus, researchers must devise a technique that will reduce dimensionality without simultaneously reducing semantic content.

Koshman (2006) distinguishes between the dimensionality of data and the dimensionality of representation, and concludes that the two need not be equivalent. (Indeed, from a theoretical perspective, they could not be: data may have hundreds of dimensions, but visual depictions, especially static ones, clearly cannot.) To that end, many in the field have turned to dimensionality reduction via clustering (Lagus, Kaski, &

[3] Consequently this study does not address applications involving active information seeking or visual analytics. Mechanisms of implementation are also beyond the scope of this investigation.

Kohonen, 2004), often on the basis of latent semantic analysis (Landauer, Laham, & Derr, 2004), or a related technique, edge compression, which has the effect of node reduction (van Ham et al, 2009). Faced with billion-record data sets, Shneiderman (2008) proposes atomic, aggregated, and density plots. To calibrate the degree of compression, Venna, Peltonen, Nybo, Aidos & Kaski (2010) present a reduction mechanism that conceptually parallels the IR concepts of precision and recall: each user decides how much dimensionality to sacrifice, balancing the desire to avoid irrelevant documents with the fear of potentially missing similar documents.

Most of these methods, however, apply only to visualization modalities that are interactive and multi-scale. When one is creating static diagrams of text documents, further constraints apply, even as the abstraction level and the dimensionality of the data remain high. Options such as motion, reorientation, temporal sequencing, filtering and "focus+context" renderings are eliminated. Static visualizations still depend on the tools and rules described by Bertin: “retinal variables” by which to differentiate graphical marks, plus guidelines for positioning arrays and other spatial displays. Rogowitz, Treinish, and Bryson (1996) list ten ways that visual encoding can differentiate among dimensions, of which eight can be applied to static visualizations. Gestalt principles — e.g., proximity, similarity, closure, symmetry — and the implications of color take on additional importance (Sayim, Westheimer, & Herzog, 2010; Rogowitz et al.).

Given these constraints, the practitioner who seeks to express dimensions of ontology (context) and taxonomy (structure) will probably have to employ verbal content. The interactive prototype WordBridge (Kim, Ko, Elmqvist, & Ebert, 2011) does so by assembling tag clouds into a node-link diagram to show relationships. Recurrent terms, in

the form of tag clouds, characterize nodes as well as links. This reliance on term frequency evinces a clear relationship to information retrieval, and thus points a way toward the depiction of a set of search results such as those being represented in this study. As a result, this study's static renderings bear some conceptual resemblance to WordBridge, in that nodes contain document titles, and edges represent frequent, shared terms.

III. STUDY DESCRIPTION

PARTICIPANTS

To assess the value of the study's static visualizations, 40 adult volunteers were recruited via social media. The principal investigator placed a solicitation notice on her Facebook page and on that of her husband. Note that this convenience sample may or may not yield generalizable results, given the group's relative homogeneity in age and education. Table 1 presents a summary of demographic characteristics. (Note that this population sample might be atypical, given the distribution of gender and educational level; further discussion of this issue follows in Chapter V.)

gender		**education level**	
male	52.5%	some college	7.5%
female	47.5%	bachelors degree	35.0%
		some graduate school	12.5%
first language		masters degree	40.0%
English	97.5%	Ph. D.	2.5%
not English	2.5%	J.D.	2.5%

TABLE 1 *Demographic Information About Study Participants*

METHODOLOGY

GENERATING THE DATA

SESSION OVERVIEW

After submitting a signed consent form, each subject used his/her own Web-enabled computer to participate in an online session that lasted roughly 30 minutes (far less in most cases). The session consisted of reading instructions, working an interface exercise that contained three informational tasks, and then completing a brief supplemental questionnaire about such personal traits as computer use, education level, gender and age. The host locations for the exercise and the questionnaire respectively were ProProfs (www.proprofs.com) and Adobe Forms Central (www.adobeformscentral.com).

TASK DESCRIPTION

The premise for the task is that participants should imagine they have been assigned to write an essay on a given topic, which is carefully defined and described on screen. The topics were wildlife extinction, international art crimes, and attacks by black bears. All exercises presented them in this order so as to eliminate any confounding effect arising from question sequence. For each of the exercise's three tasks, subjects saw nine documents and were asked to choose the three that they considered most relevant ("useful") for writing the essay. (See instructions in Appendix A.)

TASK FORMATS

In order to ascertain the effect of format on item selection, it was necessary to

present each task in one of two ways: a textual format, as shown in Figure 2, and a node-link diagram format, as shown in Figure 3. As discussed in Chapter II, the node-link format was chosen because it proved the most informative (among static diagram options) in expressing the relationships between documents.

Whether a participant saw the text or diagram format for each question was determined by random assignment. (See next section.) After each subject finished the session, his/her choices for "most useful" were compared to the subset of documents adjudged relevant by the U.S. National Institute of Standards and Technology's Text Retrieval Conference (TREC).[4] The use of TREC's materials and ratings enabled the investigator to minimize her own interpretive bias in the matters of topic selection, the aggregation of documents into sets, and judgments about a document's relevance.

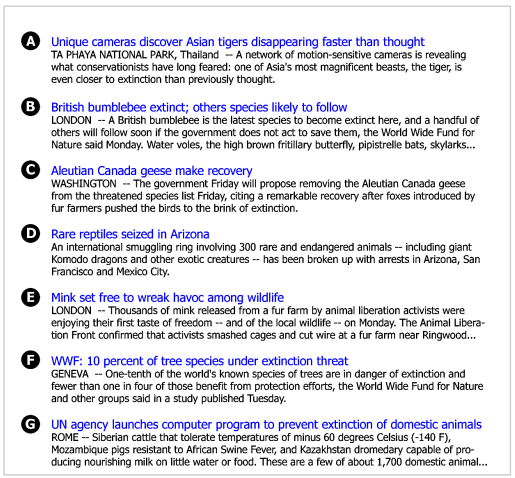
TASK PRESENTATION

To enable assessment of a subject's performance in both format categories, each participant saw at least one document set in textual format (Figure 2) and at least one document set in diagram format (Figure 3). For the third question, roughly half of the group was shown a text display and the other half was shown a diagram display.
To reiterate, each of the three questions could appear in either presentation mode.
(See Table 2, below.)

Ensuring that both presentation modes expressed the same information was essential in order to avoid skewing the results of the investigation. Thus it was necessary (if ironic) to deliberately exclude graphical aspects such as positioning, color, and shape, which could have conveyed additional information to users.

---

[4] http://trec.nist.gov/data.html



A Unique cameras discover Asian tigers disappearing faster than thought
TA PHAYA NATIONAL PARK, Thailand -- A network of motion-sensitive cameras is revealing what conservationists have long feared: one of Asia's most magnificent beasts, the tiger, is even closer to extinction than previously thought.

B British bumblebee extinct; others species likely to follow
LONDON -- A British bumblebee is the latest species to become extinct here, and a handful of others will follow soon if the government does not act to save them, the World Wide Fund for Nature said Monday. Water voles, the high brown fritillary butterfly, pipistrelle bats, skylarks...

C Aleutian Canada geese make recovery
WASHINGTON -- The government Friday will propose removing the Aleutian Canada geese from the threatened species list Friday, citing a remarkable recovery after foxes introduced by fur farmers pushed the birds to the brink of extinction.

D Rare reptiles seized in Arizona
An international smuggling ring involving 300 rare and endangered animals -- including giant Komodo dragons and other exotic creatures -- has been broken up with arrests in Arizona, San Francisco and Mexico City.

E Mink set free to wreak havoc among wildlife
LONDON -- Thousands of mink released from a fur farm by animal liberation activists were enjoying their first taste of freedom -- and of the local wildlife -- on Monday. The Animal Liberation Front confirmed that activists smashed cages and cut wire at a fur farm near Ringwood...

F WWF: 10 percent of tree species under extinction threat
GENEVA -- One-tenth of the world's known species of trees are in danger of extinction and fewer than one in four of those benefit from protection efforts, the World Wide Fund for Nature and other groups said in a study published Tuesday.

G UN agency launches computer program to prevent extinction of domestic animals
ROME -- Siberian cattle that tolerate temperatures of minus 60 degrees Celsius (-140 F), Mozambique pigs resistant to African Swine Fever, and Kazakhstan dromedary capable of producing nourishing milk on little water or food. These are a few of about 1,700 domestic animal...

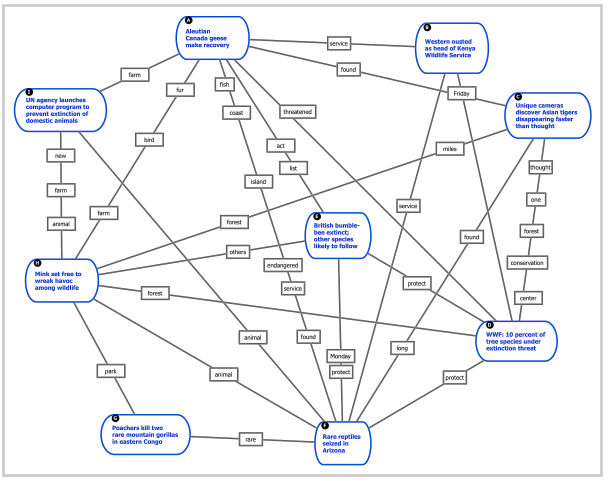


FIGURE 2

Text: Wildlife Extinction Task

FIGURE 3

Diagram: Wildlife Extinction Task

Within the node-link network diagram format (Figure 3), each node contains a document's title, and connections ("linkages") derive from recurrent terms that co-occur in more than one document. Proximity and placement have no particular meaning; this was intended to minimize the possibility of positional trust bias.

The text interface (Figure 2), used to generate a performance baseline, was styled to resemble the familiar Google output shown in Figure 1. The list contains document titles in a large, blue font, each with a snippet of body text in a smaller black typeface. Certain elements of the Google-style list (document URLs, links to relevant domains, image links, faceted index along the left margin) were left out to ensure that the text format would provide only the information that would also appear in the diagrams.

To minimize confounding effects arising from the order in which the formats were shown, the investigator created four versions of the exercise. As noted above, each version contained a mix of formats, as shown in Table 2. (All four versions may be seen

in their entirety in Appendix A.) Versions were randomly assigned to participants as follows: Ten instances of a set of four (A, B, C, and D) were auto-generated using the Research Randomizer Tool.[5] This sequence dictated the order for assigning versions. As each subject enrolled in the study, he or she was e-mailed the link to whatever version came next in the randomized list.

	VERSION A	VERSION B	VERSION C	VERSION D
QUESTION 1 (EXTINCTION)	TEXT	DIAGRAM	TEXT	DIAGRAM
QUESTION 2 (ART CRIME)	DIAGRAM	TEXT	DIAGRAM	TEXT
QUESTION 3 (BLACK BEARS)	TEXT	TEXT	DIAGRAM	DIAGRAM

TABLE 2 *Question Formats for Each Version of the Exercise*

TASK SCORING

The investigator then compared the users' performance with text-interface tasks to that with graphical-interface tasks. Performance was scored as follows: For each information task, an efficiency score was calculated. Within this study, "efficiency" is defined as accuracy divided by speed. Accuracy, the fraction's numerator, equals the number of correctly selected documents per task, i.e., the number of "useful" documents selected by the participant that match those labeled "relevant" by TREC judges. Speed, the fraction's denominator, equals the elapsed time (in seconds) for completing the task,

[5] www.researchrandomizer.com

i.e., selecting three documents. As a result, for one format to be more efficient than the other, one of two situations must pertain: 1) greater accuracy and equal speed; or 2) equal accuracy and greater speed, as shown in the efficiency scoring matrix (Figure 4). In cases where both formats yield equal efficiency, that outcome was not considered affirmative support for the research hypothesis, but was considered satisfactory from a usability perspective. The results of the three tasks executed by 40 participants (120 measurable instances) were then subjected to statistical analysis to assess outcomes and potential correlations.

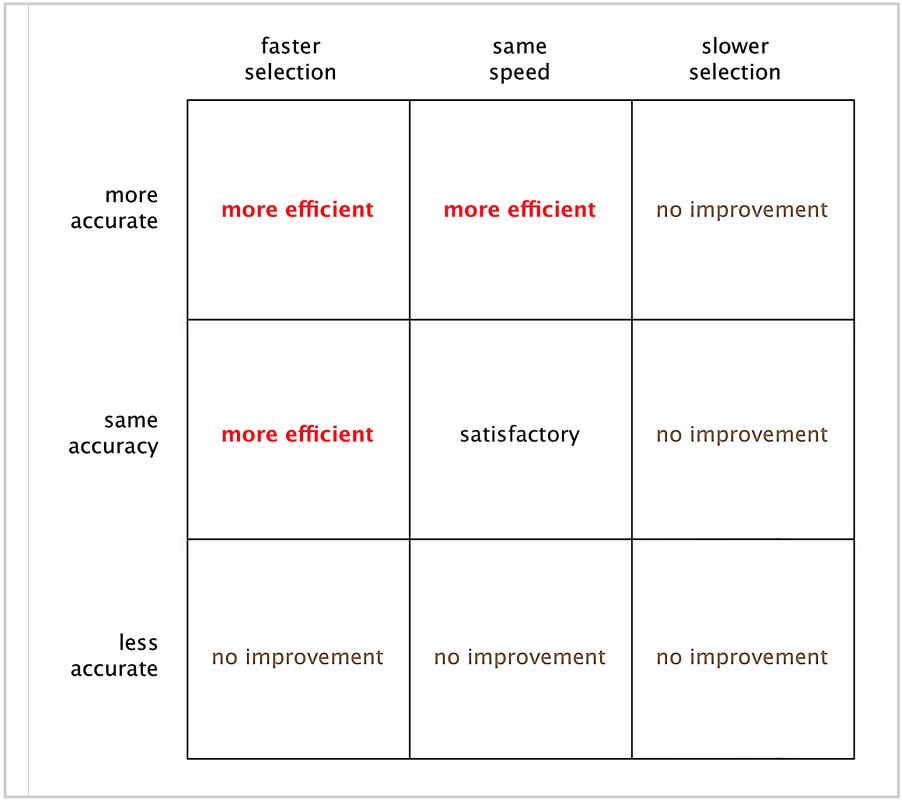


FIGURE 4 *Efficiency Scoring Matrix*

QUESTIONNAIRE

Also analyzed in conjunction with task performance were some personal characteristics of the participants, elicited via a brief questionnaire at the end of the exercise. The eight-question form asked subjects to provide information about age, gender, first language, education level, document search frequency, daily amount of computer use, favored search engine, and typical level of focus during computer work. (The questionnaire can be found in Appendix A.) Participants were free to answer as many or as few of these questions as they desired; indeed, one subject declined to fill out the questionnaire entirely.

COLLECTING AND COMPILING THE DATA

Commercial vendors ProProfs and Adobe Forms Central, respectively, hosted the main exercise and the exit questionnaire. Task answer data, including selections, accuracy, and elapsed time per task, was logged by ProProfs and extracted by the investigator. Questionnaire answers were compiled by Adobe and similarly downloaded by the investigator. Processing the data with statistical software (JMP and XLSTAT) yielded a detailed depiction of the study results, as will be discussed in the following chapter.

IV. DATA ANALYSIS AND DISCUSSION

As indicated in Chapter I, the group of experimental subjects consisted of 40 adults: 21 men and 19 women between the ages of 30 and 65, with a median age of 46. With three tasks per participant, the investigator thus had the opportunity to collect 120 observations. However, in 10 instances, subjects selected more than three documents for the task. In such cases it was impossible to determine which three documents the subject preferred most, so these instances were nullified. After their removal, the final count of observations was 110. That data set was then analyzed for correlation and statistical significance.

OVERVIEW OF PERFORMANCE

As shown in the data summary (Table 3), the average time for diagram tasks was 148.11 seconds (2.47 minutes), while the average time for text tasks was 173.96 seconds (2.90 minutes); thus the average text task was more than 17% slower. Median times for text and diagram, respectively, are 136.5 seconds (2.28 minutes) and 125.5 seconds (2.09 minutes), meaning the median text task was 8.8% slower. Although both average and median figures for text-task accuracy were higher than those for diagram tasks, the speed of the diagram tasks so exceeded that of the text tasks that the diagram format (both average and median) ultimately prevailed on the efficiency measure. Average efficiency for diagram tasks was 30% higher than for text tasks; median efficiency for diagram tasks was 15.4% higher.

format		average	median
TEXT (n=54)	speed	173.9630	136.5
	accuracy	1.4815	2
	efficiency	0.0110	0.0091
DIAGRAM (n=56)	speed	148.1071	125.5
	accuracy	1.4107	1
	efficiency	0.0143	0.0105

TABLE 3 *Average and Median Scores, by Format*

Results from the three topics and two formats are compared in the three graphs in Figure 5. (Additional graphs of these results can be found in Appendix C.) From this data, several notable patterns emerge and are discussed in the following subsections.

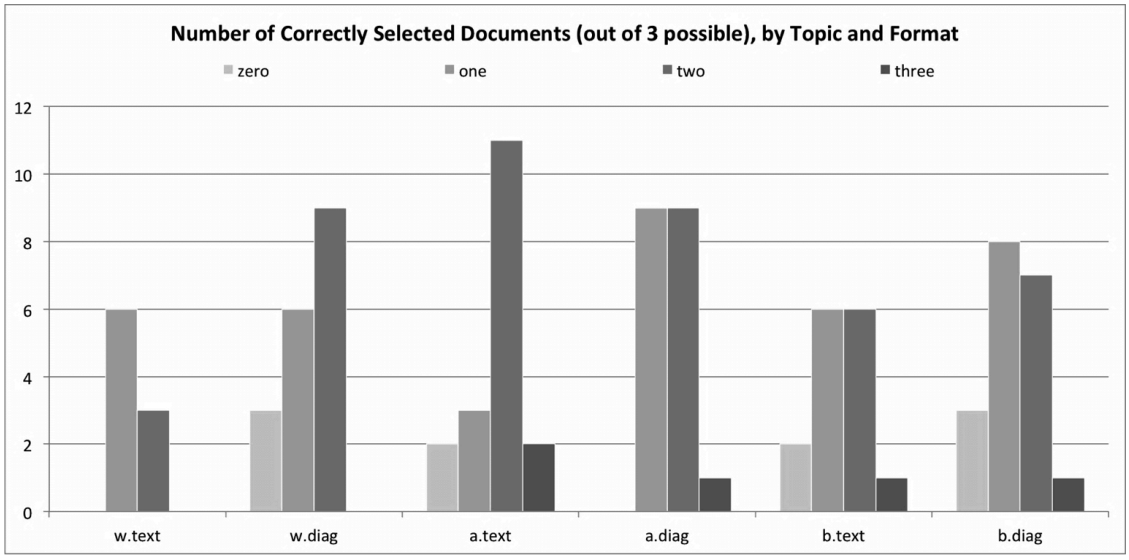


FIGURE 5a *Accuracy by Question and Format*

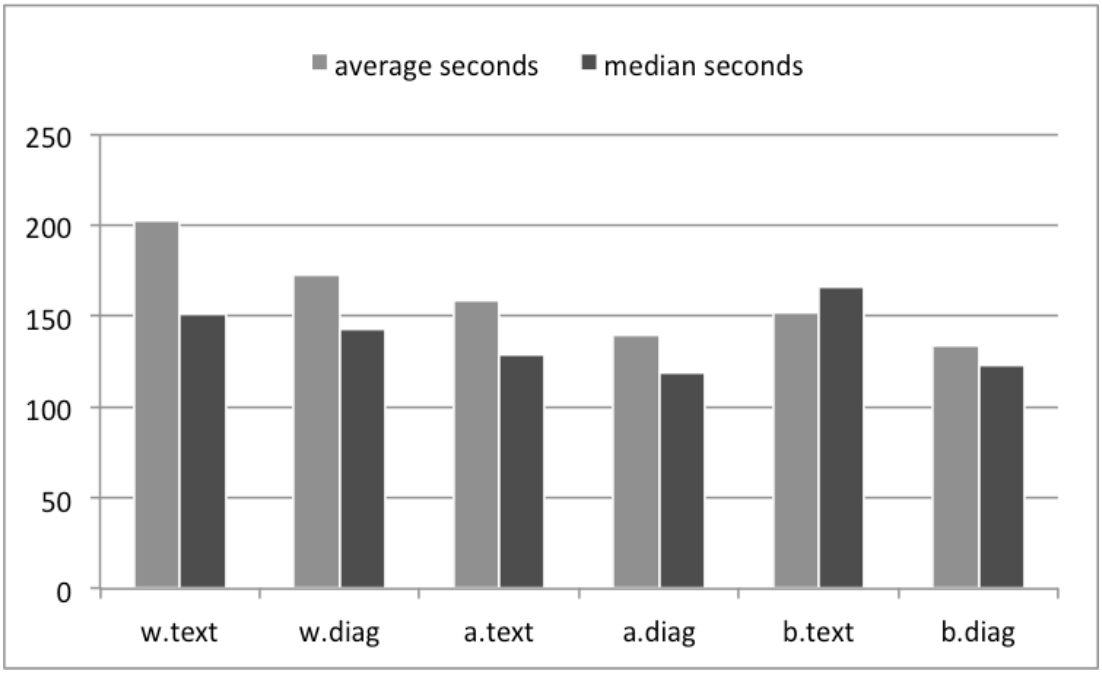


FIGURE 5b *Speed by Question and Format*

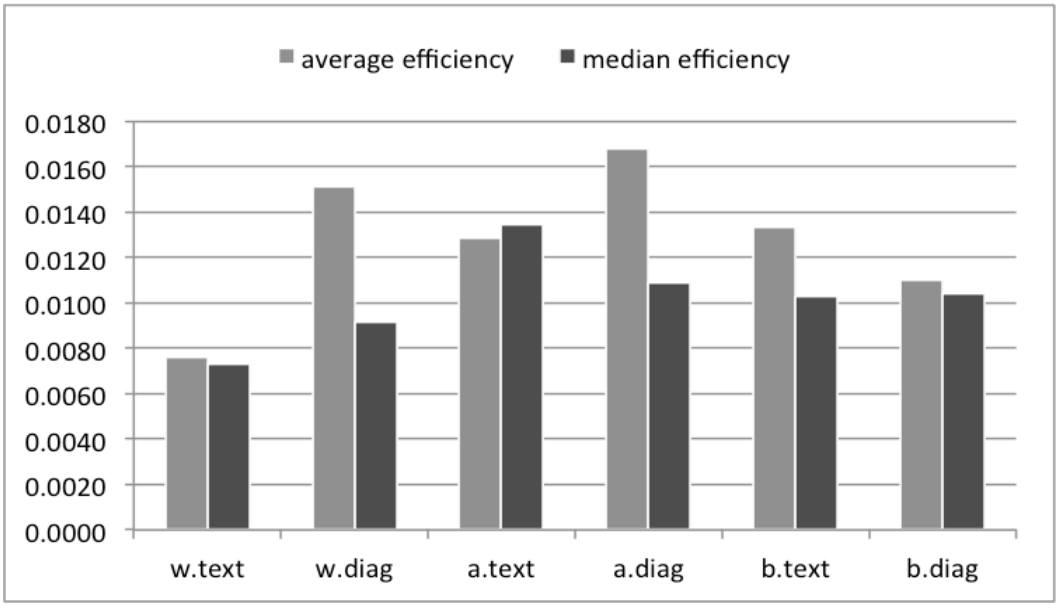


FIGURE 5c *Efficiency by Question and Format*

TEXT FORMAT

Correctly selected documents from the text format numbered 64 overall, as shown in Figure 5a. Average efficiency improved as the exercise progressed: the average efficiency score for the wildlife extinction task (.0076) was exceeded by that of the art crime task (.0129), which was exceeded by that of the black bear task (.0133). This is almost certainly a reflection of the decrease in average elapsed time, as shown in Figure 5b. Average time for the first task averaged 203 seconds; the second task, 159 seconds; and the third task, 152 seconds. For the wildlife and black bear tasks, the mean efficiency score exceeded the median score (positive skew), indicating that a small faction of participants performed efficiently enough to raise the average of the entire group. Also, the highest median efficiency score overall is associated with a text instance, that of the art crime task. Meanwhile, the highest mean efficiency score is associated with a diagram instance — also that of the art crime task. From these two results, it can be inferred that participants found the art crime task easiest to execute successfully, regardless of format.

DIAGRAM FORMAT

The diagram format resulted in the selection of 79 relevant documents, exceeding the number of text instances by 23.4%. Most of the difference in performance is associated with the wildlife question: the text format elicited only 12 correctly selected documents, versus 24 for the diagram format.

For two of the tasks (wildlife extinction and art crime), the average efficiency score exceeds the median by 20.6% and 17.6% respectively. As noted above, this

indicates that efficiency performance among a few participants far exceeded that of the sample as a whole.

As with the text format, average and median speeds within the diagram format improved from task to task through the entire exercise. Average speed for the wildlife extinction task was 173 seconds; for the art crime task, 140 seconds; for the black bear task, 133 seconds.

FORMAT COMPARISONS WITHIN TOPICS

WILDLIFE EXTINCTION

As noted above, the diagram version of this question earned twice as many correct selections as did the text version of the question. In part, this is a result of a smaller number of text instances; an unusually high number of the null results were associated with this question and format. Consequently, for this topic, the diagram format's efficiency figures proved to be 98.7% (mean) and 24.6% (median) better than those for the text format.

ART CRIME

Text and diagram formats proved roughly equal in eliciting correct selections, numbering 31 and 30, respectively. For the diagram version, however, the distribution of correct answers per participant is more uniform, with 18 of the 19 instances showing correct selections of one or two documents, and no instance where zero correct documents were chosen. The text version yields a performance distribution that is slightly closer to the normal bell curve; in two instances participants got zero selections right, and in two instances participants selected all three documents correctly. A comparison of the

results in the two formats shows that although the median duration for the text version was 8.4% longer, its median efficiency was also 23.8% better.

BLACK BEAR ATTACKS

For this task, the text format's median duration is 14 seconds longer than its mean, while the diagram format's median duration is 10 seconds shorter. No explanation can be offered for why the skew is so large, or why the direction differs between formats. Meanwhile, accuracy results for both formats reveal themselves to be normally distributed around the mean.

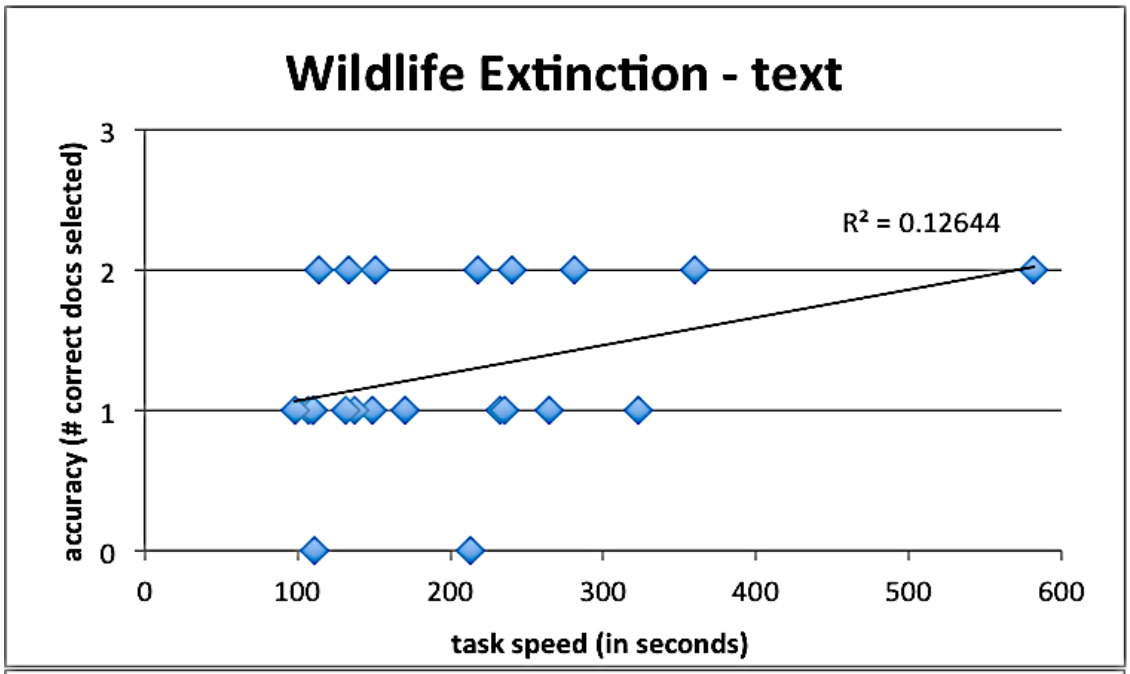
OVERVIEW OF CORRELATIONS

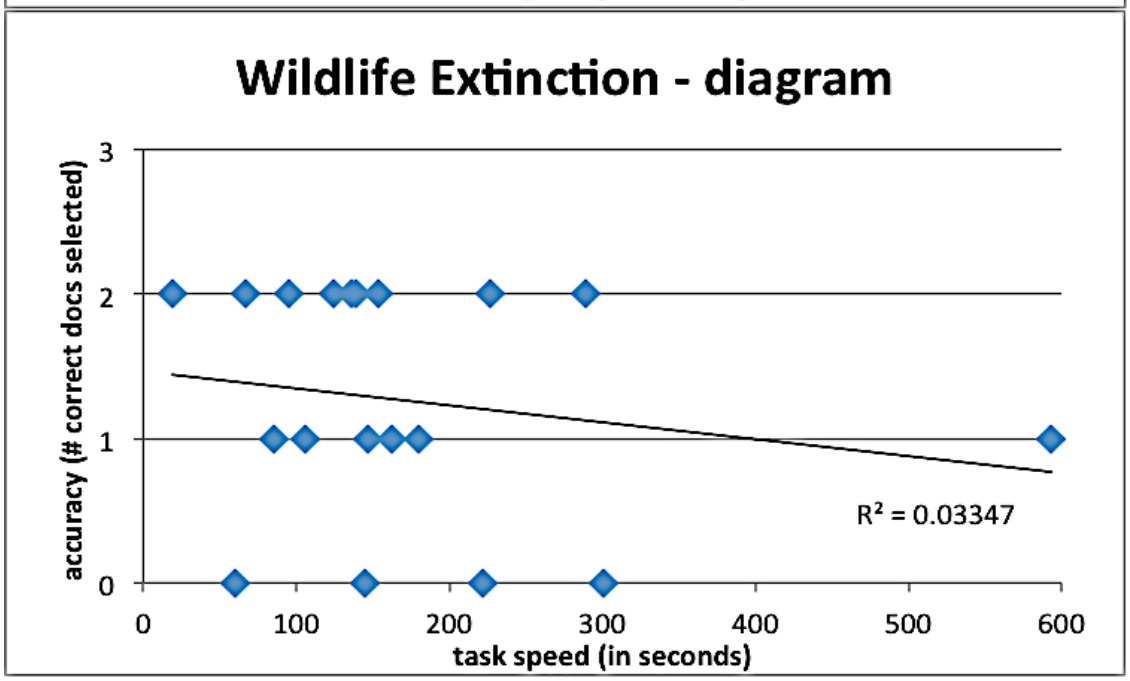
Correlation matrix - Pearson (n)

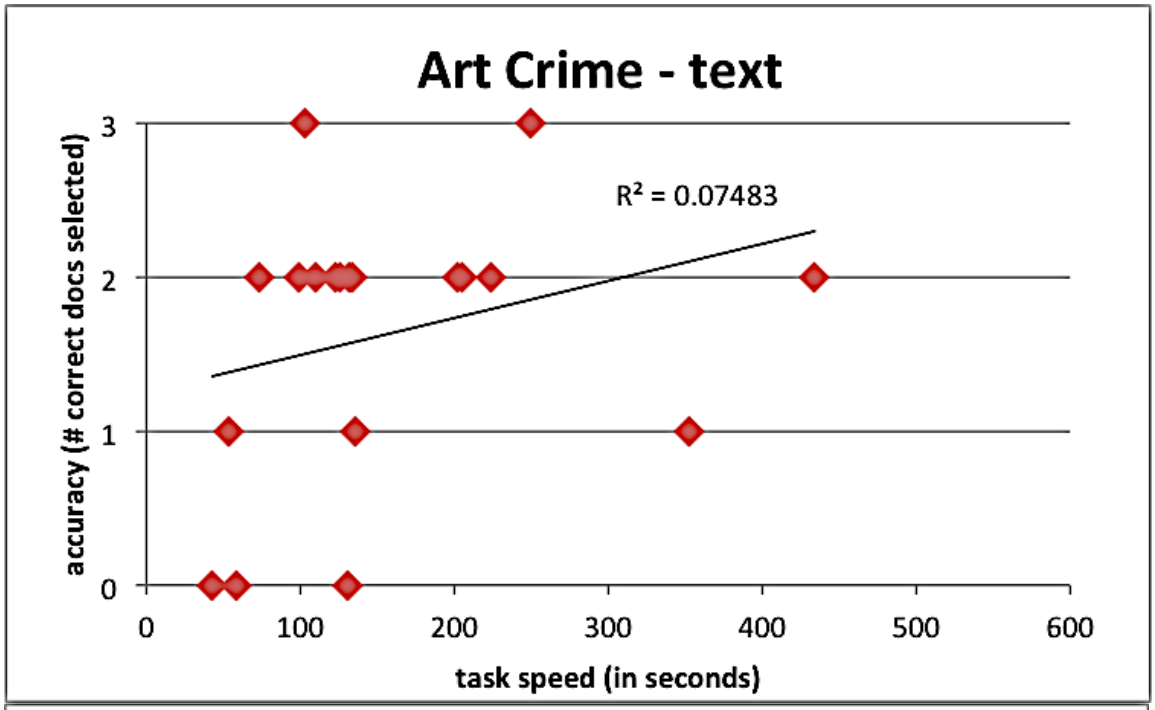
Variables	speed (in seconds)	accuracy (# correct)	efficiency (accuracy/speed)
text	0.112	0.041	-0.101
diagram	-0.112	-0.041	0.101
approx. age	0.320	0.074	-0.118
female	0.062	-0.046	-0.121
male	-0.062	0.046	0.121
English is first language	-0.089	0.275	0.172
English is not first language	0.089	-0.275	-0.172
some college	0.155	0.140	0.018
bachelors degree	0.052	-0.012	0.101
some graduate school	-0.173	0.064	0.143
masters degree	-0.036	-0.121	-0.174
doctorate	-0.057	0.051	0.035
professional degree	0.092	-0.031	-0.103
focus mostly on one task	0.079	-0.079	-0.157
multitask	-0.057	0.026	0.124
no focus in particular	-0.053	0.132	0.071
Google	-0.178	-0.015	0.085
Bing	0.178	0.015	-0.085
Est # searches yesterday	0.037	0.043	0.111
Est # hrs/day online (non-game)	-0.093	0.096	0.123

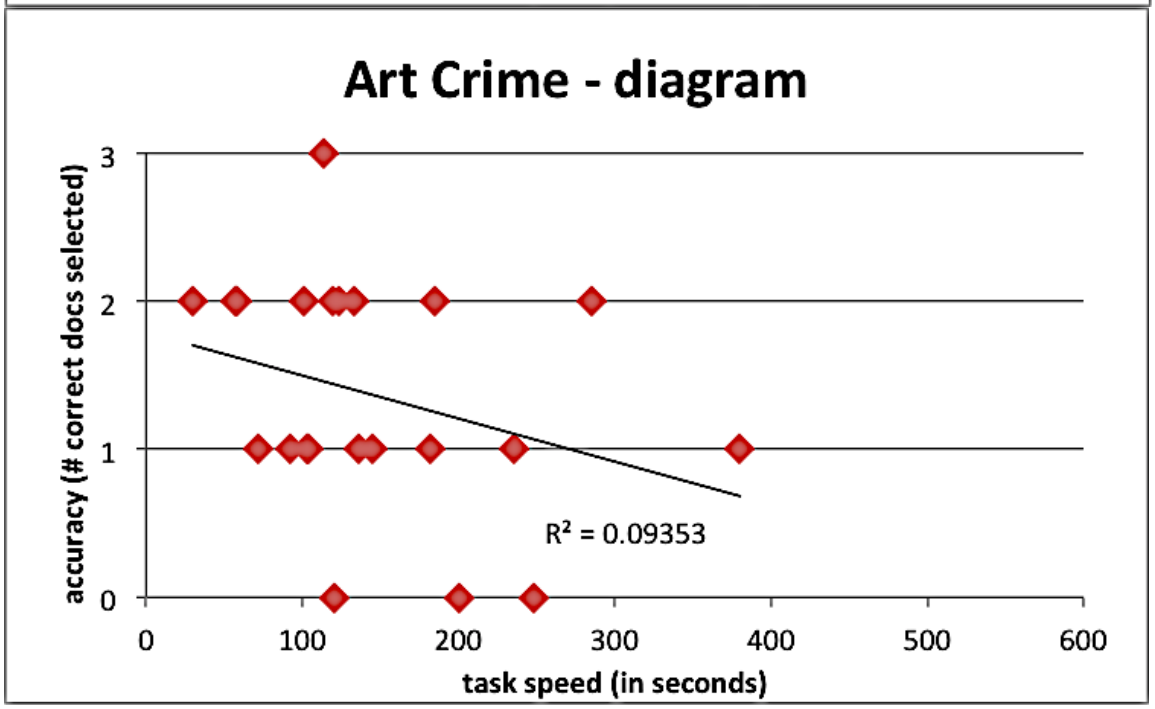
TABLE 4 *Correlation Matrix Containing All Factors Under Consideration*

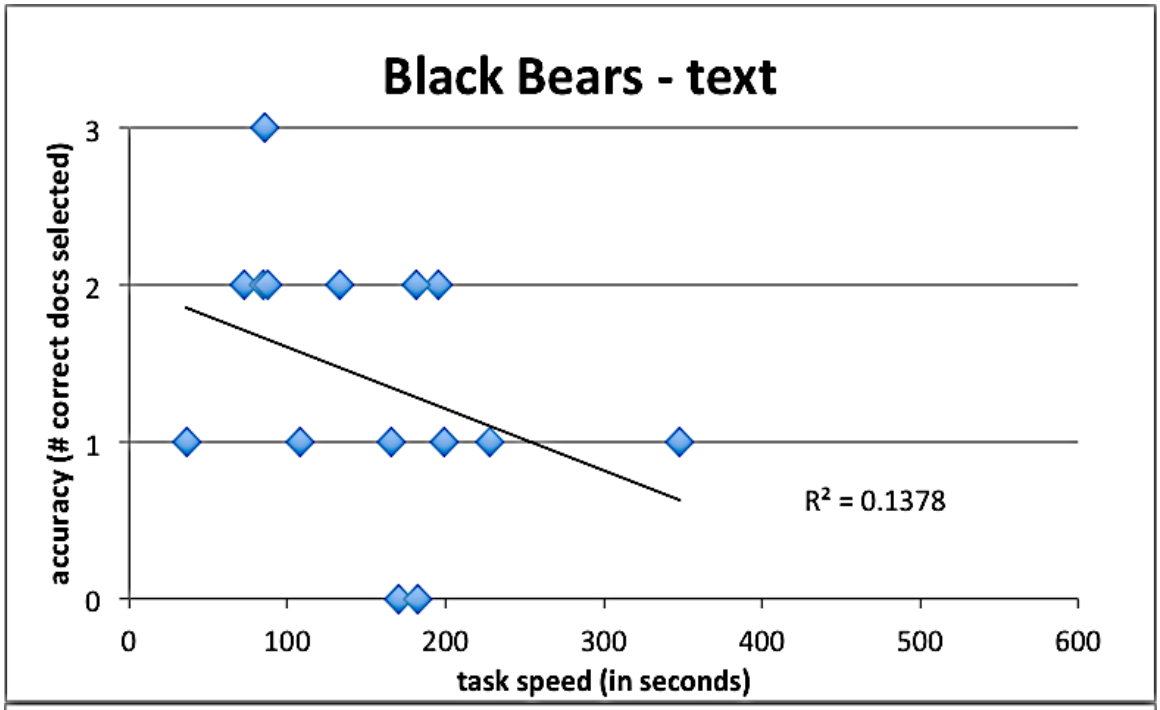
FIGURE 6 *Correlation Between Speed and Accuracy for Each Task in Each Format*

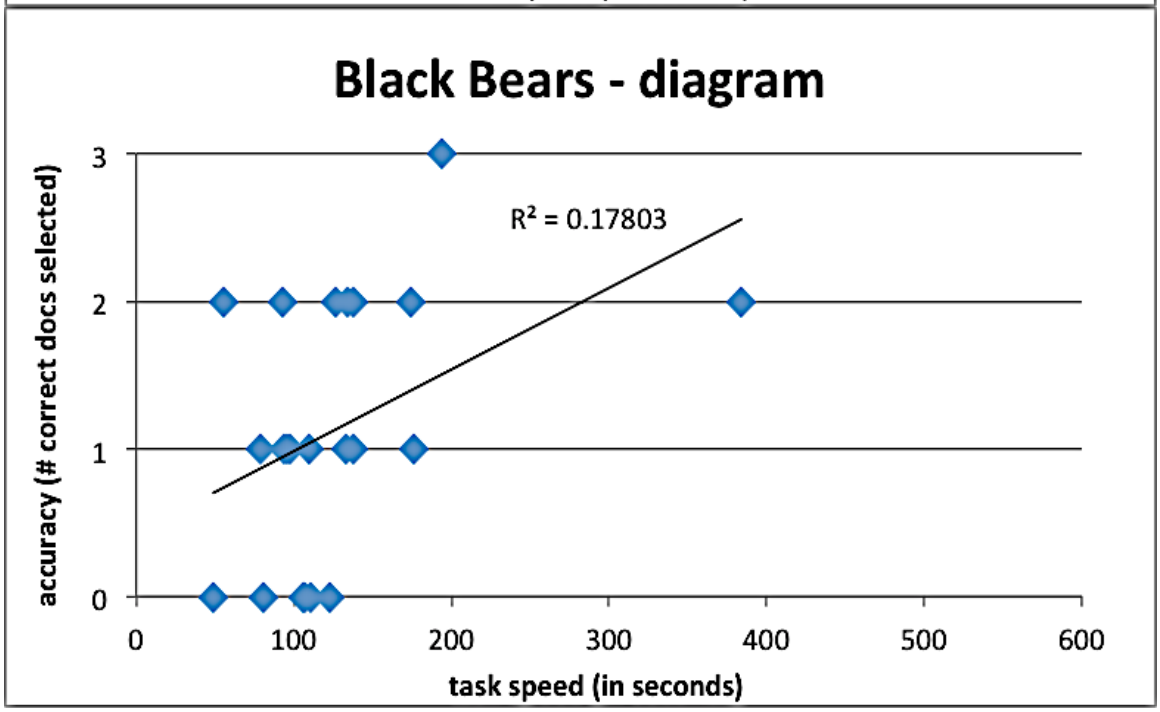





Art Crime - text
accuracy (# correct docs selected)
task speed (in seconds)
R² = 0.07483
0
1
2
3
0
100
200
300
400
500
600


Art Crime - diagram
accuracy (# correct docs selected)
task speed (in seconds)
R² = 0.09353
0
1
2
3
0
100
200
300
400
500
600


Black Bears - text
accuracy (# correct docs selected)
task speed (in seconds)
R² = 0.1378
0
1
2
3
0
100
200
300
400
500
600


Black Bears - diagram
accuracy (# correct docs selected)
task speed (in seconds)
R² = 0.17803
0
1
2
3
0
100
200
300
400
500
600

Correlation matrix (Pearson):

Variables	accuracy	interface
accuracy	**1**	0.048
interface	0.048	**1**

Values in bold are different from 0 with a significance level alpha=0.05

p-values:

Variables	accuracy	interface
accuracy	**0**	0.617
interface	0.617	**0**

Values in bold are different from 0 with a significance level alpha=0.05

Correlation matrix (Pearson):

Variables	interface	total seconds
interface	**1**	0.129
total seconds	0.129	**1**

Values in bold are different from 0 with a significance level alpha=0.05

p-values:

Variables	interface	total seconds
interface	**0**	0.180
total seconds	0.180	**0**

Values in bold are different from 0 with a significance level alpha=0.05

Correlation matrix (Pearson):

Variables	efficiency	interface
efficiency	**1**	-0.125
interface	-0.125	**1**

Values in bold are different from 0 with a significance level alpha=0.05

p-values:

Variables	efficiency	interface
efficiency	**0**	0.194
interface	0.194	**0**

Values in bold are different from 0 with a significance level alpha=0.05

FIGURE 7 *Correlation of Interface Format with Accuracy, Speed and Efficiency*

By task topic. The black bear task results for both formats present the tightest fit around the trend line (Figure 6), with r-values of 0.14 and 0.18. Of course, these and all other r-values shown on the graphs are far too low to be statistically significant.

By interface. Overall the interface type correlated very little with speed, accuracy, or efficiency (Table 4). None of the correlation factors exceeded an absolute value of 0.112.

By gender. Men performed slightly better overall; gender correlation was 0.121 for men and -0.121 for women.

By age. Younger people performed slightly better, with age having a correlation value of -0.118 with accuracy.[6] However, age had a comparatively strong positive correlation (0.320) with speed. This correlation factor is the largest value of any shown in Table 4.

By education level. The greatest correlation was in fact a negligible one (-0.174), relating performance to the possession of a master's degree.

By computer habits. Participants who typically spend more hours per day online achieved greater efficiency, but the correlation factor was a mere 0.123, again far below the level of statistical significance.[7] Also, task efficiency correlated with attention habits to a surprisingly small degree; cases where subjects reported the habit of focusing on one task at a time ended up with the largest negative correlation (-0.157), while multitasked instances had the largest positive correlation (0.124).

Unfortunately, certain factors could not be considered as possible correlates (or even meaningful indicators), because participant counts for those aspects skewed overwhelmingly in one direction. One such factor was the subject's first language. In all cases but one, the subject's first language was English; therefore it was not feasible to generalize based on that attribute. Similarly, only three of the 40 subjects (7.5%) prefer the search engine Bing to Google. Given such a small Bing cohort, generalizing on the

---

[6] The statistical power of this observation is uncertain, given the overrepresentation of people in their 40s within the sample. Further discussion on the generalizability of results from this sample population can be found at the end of Chapter IV.

[7] Note that, where subjects estimated their number of searches per day and the number of hours per day spent online as "7+", the value 8 was used in calculations.

basis of that preference is statistically insupportable.

For those correlations it was possible to calculate, the r-values fell far below the .95 level indicating statistical significance. Thus no clear association of efficiency with any of the above factors can be assumed, and based upon the results of this study, we cannot reject the null hypothesis.

Upon examination, it became clear that the first task (wildlife extinction) exhibited the widest range in answer speed. This finding is unsurprising, as getting used to the task and/or the format early in the exercise might well take a bit of extra time.

Further, as shown in Figure 6, the relationship between speed and accuracy was negligible. Participants generally achieved similar efficiency regardless of speed; hence the trend lines on these graphs demonstrate very little correlation.

Finding so little difference in the effects of the two interfaces was surprising; this outcome is promising from a usability standpoint if not from an experimental one. Admittedly, the node-link diagram as constituted here represents a very crude effort toward graphical representation of a document set. As noted in Chapter II, many more sophisticated and dynamic versions of this interface have been introduced, and yet it appears that few have been assessed against baseline data derived from text formats.

Ironically, in order to make this comparative assessment, the degree of semantic expression (i.e., dimensionality of data) had to be kept equal between formats; consequently this diagrammatic presentation sacrificed many of the connotative opportunities offered by graphical presentation. The main value of the diagram interface is that the system can embed many more semantic clues via color, shape, position, and other indicators. This is precisely the reason to employ a diagram interface, after all: to

express various attributes in an immediate and comprehensible fashion, without necessitating that users click a link and/or leave the web page.

As for assessments of this particular diagram format, no participant feedback was solicited or recorded. The investigator did not inquire about this aspect because ultimately this study was not conducted to assess this specific diagram interface, but rather to assess diagram interfaces generally. To that end, future investigators who seek to compare the effects of textual and graphical presentation may wish to improve not only on this study's diagram renderings, but also its experimental design and data collection.

Given this study's convenience sample, which was solicited via social media — specifically, a notice on the principal investigator's Facebook page and that of her husband — it must be stressed that generalizability of these results is limited. By comparison to the U.S. population as a whole,[8] this sample was skewed toward people with a higher level of education (92.5% with a bachelors degree or higher, vs. 27.9% for the population at large); also, participants' median age of 46 was significantly higher than that of the entire U.S. population (37.2%).

Most fundamentally, assessments of document relevance are by no means hard and fast. In a study by Al-Maskari, Sanderson and Clough (2008), "63% of documents judged relevant by our users matched official TREC judgments." The level of contradiction was highest among those who, like the participants in this study, were given a small set of documents to judge. "Therefore, in interactive IR studies which make use of TREC test collections ... care should be taken when comparing user effectiveness with system effectiveness" (p. 684).

---

[8] 2006-2010 American Community Survey 5-Year Estimates; www.census.gov.

V. CONCLUSION

In an era when billion-record data sets not only exist but also abound, users cannot begin to comprehend a large collection of documents just by reading or clicking one at a time. Thus graphical presentations, with their dimensional richness and information density, are becoming a necessity. Researchers such as those mentioned in Chapter II continue working to refine interfaces that help users orient themselves within a document set, so that they can perceive and understand relationships among the retrieved items.

Against that backdrop, this study sought to determine the degree (if any) to which a graphical interface would boost user efficiency in finding relevant documents. Experimental results were inconclusive, indicating that further study of this question is needed. One can posit that the diagrams are not a hindrance, precisely because the two formats generated no statistically significant difference in user efficiency.[9] In fact there are good reasons to continue testing and refining these interfaces. Using Bertin's retinal variables (color, shape, size, position, etc.) to encode various dimensions of the data (i.e., metadata such as file format, top-level domain, recency, or file size) would, in theory, give users a context within which to work more efficiently, in the same way that faceted displays or browsers can do.

Such a visual system might require a significant time investment for the user to

[9] This statement must be qualified: results from this population sample may not be generalizable, as discussed in Chapter IV.

become proficient, however, and more reliable empirical performance data should be collected in order to determine whether 1) efficiency gains exist, and 2) the gains are large enough to make that upfront time investment worthwhile. This calculation would depend, of course, on the quality of the display and the information architecture, which in turn depend on the quality of the information professional.

From a broader perspective, given the fundamental importance of efficient retrieval and identification of relevant documents, it behooves information professionals to portray document sets in more than one way, so as to address various cognitive styles of as many seekers as possible. Static diagrammatic portrayal constitutes one possible method; it may prove useful in situations where users do not have either the connective bandwidth or the computing power to avail themselves of interactive visualization applications. When practicable, interactive presentations can be excellent.

VII. APPENDICES

APPENDIX A: THE EXERCISE

Instructions

Text and Diagram Interfaces of the Three Tasks

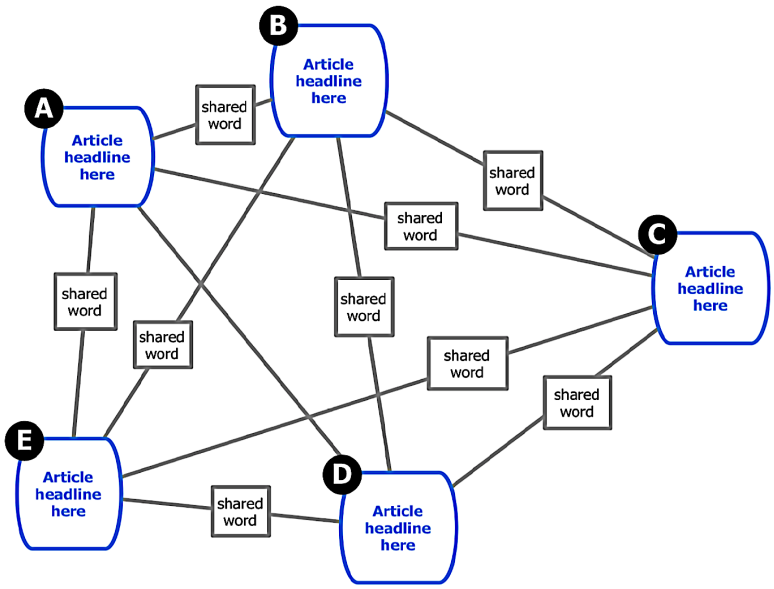
Questionnaire

Instructions:

Each of the following screens will include a topic. In each case, imagine that you have been assigned to write a term paper on that topic. Your task for each screen is to look over the articles presented there and select the three articles that would be most useful for your assignment. Check three boxes to indicate your selections.

Sometimes the group of articles will be presented as a list.

Sometimes the group of articles will be presented within a diagram that looks like this:



How to interpret the diagram: The articles' titles appear in the big nodes. When two nodes are connected by a line, that means those articles share at least one important word. The line labels tell you what the shared words are.

TOPIC: WILDLIFE EXTINCTION Write about the efforts of other (non-U.S.) countries to prevent the extinction of their native species. What have they done to halt the declines? Specify the country, the involved species, and steps taken to save the species.

A **Unique cameras discover Asian tigers disappearing faster than thought**
TA PHAYA NATIONAL PARK, Thailand -- A network of motion-sensitive cameras is revealing what conservationists have long feared: one of Asia's most magnificent beasts, the tiger, is even closer to extinction than previously thought.

B **British bumblebee extinct; other species likely to follow**
LONDON -- A British bumblebee is the latest species to become extinct here, and a handful of others will follow soon if the government does not act to save them, the World Wide Fund for Nature said Monday. Water voles, the high brown fritillary butterfly, pipistrelle bats, skylarks...

C **Aleutian Canada geese make recovery**
WASHINGTON -- The government Friday will propose removing the Aleutian Canada geese from the threatened species list Friday, citing a remarkable recovery after foxes introduced by fur farmers pushed the birds to the brink of extinction.

D **Rare reptiles seized in Arizona**
An international smuggling ring involving 300 rare and endangered animals -- including giant Komodo dragons and other exotic creatures -- has been broken up with arrests in Arizona, San Francisco and Mexico City.

E **Mink set free to wreak havoc among wildlife**
LONDON -- Thousands of mink released from a fur farm by animal liberation activists were enjoying their first taste of freedom -- and of the local wildlife -- on Monday. The Animal Liberation Front confirmed that activists smashed cages and cut wire at a fur farm near Ringwood...

F **WWF: 10 percent of tree species under extinction threat**
GENEVA -- One-tenth of the world's known species of trees are in danger of extinction and fewer than one in four of those benefit from protection efforts, the World Wide Fund for Nature and other groups said in a study published Tuesday.

G **UN agency launches computer program to prevent extinction of domestic animals**
ROME -- Siberian cattle that tolerate temperatures of minus 60 degrees Celsius (-140 F), Mozambique pigs resistant to African Swine Fever, and Kazakhstan dromedary capable of producing nourishing milk on little water or food. These are a few of about 1,700 domestic animal...

H **Western ousted as head of Kenya Wildlife Service**
NAIROBI, Kenya -- David Western was ousted as director of the Kenya Wildlife Service for the second time in three months Friday and temporarily replaced by one of his deputies. "His services have been terminated," said David Kioko, "and I have been told to be in charge..."

I **Poachers kill two rare mountain gorillas in eastern Congo**
NAIROBI, Kenya -- Poachers accidentally shot dead two rare mountain gorillas in eastern Congo early this month, the African Wildlife Foundation said Friday. The gorillas -- a juvenile male named Birori and a 3-year-old female -- were killed Sept. 3 in Virunga National Park...

A. ☐ Unique cameras discover Asian tigers disappearing faster than thought

B. ☐ British bumblebee extinct; other species likely to follow

C. ☐ Aleutian geese make recovery

D. ☐ Rare reptiles seized in Arizona

E. ☐ Mink set free to wreak havoc among wildlife

F. ☐ WWF: 10 percent of tree species under extinction threat

G. ☐ UN agency launches computer program to prevent extinction of domestic animals

H. ☐ Western ousted as head of Kenya Wildlife Service

I. ☐ Poachers kill two rare mountain gorillas in eastern Congo

Next

TOPIC: INTERNATIONAL ART CRIME Write about fraud or embezzlement in the international art trade. Objects include paintings, jewelry, sculptures and any other valuable works of art. You must include specific instances; avoid general statements.

A Police recover stolen Mondrian, arrest suspects
MIDDELBURG, Netherlands -- Dutch police recovered a painting by abstract pioneer Piet Mondrian and arrested three people suspected of snatching it from a museum, officials said Friday.

B Some of expanded Uffizi in Florence to open in December
FLORENCE, Italy -- The first of new gallery rooms in the Uffizi museum's expansion will open to the public on Dec. 16, Italy's culture minister said Saturday. Minister Veltroni, visiting the musuem, did not say which rooms or which works of art would be able to be seen on that day.

C Businessman killed in Siberian town
MOSCOW -- A Russian businessman was killed on his way to work Wednesday by an unidentified gunman in what appeared to be an organized crime killing, according to news reports. Yuri Zinin, the head of Sibneft Company was shot in the back...

D Cartoon theft no laughing matter
NEW YORK -- Stolen animation cells of the Grinch showed up for auction on the eBay Web site and two teen-age clerks at an art gallery were arrested Friday, police said.

E International police seek to stymie illegal art trafficking
BUDAPEST, Hungary -- Police officers from 23 countries met in Budapest to seek more efficient ways to combat illegal traffic in art treasures, a police official said Thursday. Held under the auspices of Interpol, the International Police Organization, the three-day closed meeting allowed the officers to make personal contact...

F Spirited bidders snap up art seized from failed finance firms
BANGKOK, Thailand -- Paintings and other art objects once displayed in elegant offices were snapped up this weekend in Thailand at an auction of assets of collapsed financial firms, Thai media reported Sunday. The auction of more than 1,000 items, which ended Sunday...

G New Orleans cemetery trial begins
NEW ORLEANS -- Two prominent antiques dealers and one of their customers are on trial, accused of putting in orders for thefts of sculptures, urns and benches from ornate above-ground tombs.

H Islamic art and soul
In a Muslim society, the mosque is the hub of the "Downtown" community around which all the hustle and bustle of activities takes place. There is a certain aura about it which makes it stand out from among the rest of the buildings.

I Police seek two men for political bombing in Hungary
BUDAPEST, Hungary -- Police said Wednesday they are seeking two men suspected in this week's bomb attack against Hungary's leading parliamentary party. Composite sketches have been issued of the two men, whom police suspect of placing the bomb on a balcony...

A. ☐ Police recover stolen Mondrian, arrest suspects

B. ☐ Some of expanded Uffizi in Florence to open in December

C. ☐ Businessman killed in Siberian town

D. ☐ Cartoon theft no laughing matter

E. ☐ International police seek to stymie illegal art trafficking

F. ☐ Spirited bidders snap up art seized from failed finance firms

G. ☐ New Orleans cemetery trial begins

H. ☐ Islamic art and soul

I. ☐ Police seek two men for political bombing in Hungary

Next

TOPIC: BLACK BEAR ATTACKS Discuss the frequency of vicious black bear attacks worldwide. Include input from the scientific community on possible causes of vicious attacks, plus details about steps taken by wildlife officials to control such attacks.

A El Nino brings predatory feast; population level unlikely to spike
It's about 5:30 a.m. on Memorial Day. The driver heads north on Arizona 87 toward Payson. He is beyond the Saguaro Lake turnoff and enjoying the cool wind coming through the window.

B Boy attacked by bear at scout camp
SHELL LAKE, Wis. -- A black bear attacked a 14-year-old boy at a Boy Scout camp, dragging him into the woods and injuring him before his father drove the animal off by pelting him with rocks and wood. Matthew Murphy, 14, had bite wounds to his head and shoulder area...

C 53 killed in three days of factional fighting in Somalia
MOGADISHU, Somalia -- The death toll in three days of factional fighting for control of a strategic town northwest of Mogadishu rose to 56 Friday. At least 18 people were reported killed...

D Bears become uncomfortably close neighbors in northern New Jersey
STILLWATER, N.J. -- Danette Conroy was upstairs cleaning kitchen cabinets and her year-old daughter Kayleigh was asleep when the black bear somehow got into the ground floor of their new home. They never heard a thing until the crash of breaking glass.

E Two bear attacks on Sakhalin Island
MOSCOW -- The Hunting Department warned residents on Sakhalin Island to be careful while walking in forests after two people were mauled by bears. In one incident, a bear mauled a hunter, who was taken to the intensive care unit of a hospital...

F Bear kills two campers in B.C. park, Canada
OTTAWA -- A woman and a man were killed by a black bear in a remote northern park in Canada's British Columbia (B.C.) province, a report said today. The woman's 13-year-old son and a man were also attacked and suffered injuries. The attack occurred Thursday night...

G This "Dracula" kills cattle and deer
CLUJ, Romania -- A giant black bear, whom peasants have nicknamed "Dracula," is creating anxiety in western Romania because of his taste for cattle and sheep, an official said Thursday.

H Hunters pursue bear that mauled jogger to death
HELSINKI, Finland -- Hunters and border officials were searching for a bear Thursday that killed a man while he was jogging in southeastern Finland near the Russian frontier, police said. A woman sounded the alarm after her 33-year-old husband did not return home from jogging...

I Tina tucks in her little honey to make sure she is safe
Twenty-eight-year-old sun bear Tina cuddles her healthy newborn, Juma, at Zoo Negara yesterday. Juma was named after the day and month (in Bahasa Malaysia) she was born. She was delivered on Friday, March 27, and is the first surviving sun bear cub at Zoo Negara in 21 years.

A. ☐ El Nino brings predatory feast; population level unlikely to spike

B. ☐ Boy attacked by bear at scout camp

C. ☐ 53 killed in three days of factional fighting in Somalia

D. ☐ Bears become uncomfortably close neighbors in northern New Jersey

E. ☐ Two bear attacks on Sakhalin Island

F. ☐ Bear kills two campers in B.C. park, Canada

G. ☐ This "Dracula" kills cattle and deer

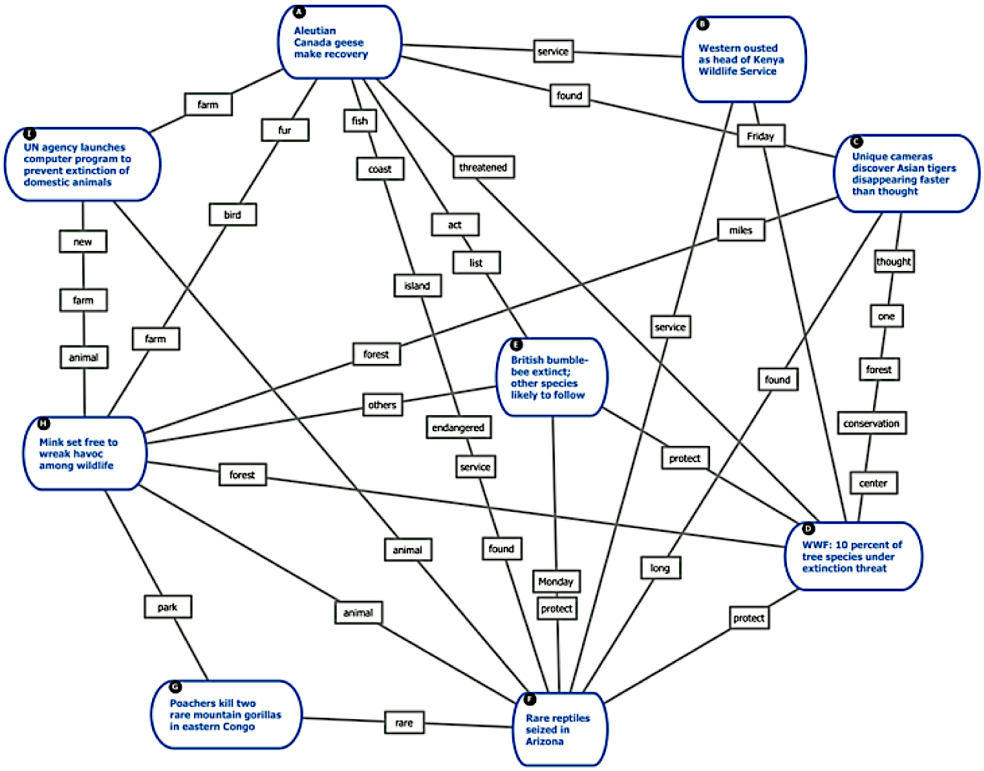
H. ☐ Hunters pursue bear that mauled jogger to death

I. ☐ Tina tucks in her little honey to make sure she is safe

Next

TOPIC: WILDLIFE EXTINCTION

Write about the efforts of other (non-U.S.) countries to prevent the extinction of their native species. What have they done to halt the declines? Specify the country, the involved species, and steps taken to save the species.



A. ☐ Aleutian Canada geese make recovery

B. ☐ Western ousted as head of Kenya Wildlife Service

C. ☐ Unique cameras discover Asian tigers disappearing faster than thought

D. ☐ WWF: 10 percent of tree species under extinction threat

E. ☐ British bumblebee extinct; other species likely to follow

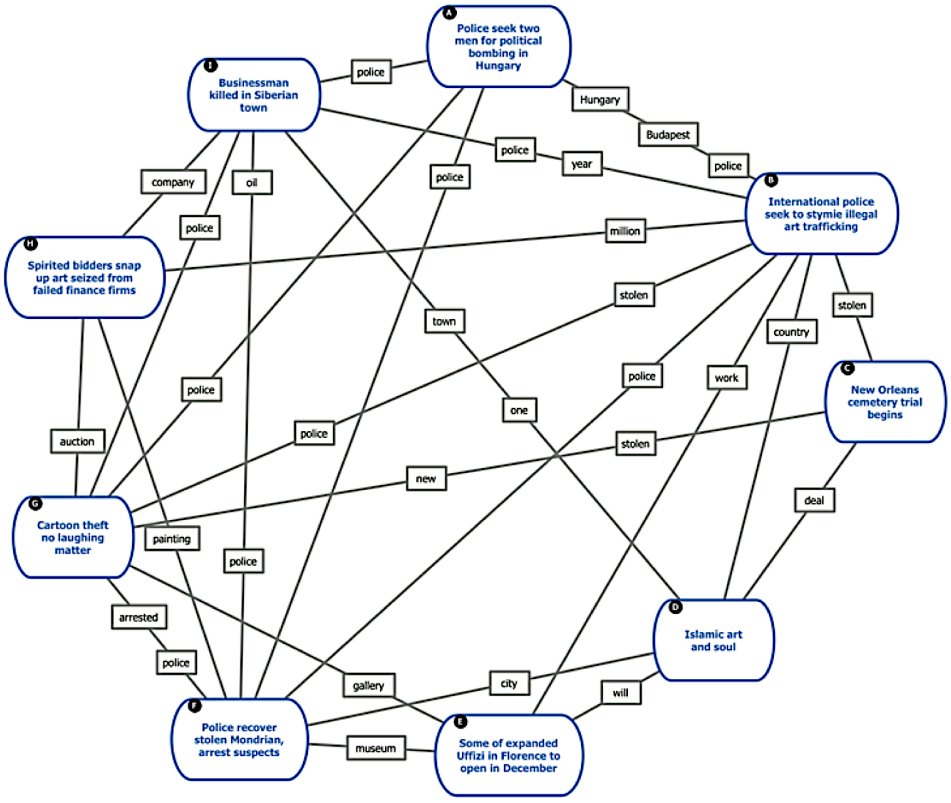
F. ☐ Rare reptiles seized in Arizona

G. ☐ Poachers kill two rare mountain gorillas in eastern Congo

H. ☐ Mink set free to wreak havoc among wildlife

I. ☐ UN agency launches computer program to prevent extinction of domestic animals

TOPIC: INTERNATIONAL ART CRIME Write about fraud or embezzlement in the international art trade. Objects include paintings, jewelry, sculptures and any other valuable works of art. You must include specific instances; avoid general statements.



A. ☐ Police seek two men for political bombing in Hungary

B. ☐ International police seek to stymie illegal art trafficking

C. ☐ New Orleans cemetery trial begins

D. ☐ Islamic art and soul

E. ☐ Some of expanded Uffizi in Florence to open in December

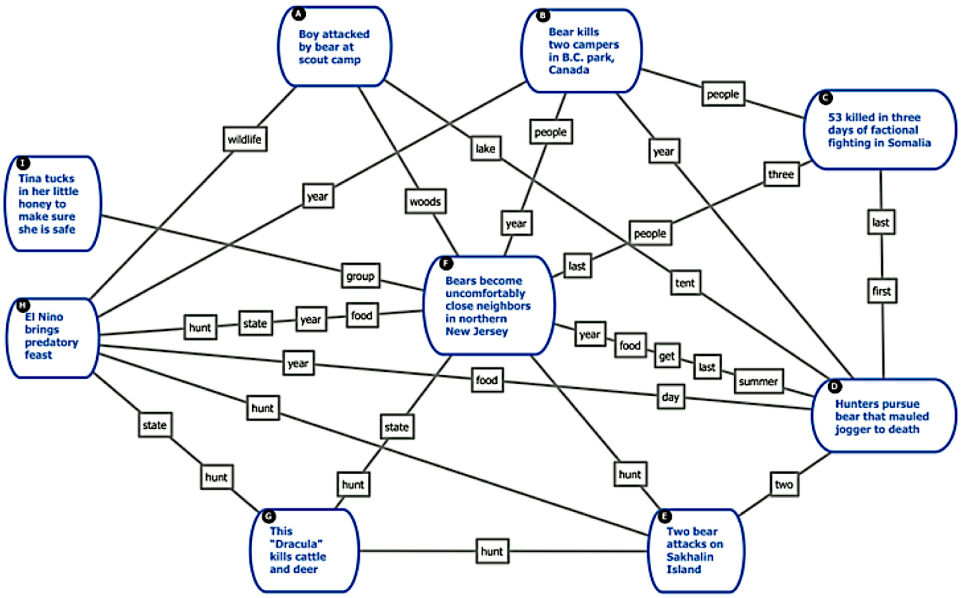
F. ☐ Police recover stolen Mondrian, arrest suspects

G. ☐ Cartoon theft no laughing matter

H. ☐ Spirited bidders snap up art seized from failed finance firms

I. ☐ Businessman killed in Siberian town

TOPIC: BLACK BEAR ATTACKS Discuss the frequency of vicious black bear attacks worldwide. Include input from the scientific community on possible causes of vicious attacks, plus details about steps taken by wildlife officials to control such attacks.



A. ☐ Boy attacked by bear at scout camp

B. ☐ Bear kills two campers in B.C. park, Canada

C. ☐ 53 killed in three days of factional fighting in Somalia

D. ☐ Hunters pursue bear that mauled jogger to death

E. ☐ Two bear attacks on Sakhalin Island

F. ☐ Bears become uncomfortably close neighbors in northern New Jersey

G. ☐ This "Dracula" kills cattle and deer

H. ☐ El Nino brings predatory feast

I. ☐ Tina tucks in her little honey to make sure she is safe

Next

Questionnaire

E-mail address:*

Date of birth:

Please provide some additional information.

I am:

My first language is:

My education level is:

When I use a computer I usually:

My preferred search engine is:

Recent computer use

	0	1	2	3	4	5	6	7+
Estimated number of Web searches yesterday:	○	○	○	○	○	○	○	○
Estimated number of hours per day spent online in non-game activity:	○	○	○	○	○	○	○	○

Thanks again for your participation in this study.

APPENDIX B: RESULTS

Results for Each Version

Questionnaire Responses from Adobe Forms Central

Selecting Useful Documents A

Times taken
10

Average Time(HH:MM:SS)
00:15:29

Report Summary

Question	% Correct	Correct	Incorrect	Responses
1. TOPIC: WILDLIFE EXTINCTION Write about t...	0%	--	10	10
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	0%	--	10	10
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	0%	--	10	10

Average time taken per question

Question	Average Question Time(secs)
1. TOPIC: WILDLIFE EXTINCTION Write about t...	6 mins 11 secs
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	4 mins 26 secs
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	4 mins 16 secs

Selecting Useful Documents B

Times taken
10

Average Time(HH:MM:SS)
00:15:23

Report Summary

Question	% Correct	Correct	Incorrect	Responses
1. TOPIC: WILDLIFE EXTINCTION Write about the eff...	0%	--	10	10
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	10%	1	9	10
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	10%	1	9	10

Average time taken per question

Question	Average Question Time(secs)
1. TOPIC: WILDLIFE EXTINCTION Write about the eff...	5 mins 14 secs
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	3 mins 43 secs
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	5 mins 4 secs

Selecting Useful Documents C

Times taken
12

Average Time(HH:MM:SS)
00:12:37

Report Summary

Question	% Correct	Correct	Incorrect	Responses
1. TOPIC: WILDLIFE EXTINCTION Write about t...	0%	--	12	12
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	8%	1	11	12
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	0%	--	12	12

Average time taken per question

Question	Average Question Time(secs)
1. TOPIC: WILDLIFE EXTINCTION Write about t...	4 mins 30 secs
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	3 mins 17 secs
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	3 mins 34 secs

Selecting Useful Documents D

Times taken
10

Average Time(HH:MM:SS)
00:12:01

Report Summary

Question	% Correct	Correct	Incorrect	Responses
1. TOPIC: WILDLIFE EXTINCTION Write about the eff...	0%	--	10	10
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	10%	1	9	10
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	10%	1	9	10

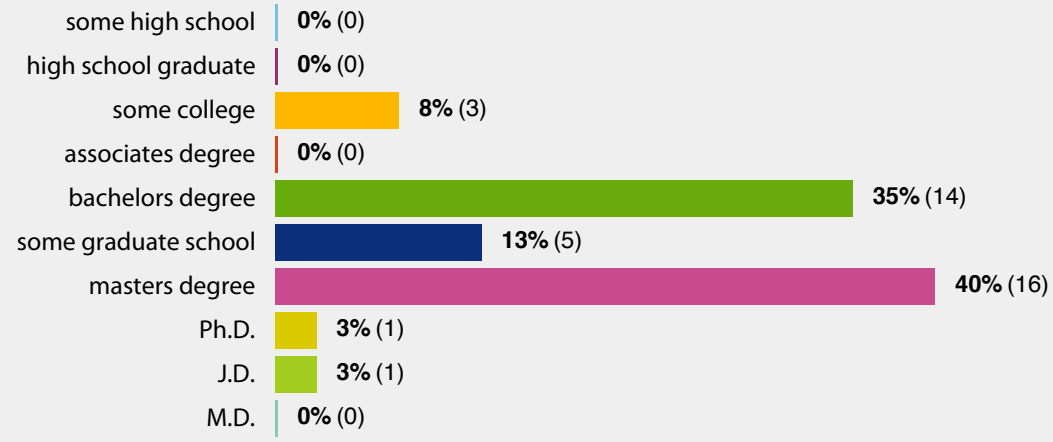
Average time taken per question

Question	Average Question Time(secs)
1. TOPIC: WILDLIFE EXTINCTION Write about the eff...	3 mins 37 secs
2. TOPIC: INTERNATIONAL ART CRIME Write abo...	4 mins 22 secs
3. TOPIC: BLACK BEAR ATTACKS Discuss the fr...	2 mins 47 secs

I am:

* Total Responses: 40, 100% of submissions

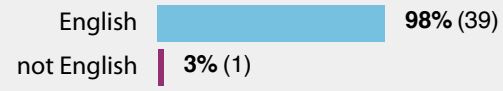
My education level is:



1 additional choice not shown

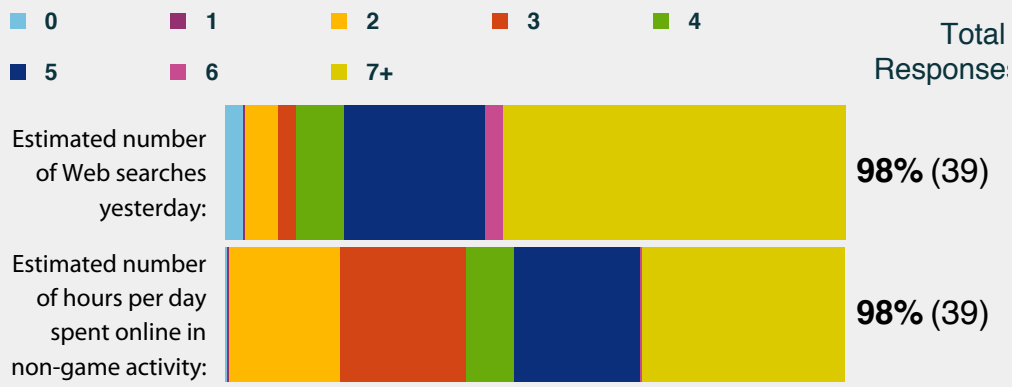
* Total Responses: 40, 100% of submissions

My first language is:



* Total Responses: 40, 100% of submissions

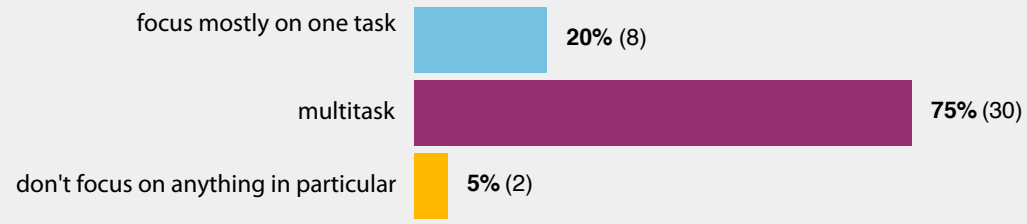
Recent computer use



My preferred search engine is:

* Total Responses: 39, 98% of submissions

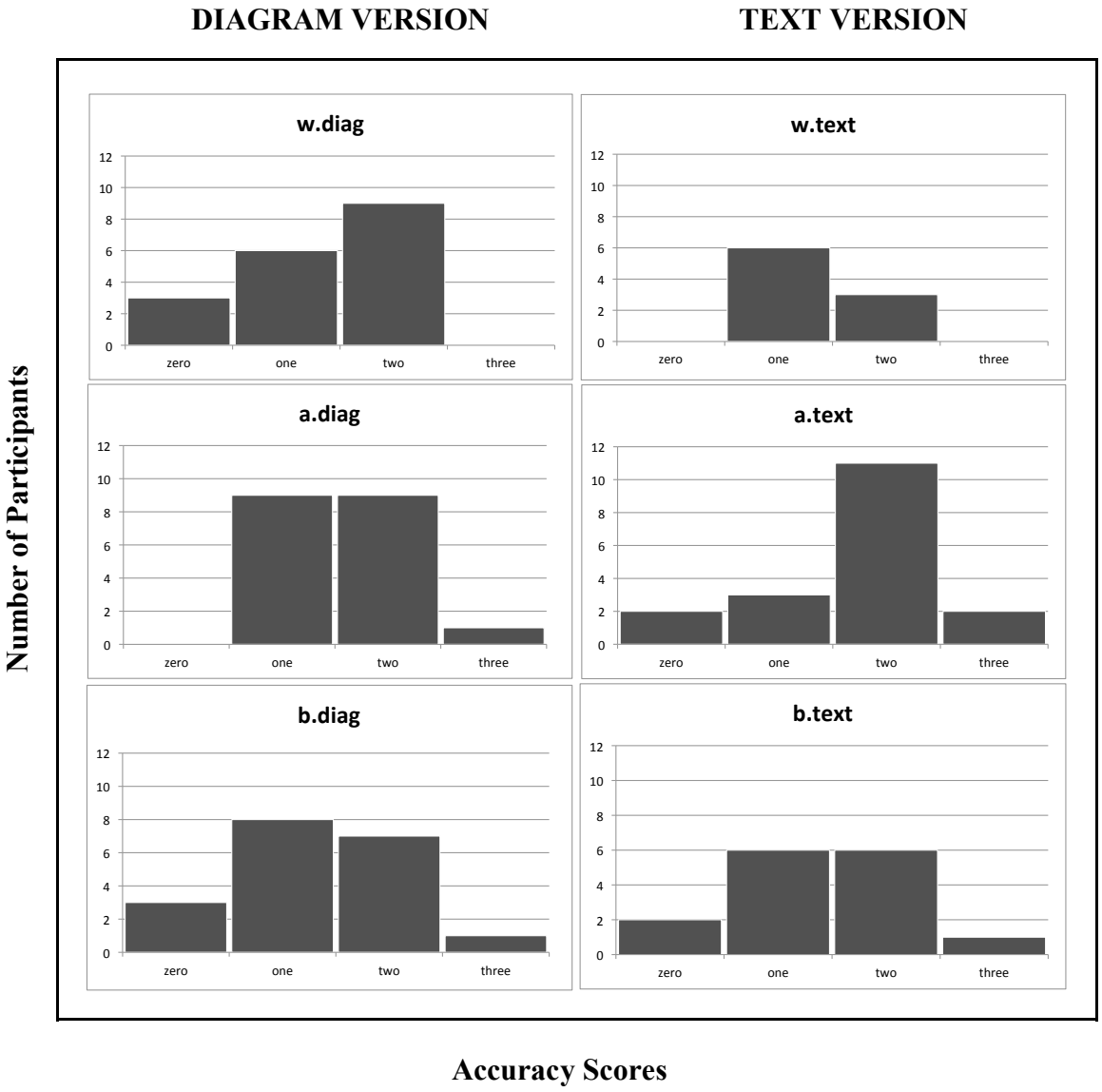
When I use a computer I usually:



* Total Responses: 40, 100% of submissions

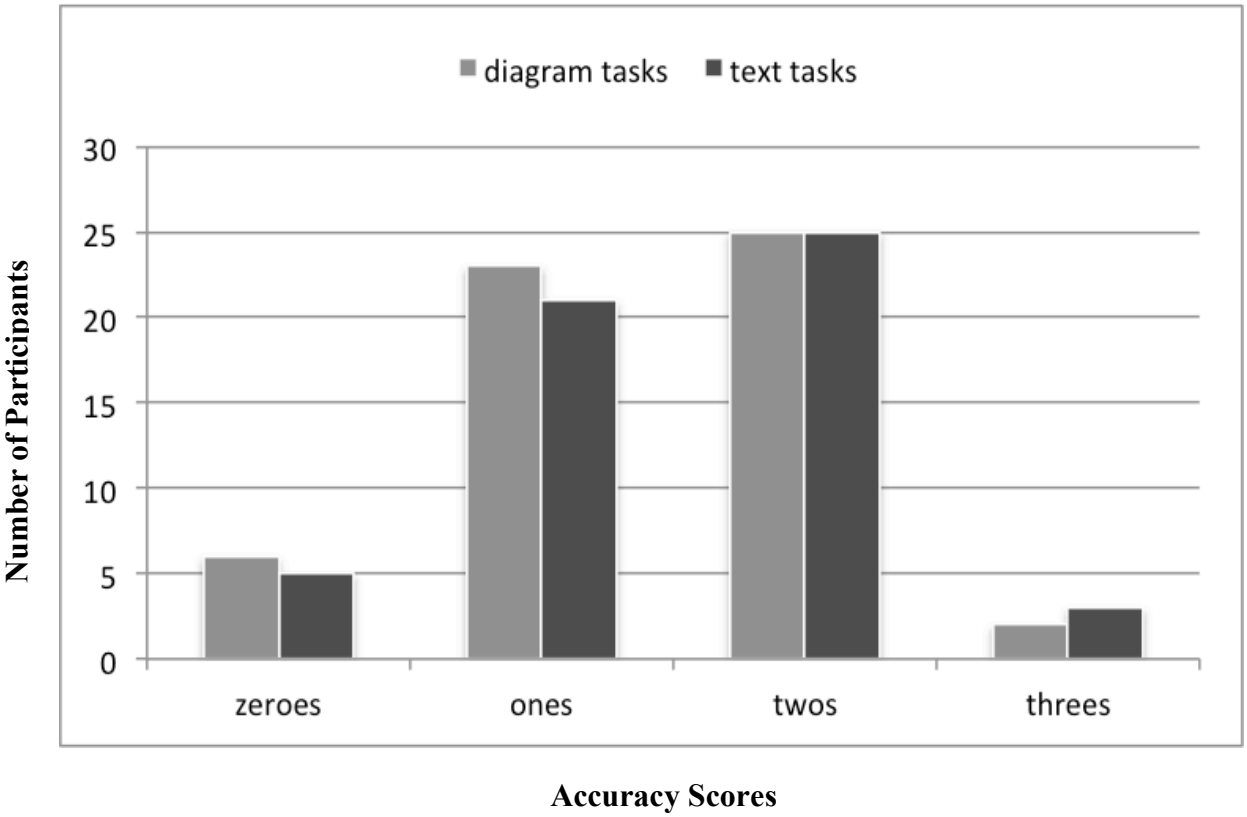
APPENDIX C: ADDITIONAL GRAPHS OF RESULTS

Documents Correctly Selected, By Question And Format



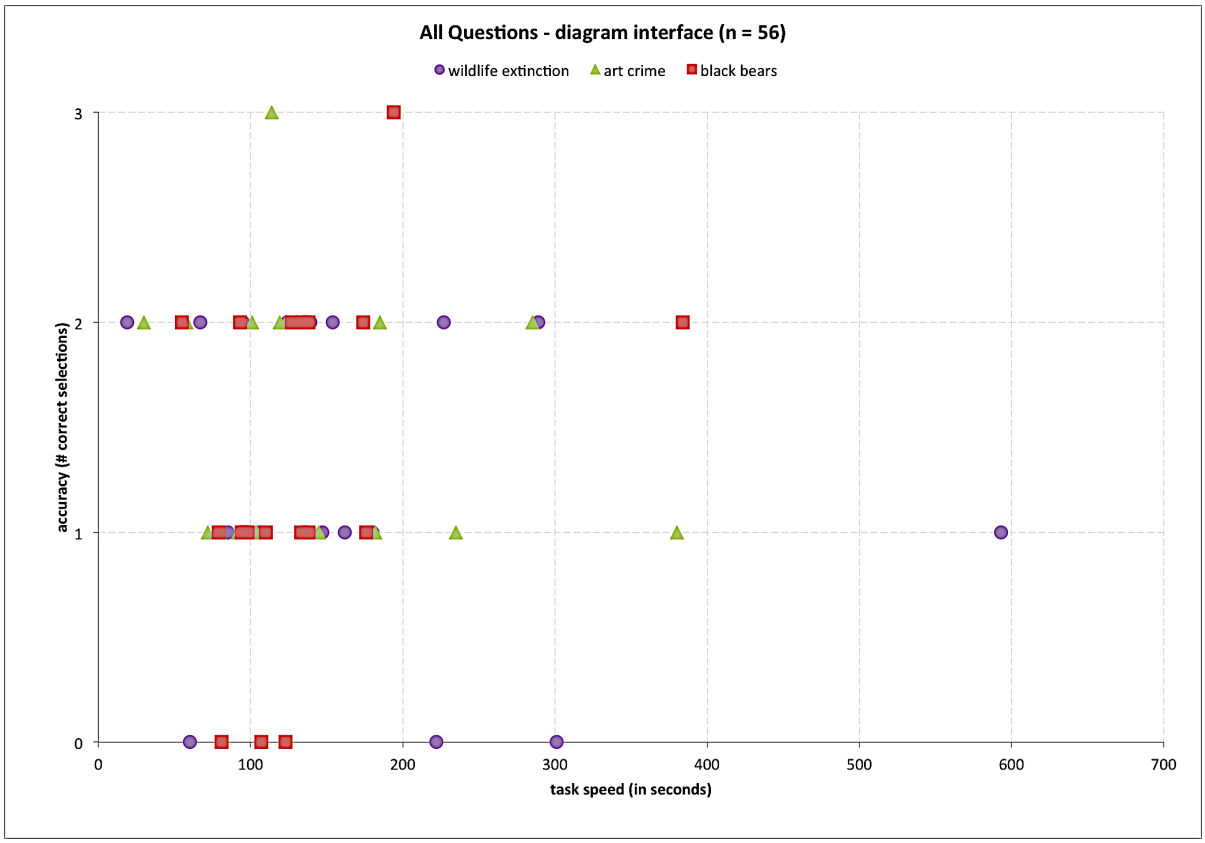
The series of bar graphs above shows, for each question and in each format, the distribution of each question interface's accuracy scores. A participant's accuracy score ("zero," "one," "two," or "three") reflects the number of relevant documents correctly

selected by a participant. On the chart labels, the wildlife extinction topic is signified by *w*, the art crime topic is signified by *a*, and the black bear attacks topic is signified by *b*. The left column represents the topic's diagram version; the right represents the text version. The topics were presented in the order shown (from top to bottom).



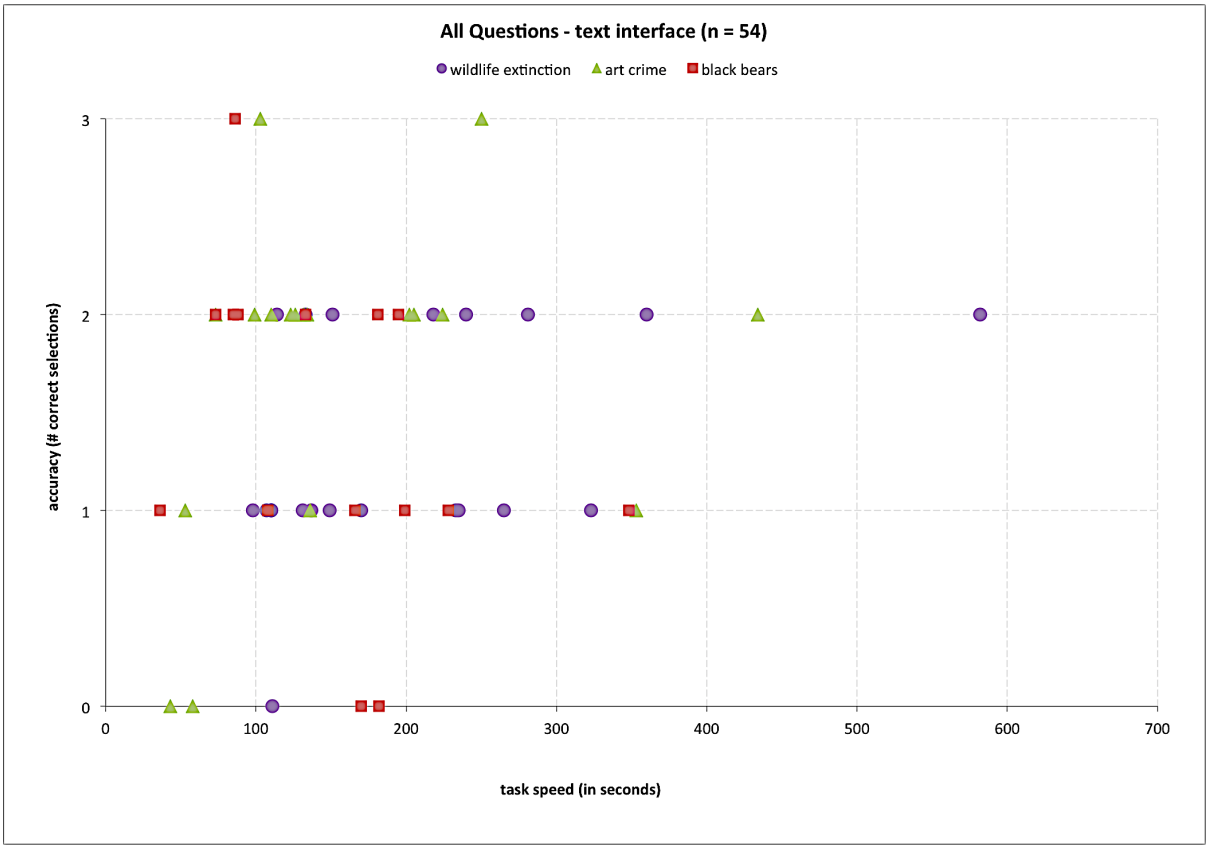
This clustered column chart simply aggregates all the data shown in the previous six charts. Despite their obvious differences, the two interfaces elicited remarkably similar outcomes.

Accuracy Vs. Speed, All Questions, Diagram Format



These scatterplot graphs show accuracy scores and speed (in seconds) for each question's diagram version. Each point is a participant's performance on a particular question. Circles indicate the wildlife topic; triangles, the art crime topic; and squares, the black bears topic. Because accuracy scores are small integers, putting them on the vertical axis has the effect of visually stratifying the data somewhat, but the clustering across the horizontal time axis is instructive. For the diagram interface, almost every instance took place in less than 400 seconds — except for one outlier, on the extinction task. That instance took roughly 600 seconds (10 min.). It ended with the participant only getting one out of three accuracy points. Coincidentally, one participant with the text

Accuracy Vs. Speed, All Questions, Text Format



version of that question also took almost 10 minutes to answer the question. That participant scored two out of three.

Additional Observations:

- The text format elicited three perfect ("3") accuracy scores; the diagram format got only two.
- All five of the zero-accuracy text instances were executed within 190 seconds, while two of the six zero-accuracy diagram instances took longer (220 and 300 seconds). The extra time was to no avail.
- 47 of the 56 diagram instances finished in less than 200 seconds, while only 34 of the 54 text instances did.

REFERENCES

Al-Maskari, A., Sanderson, M., & Clough, P. (2008) Relevance judgments between TREC and non-TREC assessors. *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* 683–684. doi: 10.1145/1390334.1390450

Amar, B., Day, J., Godfrey, J., & Plaue, C. (2004) Dart: A dynamic article research tool. doi: 10.1.1.199.8641 Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.199.8641&rep=rep1&type=pdf

Amar, R., Eagan, J., and Stasko, J. (2005). Low-level components of analytic activity in information visualization. *IEEE Symposium on Information Visualization,* 111–117.

Andrews, K., Kienreich, W., Sabol, V., Becker, J., Droschl, G., Kappe, F., ... Tochtermann, K. (2002). The InfoSky visual explorer: Exploiting hierarchical structure and document similarities. *Information Visualization, 1,* 166–181. doi:10.1057/palgrave.ivs.9500023

Bertin, J. (1983). *Semiology of graphics.* Madison, WI: University of Wisconsin Press.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes, 45*(1), 111–139.

Boechler, P.M. (2001). How spatial is hyperspace? Interacting with hypertext documents: Cognitive processes and concepts. *Cyberpsychology & Behavior, 4*(1), 23–46.

Chen, C. (2010). Information visualization. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 387–403. doi: 10.1002/wics.89

Chen, L. (2005). The topological approach to perceptual organization. *Visual Cognition, 12*(4), 553–637. doi: 10.1080/13506280444000256

Chuklin, A., & Serdyukov, P. (2012) Potential good abandonment prediction [Abstract]. *Proceedings of the 21st International Conference Companion on World Wide Web,* 485–486. doi: 10.1145/2187980.2188089.

Clarkson, E. C., Desai, K., & Foley, J. D. (2009). ResultMaps: Visualization for search interfaces. *IEEE Transactions on Visualization and Computer Graphics, 15*(6), 1057–1064.

Collins, C., Carpendale, S., & Penn, G. (2009). DocuBurst: Visualizing document content using language structure. *Computer Graphics Forum, 28*(3), 1039–1046.

Cribbin, T. (2010) Visualising the structure of document search results: A comparison of graph theoretic approaches. *Information Visualization, 9*(2), 83–97.

Gretarsson, B., O'Donovan, J., Bostandjiev, S., Höllerer, T., Asuncion, A., Newman, D., and Smyth, P. (2012). TopicNets: Visual analysis of large text corpora with topic modeling. *ACM Transactions on Intelligent Systems and Technology, 3*(2). doi: 10.1145/ 2089094.2089099

Hemmje, M., Kunkel, C., & Willett, A. (1994). Lyberworld — A visualization user interface supporting fulltext retrieval. *Proceedings of the 17th Annual International ACM SIGIR conference on Research and development in information retrieval,* 249–259.

Ingram, S., Munzner, T., Irvine, V., Tory, M., Bergner, S., & Moller, T. (2010). DimStiller: Workflows for dimensional analysis and reduction. *Proceedings of the 5th IEEE Symposium on Visual Analytics in Science and Technology (VAST),* 3–10. doi: 10.1109/VAST.2010.5652392

Joachims, T., Granka, L., Pan, B., Hembrooke, H. & Gay, G. (2005) Accurately interpreting clickthrough data as implicit feedback. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* 154–161. doi: 10.1145/1076034.1076063

Julien, C., Leide, J. E., & Bouthillier, F. (2008). Controlled user evaluations of information visualization interfaces for text retrieval: Literature review and meta-analysis. *Journal of the American Society for Information Science & Technology, 59*(6), 1012–1024. doi:10.1002/asi.20786

Kaski, S., Honkela, T., Lagus, K., & Kohonen, T. (1998). WEBSOM — Self-organizing maps of document collections. *Neurocomputing, 21,* 101–117.

Kim, K., Ko, S., Elmqvist, N., & Ebert, D. S. (2011). WordBridge: Using composite tag clouds in node-link diagrams for visualizing content and relations in text corpora. *44th Hawaii International Conference on System Sciences (HICSS),* 1–8.

Kobsa, A. (2001). An empirical comparison of three commercial information visualization systems. *Proceedings of InfoVis 2001, IEEE Symposium on Information Visualization,* 123–130.

Koshman, S. (2004). Comparing usability between a visualization and text-based system for information retrieval. *Journal of Documentation, 60*(5), 565–580. doi: 10.1108/00220410410560627

Lagus, K., Kaski, S., & Kohonen, T. (2004). Mining massive document collections by the WEBSOM method. *Information Sciences, 163,* 135–156. doi: 10.1016/j.ins.2003.03.017

Landauer, T. K., Laham, D., & Derr, M. (2004). From paragraph to graph: Latent semantic analysis for information visualization. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 5214–5219. doi:10.1073/pnas.0400341101

Lee, B., Plaisant, C., Parr, C. S., Fekete, J., & Henry, N. (2006). Task taxonomy for graph visualization. *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors:*

Novel Evaluation Methods for Information Visualization, 1–5. doi: http://doi.acm.org/ 10.1145/1168149.1168168

Lehmann, S., Schwanecke, U., & Dorner, R. (2010). Interactive visualization for opportunistic exploration of large document collections. *Information Systems, 35,* 260–269.

Nielsen, J. (2006, April 17). F-shaped pattern for reading web content. Retrieved from http://www.useit.com.libproxy.lib.unc.edu/ alertbox/reading_pattern.html

O'Brien, M. & Keane, M. T. Modeling result-list searching in the world wide web: The role of relevance topologies and trust bias. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society,* 1881-1886.

Rogowitz, B. E., Treinish, L. A., & Bryson, S. (1996). How not to lie with visualization. *Computers in Physics, 10*(3), 268–273.

Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science, 21*(5), 641–644. doi:10.1177/ 0956797610368811

Shneiderman, B. (2008). Extreme visualization: Squeezing a billion records into a million pixels. *Proceedings of the ACM SIGMOD International Conference on Management of Data,* 3–12. doi: 10.1145/1376616.1376618

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. *Proceedings of the IEEE Symposium on Visual Languages,* 336–343. doi: 10.1109/VL.1996.545307

Van Ham, F., Wattenberg, M., & Viegas, F. B. (2009). Mapping text with phrase nets. *IEEE Transactions on Visualization and Computer Graphics, 15*(6), 1169–1176.

Vegas, J., Crestani, F., & de la Fuente, P. (2007). Context representation for web search results. *Journal of Information Science, 33*(1), 77–94.

Venna, J., Peltonen, J., Nybo, K., Aidos, H., & Kaski, S. (2010). Information retrieval perspective to nonlinear dimensionality. *Journal of Machine Learning Research, 11,* 451–490.

Woodruff, A. & Plaunt, C. (1994). GIPSY: Automated geographic indexing of text documents. *Journal of the American Society for Information Science, 45*(9), 645–655.

Zhang, J., & Nguyen, T. (2005). WebStar: A visualization model for hyperlink structures. *Information Processing and Management, 41,* 1003–1018. doi: 10.1016/ j.ipm.2004.03.005

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Oxford, England: Addison-Wesley Press.