THE ROLE OF PERCEPTUAL-MOTOR AVAILABILITY IN THE INTEGRATION

OF INFORMATION ACROSS GRAPH AND TEXT

By

JASON F. RUBINSTEIN

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Eileen Kowler

And approved by

____

____

____

New Brunswick, New Jersey

October 2017

ABSTRACT OF THE THESIS

The role of perceptual-motor availability in the integration of information across graph and text

By Jason F. Rubinstein

Thesis Director:

Eileen Kowler

Gathering and integrating information from spatially distinct and qualitatively different sources requires decisions about when and where to sample information. An example is reading graphs and accompanying text to arrive at a coherent interpretation. Eye movements were recorded while subjects viewed bar graphs depicting information about two attributes of two fictitious products, along with descriptive text, after which subjects indicated which product they preferred. Three perceptual-motor configurations were tested: (1) Simultaneous: Graph and text were displayed adjacent to each other; (2) Button-Press: Graph and text appeared sequentially, with the appearance of each triggered by a button press; and (3) Eye-Contingent: Graph and text appeared sequentially, in the same spatial locations as in the Simultaneous condition, with the appearance of each triggered by a saccade into the region.

Shifts of gaze between graph and text occurred about twice as often in the Simultaneous condition than in either Button-Press or Eye-Contingent conditions. The

rate of shifts in the Button-Press and Eye-Contingent conditions did not differ, showing that the relevant factor was not motor effort (saccade vs. button press), but rather the sequential vs. simultaneous aspects of the presentations. Conditions did not differ either in trial duration or in the proportion of time spent in the graph or text.

In the Simultaneous condition, most trials began with relatively long inspections of the graph and the text, with a preference to view the text first. The initial inspections of the graph or text were then followed by one or two relatively brief visits to each. Patterns were different for the other two perceptual-motor conditions in that the initial inspections of graph or text were typically longer and the subsequent visits to each were rare. Analyses of fixated locations showed a preference to use the more frequent visits in the Simultaneous condition to re-examine previously seen material rather than to look at new material. The simultaneous availability of graph and text did not necessarily encourage a strategy of inspecting each region in segments but spread out the reviewing time across subsequent visits to graph and text.

These results show that the strategies used to integrate information across graphs and text depend on more than just the information content, and are influenced by even relatively minor variations of perceptual availability. The extra steps involved in conditions aside from the Simultaneous condition may discourage saccades as a part of an overall strategy of conserving neural costs or cognitive load. The relative perceptual availability of information in different spatial regions must be taken into account when developing models of cognitive strategies on the basis of observed saccadic patterns.

ACKNOWELDGEMENTS

I would like to thank my advisor, Eileen Kowler, for her extremely helpful guidance, training, and insightful discussion during the past several years. Perhaps most importantly, I am so very grateful you decided to invite me into your lab as an undergraduate REU participant that summer 4 years ago, setting me on this path.

I would also like to thank my committee members Melchi Michel and Pernille Hemmer for their useful comments and discussion.

Thank you to Cordelia Aitkin for the great help with generation of the stimuli for pilot version of this experiment as well as conversations regarding the design. I also want to thank present and past lab members Jie Wang and Elsie Lee for helpful discussion, and especially for Elsie's continuation of graph-related research in this lab and elsewhere.

Finally, thank you to my family for your support through the past few years. I couldn't have done it without my mother (Karen), father (Ron), brother (Adam) and dog (Mocha)! Thank you for being there for me every step of the way, from words of encouragement to delicious family dinners. I love you all.

The pilot experiment for this thesis was supported by the NSF REU Site Perceptual Science and Technology (Award Number: 1062735) with Matthew Stone as the PI.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

1. Introduction

Visual-motor tasks are an essential part of moment-to-moment activities. With these activities come various benefits, but also costs, which may be taken into account when deciding which actions to take to accomplish the task at hand. These include physical costs, such as the motor cost of executing an action, but also various types of mental costs related to processing demands, decision-making or memory storage or retrieval.

Information-gathering tasks, in particular, may be subject to a wide variety of mental and physical costs. These costs must be taken into consideration when making decisions about where to direct gaze or motor action. This thesis examines how varying the perceptual and motor costs of accessing information in the form of a graph and accompanying passage of text can influence the strategies of observers as they move towards making a decision about the information. Before describing the present experiment, I will briefly review several examples in which mental and motor costs are taken into account when choosing actions.

1.1. Background

One of the most well-known examples in which costs need to be taken into account is foraging. When foraging for food, animals take the cost of searching a patch (time and energy) into account when deciding where to search next, taking into consideration both the amount energy obtained at a particular patch and the costs of moving to a different location (Stephens, Couzin, & Giraldeau, 2012). These costs not only include the motor costs and time involved, but also the costs of unspecified cortical computation, which has been estimated to be quite high during demanding tasks as a result of an increased number of metabolically costly neural spikes (Lennie, 2003). Numerous search tasks have been modeled on foraging behavior. For example, Wolfe (2013) and Constantine & Daw (2015) have adapted foraging situations as frameworks for visual search and decision-making experiments, both utilizing marginal value theory to provide descriptions of human behavior during patch foraging.

In foraging, there are immediate costs to particular actions. The role of costs of mental computations in choosing an action has also been studied. Kool, McGuire, Rosen, & Botvinick (2010) gave participants a choice of two task sets. The alternatives were indicated by two decks of cards on the display. While the participants were aware that the card drawn corresponded to one of two tasks, they were not informed that each deck had a different probability of a task switch. The deck in which switches occurred more often thus required greater mental processing. Despite their overt lack of awareness of the differences between the demands of the decks, subjects nevertheless opted to choose the deck with lower demand significantly more often than the one with higher demand. Kool, et al. (2010) thus concluded that there exists a preference for choosing tasks with reduced cognitively demand, even when the potential processing load is not explicitly known by the observer.

The tendency to favor less costly tasks was also found for reaching tasks. Moher and Song (2014) found that when subjects had to physically reach for one of two boxes to indicate their response to a perceptual task, they were less likely to change their mind mid-reach when the boxes were further apart. Moher & Song (2014) concluded that decision-making processes were dynamically altered to avoid a higher motor cost (longer reach path) associated with a change of mind.

Real-time adjustments of decisions are not limited to reaching. In situations involving gaze, the decision of where to look next can be based on the processing demands of the task. Ballard, Hayhoe, and Pelz (1995) examined a block-copying task, recording eye movements as subjects copied a model pattern of colored blocks using a mouse to move the blocks from a source pile to a workspace in a separate area of the screen. Analysis showed a trend of frequent eye movements to the model area of the screen, usually before and after retrieving a block from the source area. This indicated that subjects were minimizing the load on working memory by re-fixating the model right before they selected a block and brought it to their workspace. This preference avoids excessive demands on memory in favor of using eye movements, which presumably make fewer demands on limited resources, i.e., the model is used as an external memory (O'Regan, 1992). The preference for eye movements over memory was reduced when the areas of the screen were moved far apart enough to require head movements, a presumably more demanding action than a gaze shift, to look between the screens.

Kibbe and Kowler (2011) also found tradeoffs between reliance on memory in motor behavior in a task that involved a somewhat higher processing load. They used a category search task in an attempt to understand how varying complexity and motor demands involved in the task could affect use of memory and exploratory strategies. The task was to find three objects that fell into a common category, where the category rules varied in complexity. Kibbe & Kowler (2011) found that both the complexity of the category rules, and the motor aspects of the task (mouse click vs. gaze shifts; gaze shifts with and without imposed delays) affected the information-gathering strategy. In particular, trials involving more complex rules yielded more revisits to the stimuli, meaning there was less of a reliance on memory and a greater emphasis on exploration. More challenging motor requirements had the opposite effect. Kibbe and Kowler (2011) concluded that cognitive demands and motor demands affected decisions about where to look or click next, indicating that the strategies of relying on memory or exploration are adaptive based on these demands.

The need to take the computational costs into account appeared in even some very simple search tasks involving saccadic eye movements. Araujo, Kowler, and Pavel (2001) studied saccadic decisions in a visual search task which required identifying the orientation of a target that could appear in one of two eccentric locations. A luminance cue indicated the location that was more likely to contain the target. Araujo et al. (2001) found that virtually all subjects preferred to ignore the cue and instead made the first (and only) saccade to the less eccentric location. This preference was attributed to the reluctance to delay the saccade long enough to identify the luminance cue. Araujo et al. (2001) concluded that observers preferred to respond quickly (even at the cost of greater search error) rather than invest more time or effort to identify the cue.

Hooge and Erkelens (1999) later examined a search task that also involved peripheral selection in addition to foveal discrimination. In this task, a single circle had to be located in an array of Cs. Subjects were tasked with finding a thin circle in an array of the thin and fat letter Cs of various orientations, so an efficient strategy involved peripherally selecting thin targets over thick distractors. Since only the thin objects could be targets, it was found that dwell time for the thick objects was significantly shorter than for the thin objects, which were more likely to be targets. Shorter dwell time was also found for Cs with a large gap, which clearly indicated that the object was a C. It was posited that this was due to the fixations on thin objects with smaller gaps requiring more foveal processing to discriminate whether the fixated object was a C or the target circle. Peripheral examination, on the other hand, was more prominent on thick object fixations, where exploratory saccadic planning in selecting the next target was more brief. This further emphasizes that on-line changes in search strategy are made to account for the increased load of the current fixation.

A similar reluctance to make full use of peripheral cues was demonstrated by Wu and Kowler (2013) in a search task involving multiple locations. Participants were instructed to fixate a sequence of 6 thin target circles (as opposed to thicker distractor circles) containing a tilted line of a particular angle drawn from a Gaussian distribution. In the first "statistical estimation" they were told to determine the mean of the tilted lines in the target circles. In the second "look-only" task, they were told to simply fixate as many thin target circles as possible during the trial. Their results showed that the fixation dwell time was significantly greater during fixations to targets than during fixations to nontargets, emphasizing the increased on-line processing when targets are fixated vs. the more exploratory nature of the distractor fixations. This result echoes previous findings that exploratory pace is not slowed in an attempt to fixate targets (Hooge & Erkelens, 1999).

In summary, in situations ranging from animal foraging to various visual-motor tasks, motor and mental costs are taken into account when deciding where to inspect next. Regardless of whether the observer is looking to gain energy, find a target, or gather information, costs and benefits must be weighed in order to optimize search.

1.2. Present study

Acquiring information by means of action carries with it a consideration of the value of information and the cost of getting it. Costs include time, cognitive resources, and motor effort. The reviewed tasks, while effective in their respective domains, took place much on shorter time-scales and were limited in scope compared to those we usually encounter every day. Furthermore, the tasks required modest cognitive load; the studies involved simple judgements and interpretations of the stimuli. This is contrasted with the more complex, open-ended, and lengthy information-gathering tasks that people do all the time.

One example of information gathering that requires considerable management of resources, time, and strategies is the integration of information across graphs and text. Textbooks, websites, and journal articles often contain graphs and accompanying passages of text. They may appear side by side or on different pages, requiring actions, a click of a mouse or a turn of a page, to switch between them. The present experiment addressed two main questions:

1) How are strategies for integrating graph and text affected by perceptual or motor factors, such as the motor or cognitive effort required to access the information? An answer to this question might provide insights into how we manage tradeoffs between the value of information and the cost of acquiring it.

2) At what level do people integrate information across graphs and text? Integration might occur at a higher, holistic level, after the entire graph and text are read. Alternatively, integration might be feature by feature, with observers looking back and forth frequently between the parts of graph and text in a piecewise fashion.

We examined the integration of information across graph and text using stimuli that described simple attributes of two fictitious products. The task required observers to pick which product they preferred. There were three types of presentations. In the *Simultaneous* condition, the graph and text were displayed side by side on the screen, requiring only a saccadic eye movement to switch between the graph and text. In the *Button-Press* condition, the graph and text were displayed sequentially, requiring the press of a button to switch between the graph and text. In order to preserve the sequential aspect of the presentation without adding motor effort beyond what was required in the *Simultaneous* condition, a third condition, *Eye-Contingent*, was included. In the *Eye-Contingent* condition, graph or text appeared sequentially in the same location as in the *Simultaneous* condition, triggered by a saccade into the corresponding region on the screen. These conditions allowed both the perceptual and motor availability of the stimuli (graph or text) to be manipulated while keeping the actual informational content consistent.

2. Methods

2.1. Eye movement recording

Eye movements were recorded using the EyeLink 1000 (SR Research, Osgoode, Canada), tower-mounted version, sampling at 1000 Hz. A chin rest was used to stabilize the head. Viewing was binocular and eye movements were recorded from the right eye.

2.2. Subjects

Subjects were 22 students at Rutgers University, 11 tested in Instruction 1 and 11 in Instruction 2 (see Procedure below for definitions of instructions). All had normal vision and were naïve to the purpose of the experiment. An additional 7 subjects were tested but data were not analyzed because: (1) at least 30% of the data were lost due to excessive blinking or head motion (5 subjects); (2) the subjects did not understand how to use the gamepad to switch between the graph and the text in the Button-Press condition until late in the experimental session (2 subjects). Testing was in accordance with the Declaration of Helsinki and approved by the Rutgers University Institutional Review Board.

2.3. Stimulus display

Stimuli were displayed on a Dell U2413 LCD monitor (refresh rate 60 Hz) viewed from a distance of 60 cm. Stimuli were displayed within a 1280 x 1024 pixel (28.2 x 22.5 deg) region of the screen. Displays consisted of a bar graph and a paragraph of text (Fig 1, S1), as described below.

2.3.1. Graph stimuli

Graphs were generated as 447 x 502 pixel (9.9 x 11.1 deg) images. The axes of the graph were contained within a 329 x 410 pixel (7.3 x 9.1 deg) rectangle.

Graphs contained 4 colored bars on a white background. The bars compared the values of two fictitious common household products along two different attributes (Fig. 1), with values along the two attributes shown on the left and right Y-axes, respectively. Lettering (legend and axis labels) were black. Bars were grouped in pairs, either according to the products or the attributes. The labels on the X-axis under each pair indicated the name of the product or attribute. The colors of the bars within each pair corresponded to a different attribute (in the case of grouping by product) or product (in the case of grouping by attribute), and the significance of the color was indicated by the legend in the upper left corner.

The values along each of the attributes were selected so that the relative merits of the two products on each attribute were either in conflict or not in conflict. In the Conflict condition, one product was better than the other on one attribute, and worse in the other attribute. In the No Conflict condition, one product was better than the other on both attributes. These two conditions were included to introduce sufficient uncertainties so that subjects would be encouraged to inspect each stimulus rather than develop beliefs about whether one or both attributes would be relevant.

Twenty-four different bar graphs were generated, each for a different pair of products. Four versions of each of these 24 graphs were generated, according to whether (1) the values of the attributes were either in conflict or not in conflict, and (2) the bars were grouped by item or by attribute. These variations were included to encourage subjects to attend to the details in each graph, rather than assuming that the information in each graph would be arranged in the same way, or that a decision about product preference could be made by examining only one of the attributes.

2.3.2. Text Stimuli

Each graph was accompanied by a paragraph of text, 264 to 387 characters in length (including spaces). Text was black 18 pt monospaced Courier New font (width = 13.6 pixels/character or 6 characters per degree) on a white background. The text was displayed within a 555 by 760 pixel region (12.3 deg x 16.8 deg). Each of the four versions of graph (see Graph Stimuli above) could be accompanied by one of two types of text (Fig. 1), redundant or non-redundant. Redundant text restated the information depicted in the graph. Non-redundant text provided information that differed in some way from the graph. This information was either irrelevant to the relative merits of the products, or added details that favored one of the products. Some of the non-redundant texts stated that the information depicted in the graph was outdated or contained an error. The use of both redundant and non-redundant text, like the four different versions of the graph, was done to discourage pre-existing beliefs about the value of the graph and text in the decision. Note that the study was focused on the eye movement strategies and not on whether the subject made a "correct" decision since in many cases (non-redundant texts, conflict between attributes), there was no "correct" choice.

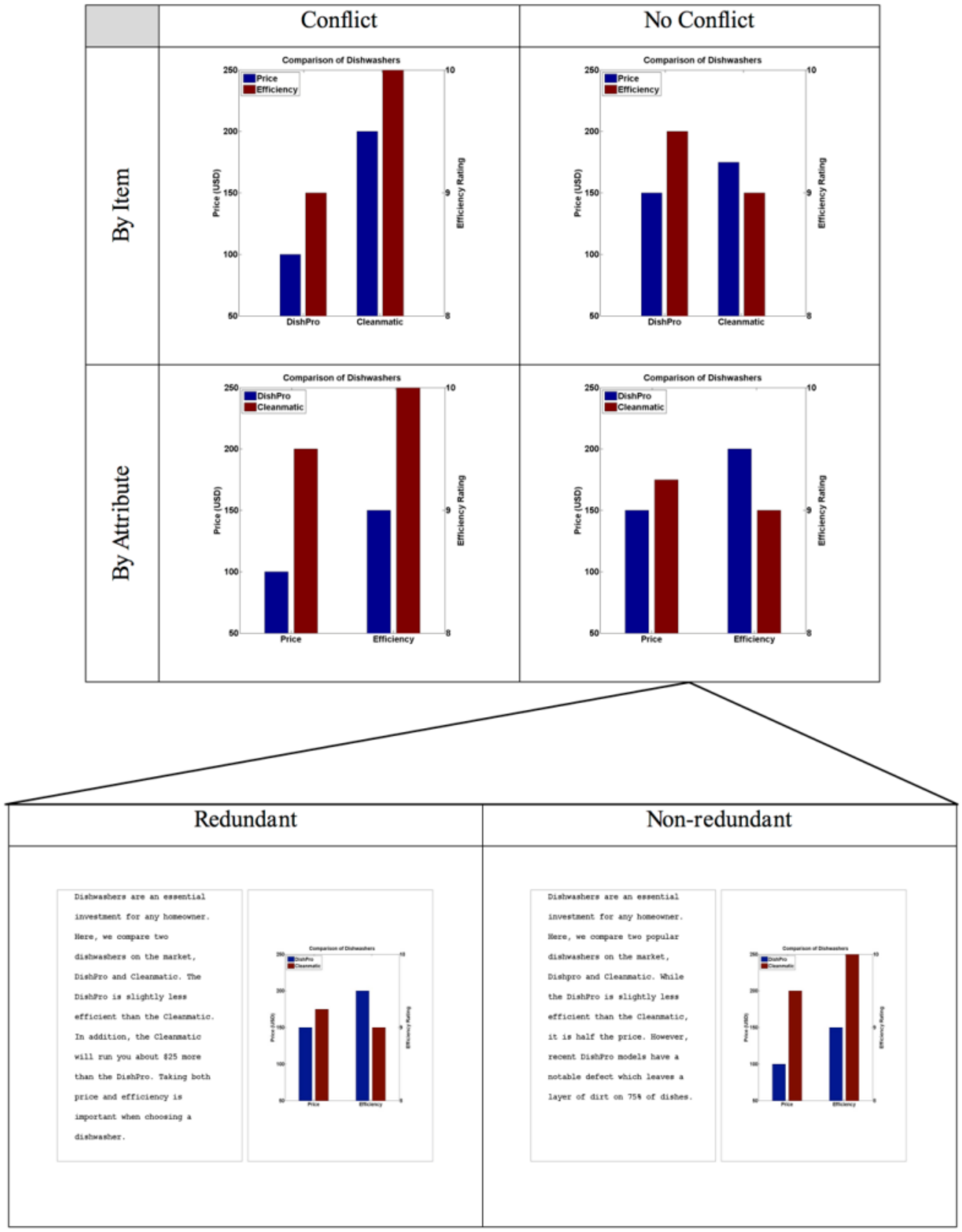

Figure 1: Examples of four types of graph stimuli and two types of associated text. The products could either be in Conflict (left), in which one product was better than the other on only one attribute, or not in conflict (right), in which one product was better than the other on both attributes. Bars on the graph could also be arranged by Item (top), in which the bars were grouped by product as indicated by the x axis label with the color of the bars corresponding to the attributes being depicted, or by Attribute (bottom), in which the bars were grouped by the attributes being depicted with the color of the bars indicating the product. The text could either be Redundant (bottom left) with the graph or Non-redundant (bottom right). Redundant text described the data in the graph. Non-redundant text provided either additional information that was irrelevant to the comparison or favored one product over the other.

2.4. Perceptual-Motor conditions

Four Perceptual-Motor conditions were tested.

(1) Simultaneous: Graph and text were displayed side by side. The regions containing the graph and the text were separated by a blank region 16 pixels (0.35 deg) wide. The graph was displayed either on the right or left side, randomly and independently chosen on each trial.

(2) Button-Press: Graph and text were display sequentially, each located in the center of the screen. Subjects pressed either the left or right trigger button on a gamepad to switch to graph or text, respectively.

(3) Eye-Contingent: Graph and text were displayed sequentially, as in the Button-Press condition, and appeared either on the right or left side of the display, randomly chosen, in the same locations as in the Simultaneous condition. The appearance of the graph or text was triggered by online detection of the offset of a saccade into the right or left side of the screen. Delays between the onset of the fixation and onset of the display were about 50 ms, except for the very first appearance of the text, for which delays were about 120 ms.

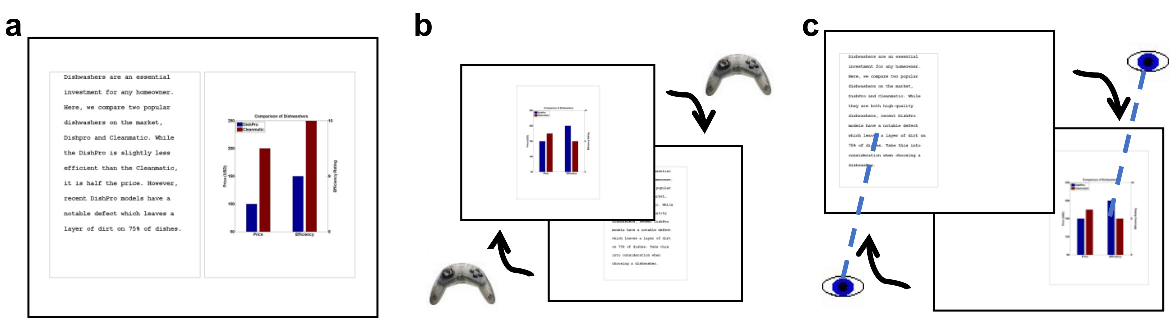

Figure 2: Examples of displays showing the 3 Perceptual-Motor conditions. a: Simultaneous: Graph and text appear side by side. b: Button-Press: Graph and text appear in the center, with the sequential appearance of each triggered via a button press. c: Eye-Contingent: Graph and text appear side by side, with the sequential appearance of each triggered via saccade into the corresponding area of the screen.

2.5. Procedure

The task that motivated reading of the graph and accompanying text was to choose the preferred product. Each subject was tested in a single 24-trial experimental session. Before testing began subjects were told they were would view a series of graphs accompanied by passages of text about two products. At the end of each trial they would be asked to indicate which they would prefer. The first group of subjects tested was told they should read the graph and text to determine their preference (Instruction 1). The second group of subjects was given instructions that did not contain the word “read” in order to avoid implying that the text was more important than the graph, and told only to indicate their preferred product based on the display (Instruction 2). Subjects were told they could end the trial by pressing a button on the gamepad when they were ready to make the decision.

Before testing began subjects were presented with three practice trials, one for each of the Perceptual-Motor conditions, in order to illustrate the Perceptual-Motor conditions and show how the display could be switched between graph and text using either the button press or a saccade.

The calibration routine built into the EyeLink software was run before the start of each experimental session and again midway through. Before each trial, the number of trial as well as a label indicating the Perceptual-Motor condition was displayed. Subjects started the trial with a button press when ready. Then, 5 crosses were presented for 5 s, one in the center and one in each corner of the display to serve as confirmation of the calibration. Subjects were told to fixate the center cross then look to each of the other 4 crosses in sequence, then back to the center cross. The crosses then disappeared, replaced by the critical display.

In the Button-Press condition the button subjects used to start the trial determined whether the graph or text appeared first. Subjects learned about which button corresponded to the graph and which to the text during the practice trials described above. In the Eye-Contingent condition, the screen was blank until subjects fixated either side, at which point the graph or the text appeared. Thus, in the Eye-Contingent condition, subjects did not know which side corresponded to each until after the first saccade.

Subjects were instructed to press a button to end the trial when they were ready to make the decision about their preferred product. The trial automatically ended after 2 minutes if the subjects did not chose to end it themselves (only 3 of the 528 total trials tested lasted the full 2 minutes). After the graph and the text were removed, the 5-point calibration was re-run. Then, the subject indicated by button presses (1) which product they preferred, (2) how confident they were in their choice on a scale of 1-4 (with 1 being least confident and 4 being most confident), and (3) which of the two attributes was more influential in their choice. Note that a detailed analysis of whether fixation location predicted the decision in any given trial was outside the scope of the present study, which was focused on the role of the Perceptual-Motor condition.

2.6. Design

The perceptual motor conditions were assigned to each trial randomly using an algorithm that employed the following constraints. (1) Each subject was tested on 8 trials for each of the three Perceptual-Motor conditions. The order of testing trials with the different Perceptual-Motor conditions was random. (2) The pairing between a given graph and a given perceptual motor condition was different for each subject. The pairing was done so that across subjects each of the 24 graphs was paired with a given perceptual motor condition at least once for Instruction 1 and at least once for Instruction 2. Table 1 shows the frequency of a particular pairing of the 24 graphs and each Perceptual-Motor condition across subjects. (4) Text conditions (redundant vs. not redundant, see Text Stimuli above) were tested in blocks of 6 trials each, with the first block chosen at random. (5) Graph type (conflict vs. no conflict; grouping by product vs. grouping by attribute) and the side of the screen containing the graph were independently chosen at random on each trial.

Table 1

Frequency of each pairing of a Perceptual-Motor with a particular graph. Each graph is numbered 1-24.

Graph	Simultaneous	Button-Press	Eye-Contingent
1	6	6	10
2	10	6	6
3	11	7	4
4	9	4	9
5	7	11	4
6	4	7	11
7	5	7	10
8	7	8	7
9	6	8	8
10	9	4	9
11	9	3	10
12	6	11	5
13	4	8	10
14	6	8	8
15	5	4	13
16	9	8	5
17	5	8	9
18	8	8	6
19	4	8	10
20	7	10	5
21	13	7	2
22	8	6	8
23	8	12	2
24	10	7	5

2.7. Analysis

The beginning and ending positions of saccades were detected offline by means of a computer algorithm employing a velocity criterion to find saccade onset and offset. The value of the criterion (a change in eye position of 22 deg in 16.7 ms) was determined empirically for individual observers by examining a large sample of analog recordings of eye positions. Portions of data containing blinks or episodes where tracker lock was lost were eliminated (Trials eliminated: Simultaneous: 4, Button-Press: 11, Eye-contingent: 4).

The location of each fixation in the display was calculated from the average position of the line of sight at the offset of saccade n-1 and the onset of saccade n. The duration of the fixation pause was the difference in the time between saccade onset and offset times. Each pause was classified as being in the graph, the text, or neither. A "visit" to graph or text was defined as a set of consecutive pauses in the same region.

Pauses and visits to graph and text were further subdivided into the following areas of interest (AOIs) using an algorithm that categorized fixations based on the boundaries of manually-defined regions in the graph or text. Boundaries are shown in Fig. 3. AOIs were denoted as:

LEG: legend area

YAR: right y-axis

YAL: left y-axis

DAR: right pair of bars

DAL: left pair of bars

XAX: x-axis

OTG: graph, not otherwise identified (for example: title area)

TXT: the top third of the text

TXB: the bottom two thirds of the text

NML: region between graph and text

OTH: none of the above regions

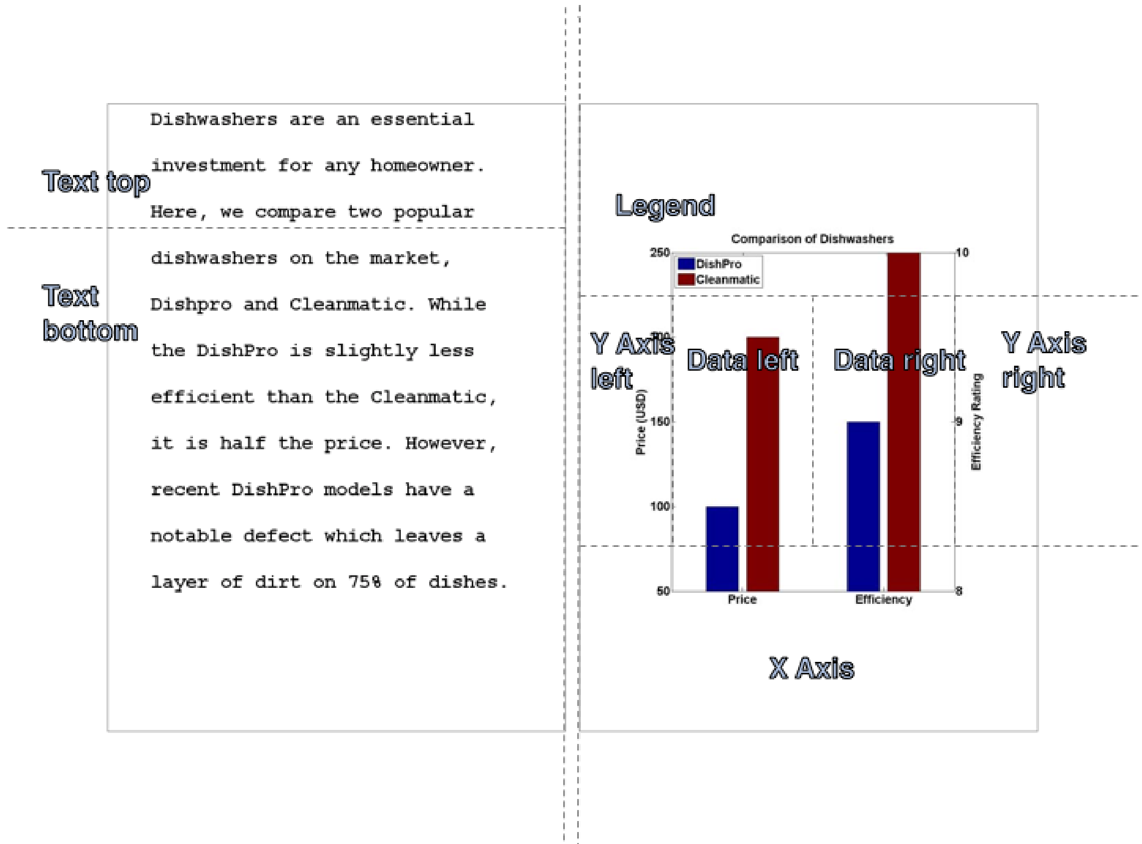

Figure 3: Location of boundaries of Areas of Interest with labels of each region. The algorithm used the location of the eye relative to these boundaries and associated the corresponding label based on the region.

3. Results

3.1. Effect of Perceptual-Motor conditions on the visits to graph and text

The Perceptual-Motor conditions had a large effect on the pattern of visits to graph and text. There were about twice as many visits to graph or text per trial in the Simultaneous condition than in either the Button-Press or the Eye-Contingent conditions. This can be seen in Figure 4 (left), which compares the number of visits/trial to graph and text (means of subject means) for the three Perceptual-Motor conditions for both instruction types. Analysis of variance (3 Perceptual-Motor conditions x 2 instruction types) showed a significant effect of Perceptual-Motor condition ($F(2, 40) = 24.01$, $p = 10^{-7}$), and no significant effect of instruction type ($F(1, 20) = 0.52$, $p = 0.48$).

Figure 4 (left) also shows that the average number of visits/trial in the Eye-Contingent condition was almost the same as in the Button-Press condition, and quite different from the number of transitions/trial in the Simultaneous condition. This means that the relevant factor discouraging transitions was not purely motor (the button press), but rather the requirement to cope with the sequential visual availability of the graph and the text.

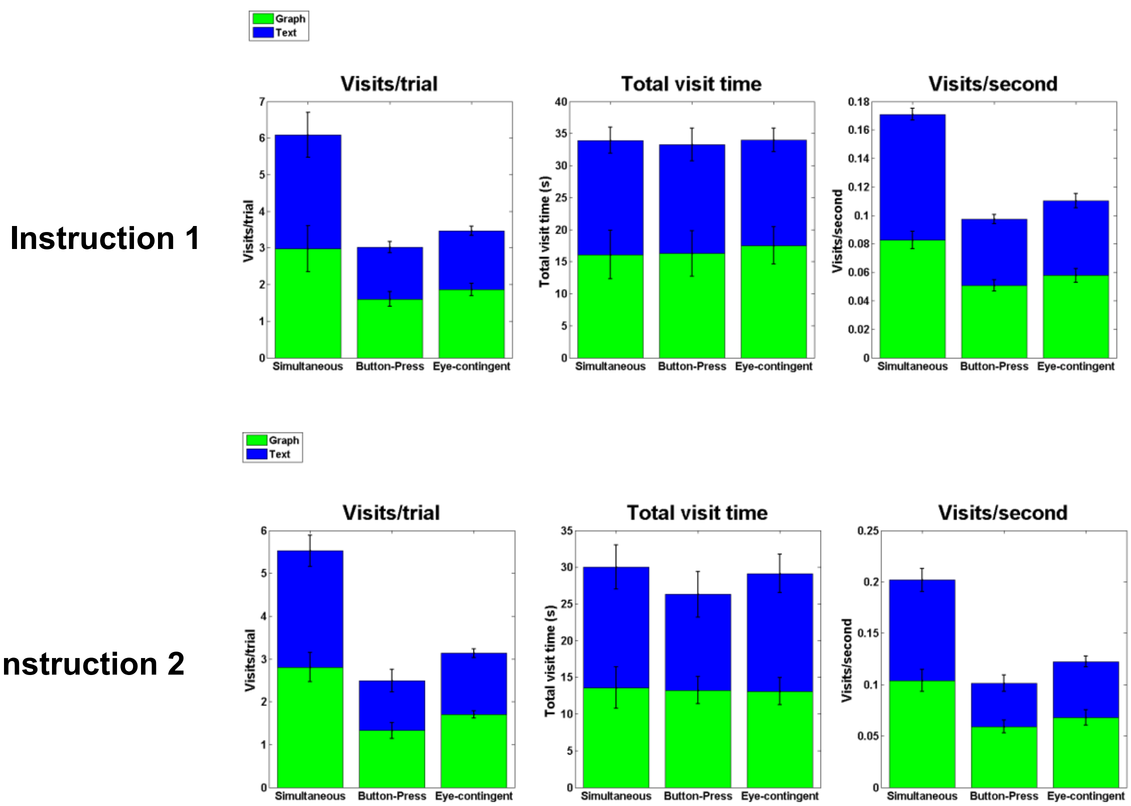

Figure 4. **Left**: Number of visits to graph and text per trial for each Perceptual-Motor condition (mean ± standard error) for Instruction 1 (top) and Instruction 2 (bottom). **Center**: Total duration per trial (s) of visits to graph and text for each Perceptual-Motor condition (mean ± standard error) for Instruction 1 (top) and Instruction 2 (bottom).

Right: Number of visits/second for each Perceptual-Motor condition (mean ± standard error). Means are based on means of subject means, with 8 trials/subject/Perceptual-Motor condition and 11 subjects for each instruction.

The greater number of visits/trial in the Simultaneous condition relative to the other two conditions was not due to effects of the Perceptual-Motor condition on viewing duration. Average trial duration did not differ across the three Perceptual-Motor conditions (Figure 4 (center); 3x2 ANOVA showed no effect of Perceptual-Motor condition, $F(2, 40) = 1.06$, $p = 0.36$, nor Instruction type, $F(1, 20) = 0.65$, $p = 0.43$), and it can be seen that there was about an equal division of time between graph and text.

The effect of Perceptual-Motor condition on the division of time between graph and text throughout each trial can be best summarized by the visit rate (visits/s), which was twice as high for the Simultaneous condition than for the other two Perceptual-Motor conditions (Figure 4 (right); Perceptual-Motor condition: $F(2, 40) = 62.56$, $p = 10^{-13}$); Instruction type: ($F(1, 20) = 1.33$, $p = 0.26$). Taking all these measures together shows that the Perceptual-Motor conditions affected how time was distributed between the graph and the text, and not in the total amount of time devoted to each.

The visit rate was not affected by the experimental manipulations of the content of the graph and text, namely, (1) whether the relative merits of the two fictitious products on each attribute was consistent or in conflict ($F(1, 20) = 0.89$, $p = 0.36$; Instruction type: $F(1, 20) = 0.95$, $p = 0.34$), (2) whether the text was redundant with the graph or introduced new or different information ($F(1, 20) = 3.74$, $p = 0.068$; Instruction type: $F(1, 20) = 1.47$, $p = 0.24$) , and (3) whether the graphs were organized by Item or by Attribute ($F(1, 20) = 0.77$, $p = 0.39$; Instruction type: $F(1, 20) = 1.17$, $p = 0.29$). As indicated in methods, these conditions were not a focus of the experiment, but were used to vary the content and organization of the information, so that subjects' strategies would not become stagnant.

3.2. Timelines

To better visualize the differences in strategy among the Perceptual-Motor conditions, timelines showing visits to the graph and the text for each trial were constructed. Figures 5 and 6 show the timelines for each Perceptual-Motor condition and each Instruction type, with trials ordered from shortest (bottom) to longest (top). A "visit" was composed of the cumulation of successive fixations in the same area, graph or text, including the gaze shift time, as well as any intervening blinks. Any visits to locations other than the graph or the text, including fixations in the blank region between graph, are shown in red. Blank regions of the graph indicate that the visit prior to the blank contained a period of lock lost greater than 2 seconds.

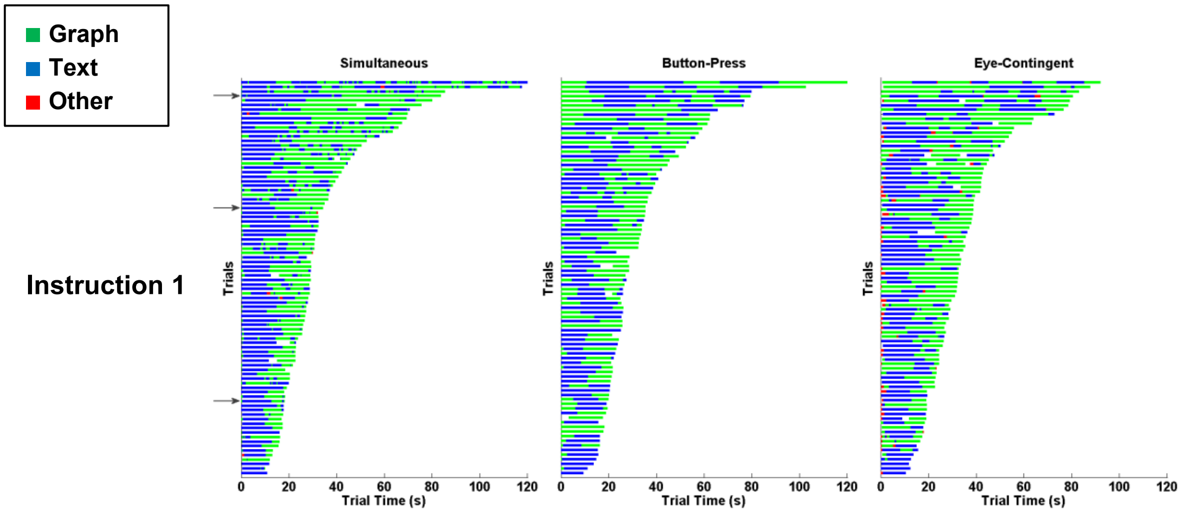

Figure 5: Timelines depicting time spent viewing the graph (green), text (blue), or other areas (red) as the trial progressed for Instruction 1 (n = 11) for the Simultaneous (left), Button-Press (center), and Eye-Contingent (right) conditions. Trials are ordered from shortest to longest. The length of each timeline bar corresponds the duration of the visit (cumulation of fixations) to the graph or text. Arrows on the Simultaneous timelines indicate trials with brief glances to graph or text at the beginning of the trial before a longer visit to the other region. Blank regions of the graph indicate that the visit prior to the blank contained a period of lock lost greater than 2 seconds.

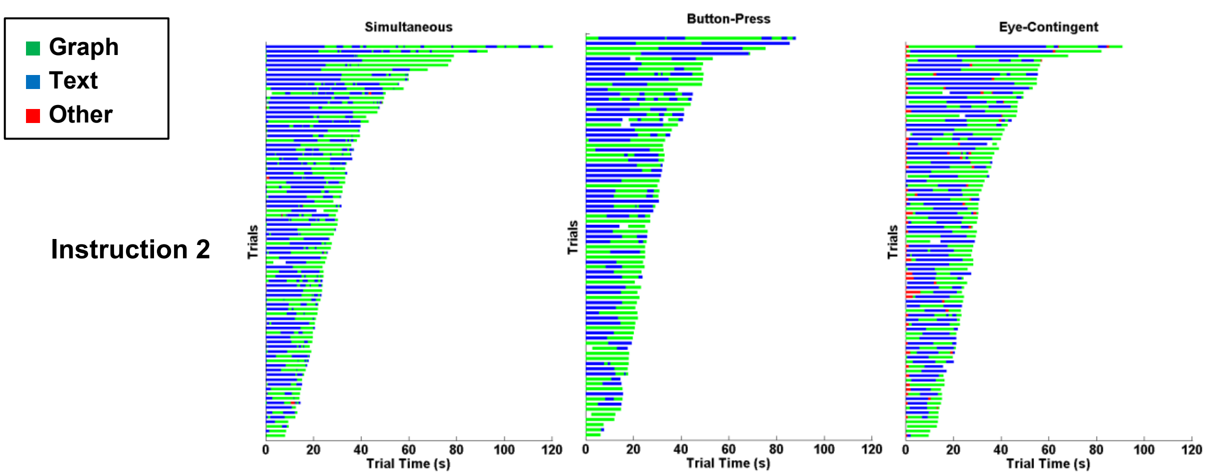

Figure 6: Timelines depicting time spent viewing the graph (green), text (blue), or other areas (red) as the trial progressed for Instruction 2 (n = 11) for the Simultaneous (left), Button-Press (center), and Eye-Contingent (right) conditions. Trials are ordered from shortest to longest. The length of each timeline bar corresponds the duration of the visit (cumulation of fixations) to the graph or text. Arrows on the Simultaneous timelines indicate trials with brief glances to graph or text at the beginning of the trial before a longer visit to the other region. Blank regions of the graph indicate that the visit prior to the blank contained a period of lock lost greater than 2 seconds

The timelines show two main trends. First, most trials started with a long visit to either the graph or the text, after which gaze switched over for a long visit to the other region. Second, the initial long visits were often followed by shorter-duration visits to the graph or the text. These shorter visits occurred more frequently in the Simultaneous condition. These trends will be examined in more detail below.

3.3. Initial visits to graph or text

The initial visit to either the graph or the text was examined first. “Initial visit” refers to the first time either the graph or text was viewed, regardless of which was visited first. To count as an initial visit, the duration had to be greater than 1 second. This value was chosen to avoid including as an initial visit the occasional brief glances to the graph or to the text that occurred early within the trial and were followed by a prolonged visit to either region. (See, for example, the trials indicated by an arrow on the Simultaneous timelines in Figure 5) In these cases, the visit following this brief glance was designated as the initial visit.

The pattern of visits to graph and text are summarized by histograms showing the number of visits of different durations to graph and text in Figures 6 and 7. These histograms were collapsed over Instruction type and show the distribution of durations for the first 3 visits to graph (Figure 6) and text (Figure 7). The bar graphs inset in Visit 1 indicate the number of trials not represented in the histogram due to the lack of visits to the graph or the text respectively. This includes first visits that were eliminated because they did not meet the requirement of initial visits to be greater than 1 second and visits with excessive amount of lock lost.

The average durations of initial visits were 7-12 for the graph and 12-15 seconds for the text, long enough to allow extensive examination of the graph or the text. These findings suggest that the preferred strategy was to attempt to extract meaning from large portions of text and entire graph, rather than a strategy based on very brief feature-by-feature comparisons of the graph or the text.

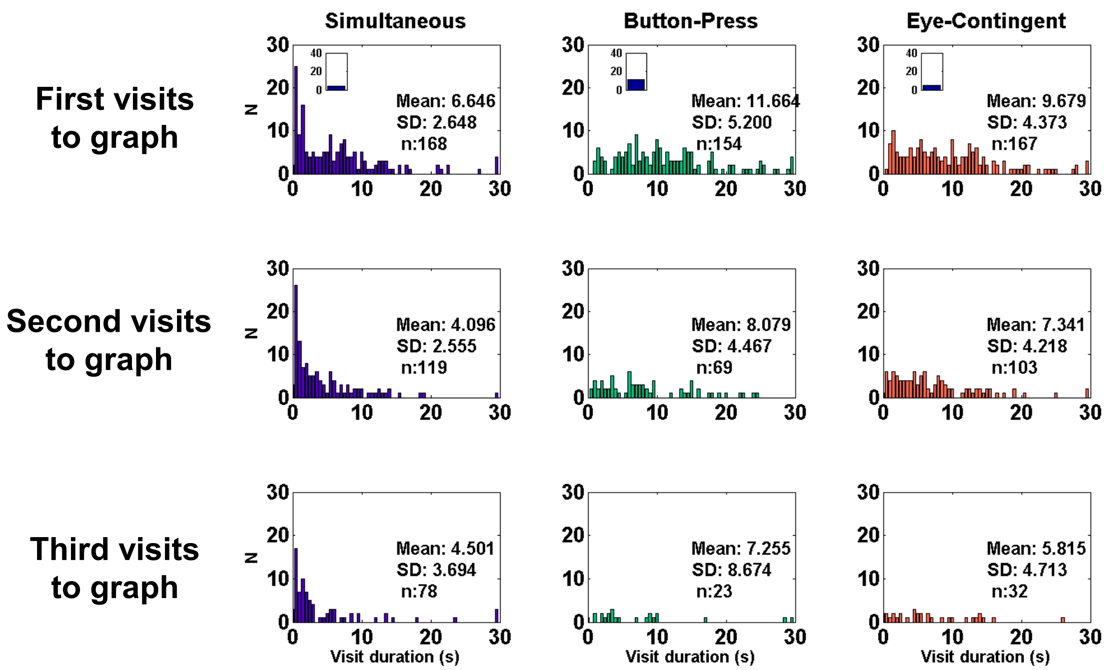

Figure 7: Histograms showing the number of visits of different durations (in seconds) for the first 3 visits to the graph for the Simultaneous, Button-Press, and Eye-Contingent conditions. Inset graphs show the number of eliminated visits (<1 s or excessive amounts of lock lost). See text for definition of a "visit".

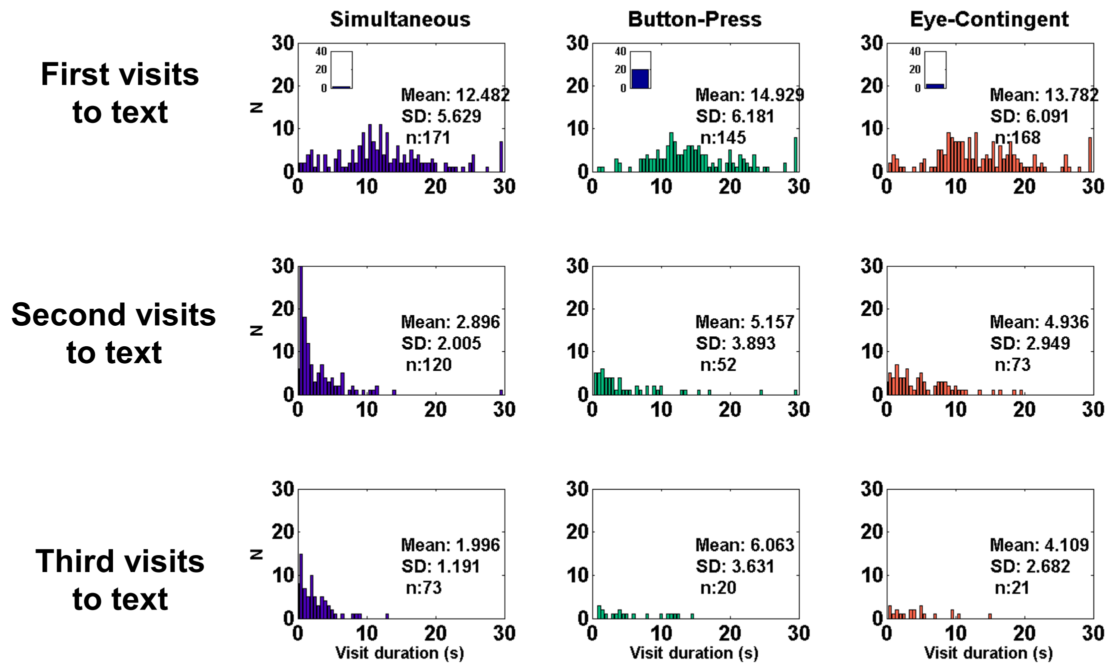

Figure 8: Histograms showing the number of visits of different durations (in seconds) for the first 3 visits to the text for the Simultaneous, Button-Press, and Eye-Contingent conditions. Inset graphs show the number of eliminated visits (<1 s or excessive amounts of lock lost). See text for definition of a "visit".

The durations of the initial visits to graph or text varied among the perceptual motor conditions ($F(2, 40) = 18.66$, $p = 10^{-6}$; Text: $F(2, 40) = 5.79$, $p = 0.0062$). The initial visits to either graph or text were shorter in the Simultaneous condition than in the other two conditions.

In the Simultaneous condition, subjects more frequently visited the text first than the graph first (Visiting text first: Instruction 1: 88%, Instruction 2: 67%). Preferences to visit graph or text first were about the same in the other two Perceptual-Motor conditions (Instruction 1: Button-Press: 43%, Eye-Contingent: 56%; Instruction 2: Button-Press 53%, Eye-Contingent 55%).

3.4. Subsequent visits to graph or text

The timelines (Figure 5) indicated that the typical pattern was an initial long visit to the graph or the text followed by several shorter revisits to either region, particularly in the Simultaneous condition. This pattern was confirmed by the histograms of the durations of second and third visits to text (Figure 6) or graph (Figure 7). The histograms show that the Simultaneous condition included a large number of relatively short duration (<5 s) second and third visits to the graph or the text. Second or third visits to graph or text in the other two perceptual motor conditions were on average longer in duration, and also occurred much less frequently.

Looking at the ns on these graphs, it appears that there were many more visits in the Eye-Contingent than in the Button-Press condition, which would be contradictory to the earlier result depicted in Figure 4 that there was only a large difference in visits per trial between the Simultaneous condition and the other two. However, this is simply due to how the means of visits per trial were calculated (means of subject means). The histograms in Figures 5 and 6 show all trials with no averaging, so variations in the number of visits per trial between subjects appear more prominently in these graphs even though they only contribute 1/22 of the mean in the calculation of visits per trial.

3.5. Integration strategy

The difference across perceptual motor conditions in the frequency and duration of the second and third visits to graph or text suggest that the perceptual motor condition influenced the strategy used to integrate information across the two regions. For example, in the Simultaneous condition, where initial visits were briefer and subsequent visits more frequent, it is possible that the viewer decided to segment the material, reading only part of the graph or text and then switching to the other region. Alternatively, it is possible that the greater numbers of subsequent visits to graph or text in the Simultaneous condition were used to revisit previously seen regions.

In order to distinguish these possibilities, the location of individual visits within graph and text were examined to distinguish between visits to new regions and revisits of previously viewed regions. A revisit within the graph was defined as a re-examination of one of the 6 Areas of Interest (AOI) of the graph that was examined on a prior visit (see Methods for a list of the AOI's and how the fixations were assigned to an AOI). A revisit within the text was defined as re-reading of a line of text that had been previously read. If the strategy were to segment the material on each visit, then the proportion of time devoted to re-reading text or to re-examining AOI's of the graph would remain low across visits to each. On the other hand, if the strategy was to examine graph and text exhaustively, or near-exhaustively, on the first visit, then the proportion of time devoted to re-reading or re-examination would increase over subsequent visits to each.

Figure 9 shows the time spent viewing new areas (blue) and re-examining areas (in red) of the graph (top) and text (bottom). Each section of a bar represents the mean of each subject's average duration over trials in that Perceptual-Motor condition. The first 3 visits to graph and text were examined for each Perceptual-Motor condition. This includes trials that had fewer than 3 visits to graph or text. Note that some subjects had no trials with 2 or 3 visits to graph or text in the Button-Press and Eye-Contingent conditions, and some subjects had more than 3 visits in the Simultaneous condition, hence the variation in the number of subjects included in the means in each bar (represented by the number above each bar). Trials with more than 3 visits to graph or text were not represented in the first 3 bars but those visits still contributed to the means for the bars representing all visits.

The time spent rereading averaged over all visits (as indicated by the bars labeled "All" in Figure 9) was about the same between the 3 conditions (In the graph: Simultaneous: 12.0s, Button-Press 13.0s, Eye-Contingent: 12.6s). In the initial visit, a substantial proportion of the (about 42% in the text and 66% in the graph) was spent re-visiting or re-reading. In these initial visits, there were significant differences in average time re-reading text between the 3 Perceptual-Motor conditions ($F(2, 40) = 6.087$, $p = 0.0049$), as well as in the average time re-examining areas of the graph ($F(2, 20) = 13.95$, 10^{-5}). There was no significant difference in time spent reading new material in initial visits to text between the 3 Perceptual-Motor conditions ($F(2, 40) = 0.55$, $p = 0.581$), and a small but significant difference in time examining new material in the graph ($F(2, 40) = 11.77$, $p = 10^{-5}$).

In the second and third visits to graph or text, the average proportion of time spent re-examining AOI's of the graph or re-reading lines of increases in all 3 Perceptual-Motor conditions, to the point where almost the entirety of the visit is spent re-viewing previously viewed areas. This is consistent with the idea that subjects spend their initial visit examining most of the graph or text and spend subsequent visits re-examining the areas they have already looked at rather than segmenting the initial read across several visits.

At first glance, the difference between the Perceptual-Motor conditions in re-examination time overall vs. in individual visits appears to indicate a change in the strategy of gathering information within the graph or text. Taking into account the greater number of visits per trial in the Simultaneous condition compared to the other two conditions (~5 visits in the Simultaneous condition vs. ~3 visits in the other two conditions; Figure 4) and shorter average visit time (Figures 7 and 8), it becomes clear that the main difference in integration strategy is in the same total duration of revisiting spread across the extra, briefer visits. This suggests that the Simultaneous condition only affected the frequency of switches between graph and text and how the time was apportioned and not the strategy of integrating information or distribution of time spent within the graph and text themselves.

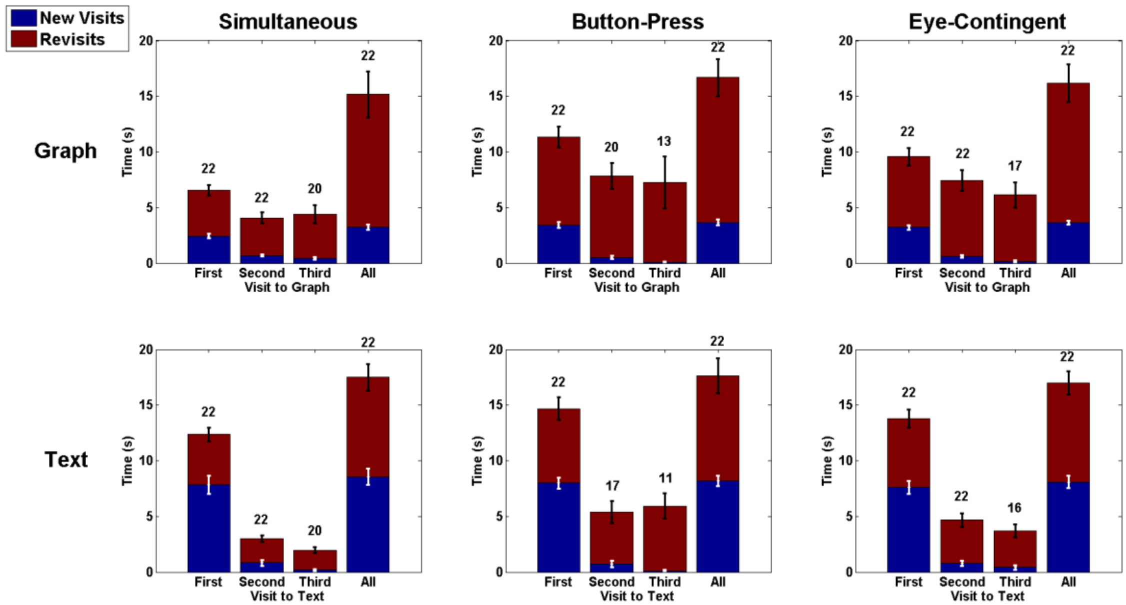

Figure 9: Time spent viewing new areas (blue) and re-examining AOIs that have already been viewed for the first 3 visits (red) for graph (top) and text (bottom) for each of the 3 Perceptual-Motor conditions (mean ± standard error). The last bar (labeled "All") is the average over all visits to graph or text. The numbers above each bar represent the number of subjects that composed the mean for the height of the bars (out of 22 subjects). A subject was not included in a bar if they did not have any second or third visits to graph or text, respectively, in that condition.

3.6. Switching over after the first visit to graph or text in the Eye-Contingent condition

Viewers could choose whether to view either graph or text first, either by where they looked in the Simultaneous condition, or which button they pressed in the Button-Press condition. In the Eye Contingent condition, however, the side of the display containing the graph or the text was not displayed until it was fixated. Thus, the behavior in the Eye Contingent condition allows examination of the question of whether viewers were willing to make an extra saccade to view a preferred region if the initial fixation landed on a non-preferred region, or whether they were willing to remain in the region that they happened to fixate first. To address this question the proportion of trials viewers stayed with or switched from the region fixated first was examined for the Eye

Contingent condition. This analysis showed that the viewers rarely decided to switch (switches occurred on 4% of the trials where text was fixated first, and 16% of the trials in which the graph was fixated first). This willingness to leave the choice of first location to chance in the Eye Contingent condition, despite a clear preference to visit the text first in the Simultaneous condition (Table 2), is another indication of the influence of perceptual availability on the viewing strategy.

Table 2

Proportion of trials in which the text was visited first (for >1 second)

Instruction Type	Simultaneous	Button-Press	Eye-Contingent
Instruction 1	.88	.49	.57
Instruction 2	.68	.54	.53

3.7. Strategies of viewing the graph

The Perceptual-Motor condition affected the strategy of integrating information across graph and text but did not affect the way the graph itself was examined. This was shown by the matrices showing the frequency of transitions between different AOIs within the graph for all three perceptual motor conditions (Figure 9). The matrices are virtually identical. The most frequent transitions occurred between the labels of the axes and the bars showing the data, and between the bars and the legend. The strategy of frequent fixations on the referents (axis labels; legends) confirms the trends described by Carpenter and Shah (1998).

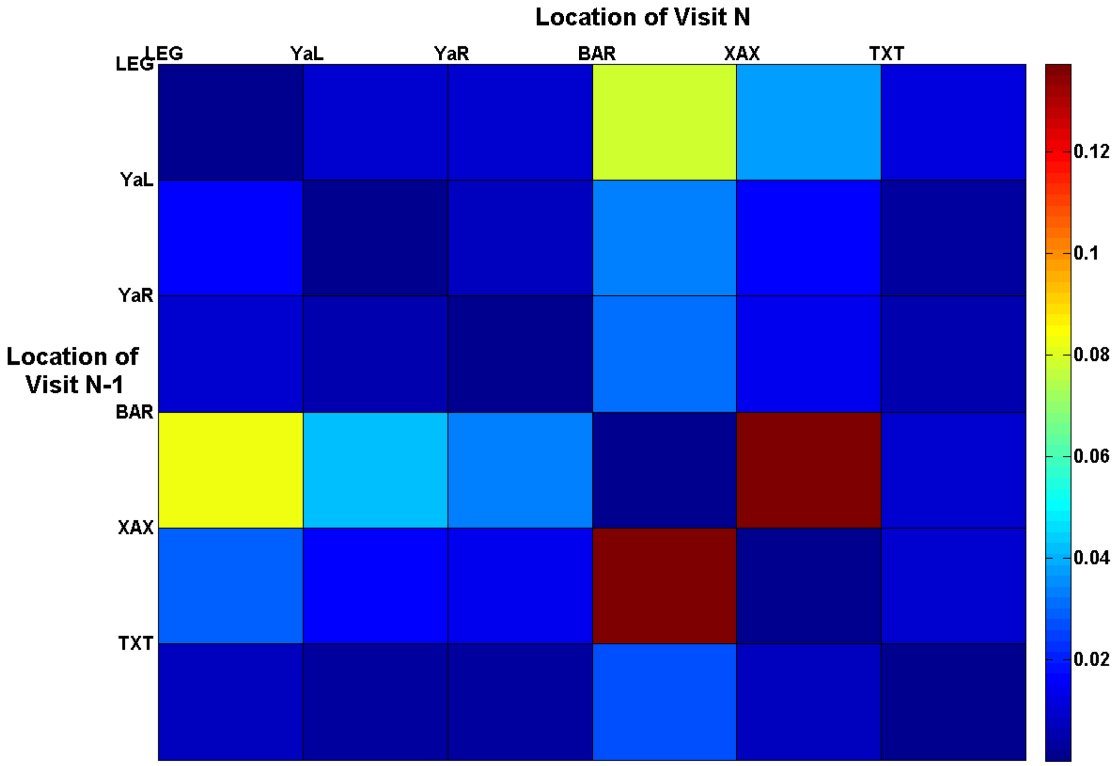

Figure 10: Matrix depicting the frequency of transitions between different areas of interest (AOI) within the graph or the text for all subjects, trials, and Perceptual-Motor conditions. The horizontal axis indicates the AOI visited on visit N and the vertical axes indicates the AOI visited on visit N-1. The color of each cell depicts the number of transitions from the AOI of visit N-1 to visit N divided by the total number of transitions (n = 12232).

A chi square analysis showed that the AOI of the current fixation was statistically dependent on the AOI of the previous fixation (Simultaneous: $X^2(36) = 1106.94$, $p < 0.0001$; Button-Press: $X^2(36) = 2820.88$, $p < 0.001$; Eye-Contingent: $X^2(36) = 1249.38$, $p < 0.0001$). Statistical dependencies of transitions between these areas of interest transitions were found not only between the AOI's one fixation back, but also two fixations back (Simultaneous: $X^2(36) = 1366.15$, $p < 0.0001$; Button-Press: $X^2(36) = 2046.90$, $p < 0.0001$); Eye-Contingent: $X^2(36) = 1977.40$, $p < 0.0001$), indicating there was some higher-order structure to the saccadic sequences.

4. Discussion

A rational strategy used to gather information from a graph and its accompanying text should be based on the value of the information contained in each. We found that strategies of gathering information from graph and text were based on more than the information content. The strategies also took into account perceptual-motor availability. Specifically, a modest change in perceptual-motor availability, namely, having the graph and text appear sequentially, rather than appearing simultaneously (side by side), affected the frequency and duration of visits to graph and. Perceptual-motor availability did not affect the total duration of the inspection, nor the overall division of time between graph and text, nor the scan paths used to inspect the graph. The main effect of the changes to perceptual motor availability was to make it more likely that the viewer would leave one region (either graph or text) and move to the other.

Comparison to prior work

These findings are in many respects similar to other examples in which the effort (motor or cognitive cost) needed to accomplish a task affected the chosen strategy. For example, Moher & Song (2014) found that decision-making processes were influenced by a tendency to avoid higher motor costs in the form of a lengthier reaching path. But mental or cognitive cost can also have a strong influence on deciding which action to take, such as with Kool et al. (2010) finding that participants avoided choosing tasks that might lead to greater mental processing. Often, tasks pit mental and motor costs against one another, requiring the observer to decide, often dynamically in real-time, which to prioritize. Ballard, et al. (1995) showed that observers preferred to minimize working memory over the addition of motor effort through the use of eye movements. Hooge and

Erkelens (1999) also found tradeoffs between memory reliance and motor effort, with the rule complexity of a task as well as the motor aspect of the task changing the strategy of gathering information. Similar tradeoffs have been observed in search tasks involving multiple locations in which observers dynamically adjust dwell time to allow for more processing as opposed to exploration via saccades (Wu & Kowler, 2013).

The present study differed from these previous findings. First, the task was at a higher level, requiring the integration of information concerning many properties of real-world items presented in a real-world format rather than a simple binary decision or motor response based on a single perceptual display. Additionally, results showed that the main property of these stimuli that affected the information-gathering strategy was not the motor effort but the simultaneous rather than sequential availability of the graph and text. In other words, the fact that the graph and text were visible on the screen at the same time affected the number of transitions between graph and text, rather than the motor action required (button press vs eye movement). This was known to be the case because the frequency of visits to graph and text was similar in the Button-Press and Eye-Contingent conditions and significantly greater in the Simultaneous condition.

Why did sequential vs. simultaneous availability affect the rate of visits to graph and text?

There were two possible ways to view this result. First, the sequential availability could have discouraged transitions from one region to the other. Why might this occur? (1) In the Eye-Contingent condition, there is no pre-saccadic image to be integrated with the post-saccadic image. (2) While regions were eccentric enough in the Simultaneous condition that individual details of the graph or text could not be resolved when fixating the opposing region, subjects may have been using the general layout of eccentric details in the Simultaneous condition to help guide the first saccade in the region. In the Eye-Contingent condition, subjects did not have visual availability of these regions, so a saccade to the region not currently being visited would require observers to either retrieve a representation of that region from memory to plan the saccade or to make a sequence of two saccades (one to make the region appear and the second to get to the desired area of the graph or text). Both of these possibilities lead to extra mental (retrieval from memory) or motor (saccades) steps, and memory-guided saccades are usually less accurate than visually-guided saccades (Gnadt, Bracewell, and Andersen 1991). This adds up to it being overall more costly to make a transition between regions when the other region is not perceptually available, even when a button press is not required to make the switch.

Another way to view the result is that the Simultaneous presentation encouraged or facilitated transitions between the graph and the text. All objects in the visual field have been shown to be represented in areas of the oculomotor network related to saccadic planning (frontal eye field, superior colliculus, parietal cortex) in the form of salience or priority maps (Itti & Koch, 2001; Fecteau & Munoz, 2006). Areas on this map with the highest neural activation will result in the system executing a saccade to the corresponding object. Anything on that map might then have some probability of encouraging a saccade. Evidence for this comes from findings that the presence of text in a static display or video tends to attract fixation even when the material is not useful (Wang & Pomplun, 2012; Ross & Kowler, 2013). Thus, it may also be possible that the presence of the graph or text region elicits saccades to those regions even when it is not necessarily necessary to do so.

Strategies of integrating across graph and text

In all 3 conditions the dominant initial strategy was to integrate at a high level rather than feature by feature. This was made evident by the average duration of initial visits to the graph or text, which were over 12 and 8 seconds for the graph and text, respectively. These durations were long enough to allow for a thorough examination of the graph and text. This meant that subjects tended to take a holistic approach of examining the entire graph or text before making a transition to the other region ration than doing a repetitive pairwise integration of each corresponding part of the graph and text. This strategy emerged throughout all 3 Perceptual-Motor conditions, with almost all of the graph or text being examined in the first visit, whereas reviewing (glances at previously examined areas or lines of the graph and text) was spread across the greater number visits in the Simultaneous condition rather than concentrated in the earlier visits as in the other two conditions. The integration strategy only changed in the Simultaneous condition in the frequency of transitions between graph and text and not in the areas examined and the distribution of time spent within the graph and text.

Eye movements during reading of graphs

When examining the pattern of transitions between areas of interest in the graph, the results were found to be consistent with prior observations on eye movements during the reading of graphs (Carpenter & Shah, 1998). In particular, it was clear that in all 3 Perceptual-Motor conditions, the most frequent transitions were from the bars (representing the data) to the axes labels and legend. This is in line with Carpenter and

Shah's model, in which the "referents" (areas of text) were fixated in order to integrate titles and other pieces of text with the information they correspond to. A high number of transitions between those referents could be due to a preference to utilize the visual display to access information that is either no longer or many or difficult to retrieve from it (Epelboim & Suppes, 2001).

Conclusions

Information gathering is essential for human life, though there are often mental and motor costs at play that influence how we may gather information exactly. We found that the sequential rather than Simultaneous presentation of a graph with associated text resulted in a decreased number of visits to the two regions. In this case, the cost of perceptual unavailability played more of a role than the motor cost due to the similarity in the average number of transitions in the Button-Press and Eye-Contingent conditions, which differed in the motor effort required to access the information. Thus, for human information gathering, even when the overall effort is modest and the times involved are short, the system takes the presumed cost (in this case, more perceptual than motor) of accessing the information into account. This may reflect a way of conserving neural operations (Lennie, 2003) or saving neural resources for more thinking about the content of the displays. Results are important both for understanding neural function as well as for designing visual displays so as to optimize information acquisition.

References

Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: The costs of choosing the optimal path. In *Vision Research* (Vol. 41, pp. 3613–3625). http://doi.org/10.1016/S0042-6989(01)00196-1

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory Representations in Natural Tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66–80. http://doi.org/10.1162/jocn.1995.7.1.66

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, *4*(2), 75–100. http://doi.org/10.1037/1076-898X.4.2.75

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective & Behavioral Neuroscience*, *15*(4), 837–53. http://doi.org/10.3758/s13415-015-0350-y

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research*, *41*(12), 1561–1574. http://doi.org/10.1016/S0042-6989(00)00256-X

Hooge, I. T., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, *58*(7), 969–76. http://doi.org/10.1017/CBO9781107415324.004

Kibbe, M. M., & Kowler, E. (2011). Visual search for category sets: tradeoffs between exploration and memory. *Journal of Vision*, *11*(3), 1–21. http://doi.org/10.1167/11.3.14

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *J Exp Psychol Gen*, *139*(4), 665–682. http://doi.org/10.1037/a0020198

Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497. http://doi.org/10.1016/S0960-9822(03)00135-0

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. In *Vision Research* (Vol. 41, pp. 3597–3611). http://doi.org/10.1016/S0042-6989(01)00203-6

Moher, J., & Song, J.-H. (2014). Perceptual decision processes flexibly adapt to avoid change-of-mind motor costs. *Journal of Vision*, *14*(8), 1-. http://doi.org/10.1167/14.8.1

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *46*(3), 461–488. http://doi.org/10.1037/h0084327

Ross, N. M., & Kowler, E. (2010). Eye movements while viewing captioned and narrated videos. *Journal of Vision*, *10*(7), 141–141. http://doi.org/10.1167/10.7.141

Stephens, D. W., Couzin, I., & Giraldeau, L. (2012). Ecological and behavioral approaches to search behavior. In Todd, P. M., Hills, T. T., & Robbins, T. W. (Eds.), *Evolution, algorithms, and the brain* (pp. 25-46). Cambridge, Massachusetts: The MIT Press.

Wang, H.-C., & Pomplun, M. (2012). The attraction of visual attention to texts in real-world scenes. *Journal of Vision*, *12*(6), 26–26. http://doi.org/10.1167/12.6.26

Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules

in human visual search. *Journal of Vision*, *13*(3), 10. http://doi.org/10.1167/13.3.10
Wu, C.-C., & Kowler, E. (2013). Timing of saccadic eye movements during visual search for multiple targets. *Journal of Vision*, *13*(11), 1–21. http://doi.org/10.1167/13.11.11.doi