THE INFORMATION THEORY OF VISION:
EVIDENCE FROM EYE-MOVEMENTS

BY

DEBORAH ANN CRONIN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Master's Committee:

    Associate Professor Alejandro Lleras, Chair
    Associate Professor Diane M. Beck

Abstract

Recent evidence suggests that target-dissimilar items that are typically regarded as irrelevant to reaction times during visual search do, in fact, affect reaction times (Buetti et al., in revision). This evidence suggests that the effect on reaction time of target-dissimilar items (lures) increases logarithmically as the number of lures present in a display increases. In response to these findings, Buetti & Lleras (in preparation) developed a model of visual search, the Information Theory of Vision (ITV). ITV uniquely proposes that the time it takes to perform the initial stage of visual search, during which likely non-targets (lures) are separated from possible targets (candidates), will increase with the amount of information present in the display. ITV further employs Information Theory (Shannon, 1948), Signal Detection Theory (Green & Swets, 1966), and Hick's Law (Hick, 1952) to support its predictions. In this study, we extend these predictions to eye-movements and find further support for ITV. Predictions of Guided Search (Wolfe, 1994) and Target Acqusition Model (Zelinsky, 2008) are also discussed.

Table of Contents

CHAPTER 1: INTRODUCTION

Visual search is a fundamental part of our daily lives. We perform dozens of searches in a typical day for objects as mundane as our keys or, if you are a radiologist, objects as critical as a tumor on a patient's scan. Understanding the factors contributing to performing an efficient, successful search are therefore of interest to the general public and to visual search professionals (e.g., radiologists, baggage screeners) alike.

Decades of research have yielded many theories describing how visual search unfolds. Common to many of these theories is an understanding that the time it takes to complete a search can be described as a linear function of the number of items to be searched through: as the number of items in a display increases, so too do search times. The slope of the search function is thought to indicate the efficiency of the search in terms of time spent processing each item. A display filled with target-similar items will result in steeper search slopes (indicating less efficient, more difficult search) than a display filled with target-dissimilar items (Treisman & Gelade, 1980). Feature Integration Theory (FIT; Treisman & Gelade, 1980), Guided Search (Wolfe, 1994), and Similarity Theory (Duncan & Humphreys, 1989) all subscribe to this understanding of the relationship between search times and the size of the search-set.

Bundesen (1990) suggested that the slope of the search function was a good measure of the amount of attentional processing a display required. Displays with more target-similar items or with more heterogeneous items require more attentional processing, and therefore result in steeper search slopes. Wolfe (1994) incorporated this idea into his Guided Search model, as well as ideas from FIT, Similarity Theory, and others. For the sake of simplicity, we will here explore the mechanism of Guided Search as a proxy

for its predecessors. In Guided Search, a display is processed by two stages. The first, a pre-attentive stage, develops feature maps for various display characteristics (color, orientation, etc.) and combines them into a priority or salience map. The salience map is a representation of where in the display there are items that share features with the target. Items that are highly target-similar yield high 'peaks' on this salience map. Other models of search have also employed a salience map to identify likely target locations or areas of interest (notably, Koch & Ullman, 1985; Itti & Koch, 2000).

In Guided Search, Stage Two processing begins after the completion of the salience map. During Stage Two, focused spatial attention inspects the objects indicated by the highest peaks on the salience map. This inspection occurs in order from most target-similar item to least-target similar item (that is, highest peak on the activation map to lowest peak) until the target is found or a certain threshold is passed and search is abandoned. The second stage, therefore, is the source of linear variation in reaction time, with more time needed for visual attention to inspect more target-similar items. The first stage is assumed to make a constant contribution to reaction time (400ms was used in simulations of the Guided Search model to represent both Stage 1 processing as well as all response-related processing; Wolfe, 1994).

Importantly, not all items must be inspected by focused attention in order for a search to be completed. In Guided Search, search is abandoned after the target is found or if all of the salience map peaks falling above a certain threshold have been inspected. Neider and Zelinski (2008) proposed that any items whose peaks fall below threshold for inspection should be excluded from the search set size, leaving only the target-similar items in the 'functional set-size'. This distinction is particularly important in real-scene

search where there can be hundreds of objects in an image. If we use the typical measure of search efficiency (the slope of the RT X set-size function), we would find extremely efficient search (~9ms/item; Wolfe, et al., 2011). The visual system cannot plan and execute an eye-movement to a new object in 9ms, let alone process the current item at fixation. In fact, the time it takes to process the category of an item has been estimated at ~75-80ms (VanRullen & Thorpe, 2001). As Neider and Zelinski (2008) suggest, per-item processing times are more realistic if we plot reaction time as a function of the number of items likely to be inspected by focused spatial attention while in search for a specific target as opposed to the scene's total set-size.

This modified measure of search efficiency seems reasonable as long as Wolfe's (1994) assumption, that Stage 1 processing occurs at a constant rate, holds true. Recently, our lab called this assumption into question by proposing a new theory of visual search, the Information Theory of Vision (ITV; Buetti et al., in revision; Buetti & Lleras, in preparation). Like other theories of visual search, ITV predicts a linear component of the search function that is driven by the scrutiny of target-similar items (that is, items that are part of Neider and Zelinski's functional set-size). A major point where ITV differs from other models of visual search is in its description of a second source of variation in RTs—ITV predicts the target-dissimilar items that are never (or rarely) inspected by focused attention do in fact affect RT in a systematic fashion. Data from our lab suggests and ITV predicts that instead of a flat cost to screen out target-dissimilar items during the first stage of processing, items that are highly dissimilar to the target affect reaction times logarithmically. That is, reaction times increase logarithmically as a function of the number of these target-dissimilar items in the display. ITV is not the first model to emphasize the effects of

3

distractor-target similarity on RTs (Duncan & Humphreys, 1989), but it is the first to propose a *logarithmic* contribution to RT from target-dissimilar distractors that arises during Stage 1 processing (not Stage 2). In terms of real-world scene search, a logarithmic contribution to reaction time of definite non-targets offers an explanation for why a scene with 10x more items will not have a RT 10x as long (e.g., Wolfe et al., 2011).

The current paper seeks to further support ITV with evidence from eye-movements using lab-developed stimuli.

The Information Theory of Vision

Like many models of visual search, ITV describes search in terms of two stages. The first stage is an unlimited capacity, resolution-limited process that evaluates the target-similarity of all display elements in parallel. Because this 'Screening' stage is resolution-limited, it is only able to screen out distractors that are sufficiently target-dissimilar. We refer to these target-dissimilar distractors as 'lures'. Lures that are very dissimilar to the target are discounted quickly, while lures that are more target-similar are discounted more slowly, leading to longer RTs. If any items are too similar to the target to be discounted by the Screening stage, information about their locations is passed on to the second stage. The second stage is a resolution-unlimited, but limited capacity process that scrutinizes each of the locations passed on by Stage 1 individually and in a random order until the target is found or all locations have been inspected. Due to their high similarity to the target, we term the distractors passed to this 'Scrutiny' stage 'candidates'.

During the Screening stage of ITV, information is gathered about each location in the display in parallel. ITV assumes that this information gathering occurs in the presence of noise, making vision a signal detection problem (Green & Swets, 1966)—and that the more

4

information that exists in a display (number of items, heterogeneity of lures, etc.), the more time it will take to process that display (Hick's Law; Hick, 1952). Using these assumptions we can make several predictions about behavior. First, reaction times should increase with increasing search set-sizes. While this is not a unique prediction of ITV, Buetti and colleagues (in revision) have found RT increases by increasing the number of lures in a display even in displays that would classically be thought of as eliciting 'pop-out' search. Furthermore, ITV predicts that RTs will increase logarithmically as the number of lure elements increases *and* that the slope of the logarithm will be influenced by how similar those lure items are to the target. Buetti and colleagues confirmed these predictions. As the number of target-dissimilar lures increased, RTs increased logarithmically. The slope of that logarithm was manipulated by altering the similarity of the lure items to the target (e.g., using orange lures will result in a shallower log slope than using red lures if a participant is searching for a red target; Buetti et al., in revision, Experiments 3A & 3C).

ITV can also make several predictions about eye-movement behavior. First, if the delay between the onset of the display and the first eye-movement (the initial saccade latency; ISL) can be considered a measure of Stage 1 processing times, displays containing more target-similar lures should result in longer ISLs than displays containing less target-similar lures as it would take more time for Stage 1 to collect enough information to reject target-similar lures as non-targets. Second, as in any signal-detection problem, there is always a possibility for false alarms. In terms of ITV, a false alarm would be a case in which a lure's location was passed from Stage 1 to Stage 2 for inspection. ITV predicts that false alarm eye-movements to lures would occur more frequently during inspection of displays containing lures similar to the target than in displays containing lures that are dissimilar to

the target. These false alarm eye-movements should result in more fixations overall and longer total scan-paths in displays with target-similar lures. ITV also assumes that the items passed to Stage 2 are inspected in a random order. Because we have not yet directly tested the effect of candidate similarity on Stage 2 processing, assuming random inspection is most parsimonious—further study may demonstrate that candidates can be processed in order of target-similarity. However, this is irrelevant to the present study as all candidate items are equally similar to the target. If candidates in the present experiment are inspected in a random order until the target is found (that is, in a random, serial, self-terminating manner), the number of candidates inspected on average should equal approximately $(n+1)/2$, where n is the number of candidates in the display. Finally, the overall number of fixations and the total length of scan paths should increase with increasing number of items in the display and with increasing lure-target similarity. As the number of items in the display increases and as the lure-target similarity increase, so does the amount of information and the amount of uncertainty in the display. More eye-movements (and therefore, longer scan paths) would be needed to resolve this uncertainty.

In the present study, we first sought to replicate the results of Experiments 3A and 3B of Buetti and colleague's behavioral study while simultaneously gathering information about eye-movements. The result is three experiments—Experiments 1 and 2 replicate Buetti and colleagues' Experiment 3A and 3B, respectively while Experiment 3 is a within-subjects manipulation using the stimuli of Experiments 1 and 2. This within-subjects manipulation allowed us to compare the effects of target-lure similarity with more power than a between-subjects comparison of the data from Experiments 1 and 2.

CHAPTER 2: EXPERIMENT 1

Experiment 1 is a replication of Buetti and colleagues (in revision) experiment 3A. In this experiment, participants searched for a red 'T' amongst a number of red 'L' shaped candidate items and orange, thick weighted, cross-shaped lure items. This experiment further served as the basis for a power analysis, in conjunction with Experiment 2, to determine the number of subjects necessary for a within-subjects comparison of the effects of different lure-types (Experiment 3).

Method

Participants. Twenty undergraduate students from the University of Illinois at Urbana-Champaign participated in exchange for course credit. All participants reported normal or corrected-to-normal vision. Five subjects were replaced due to low accuracy.

Stimuli and apparatus. Stimuli consisted of arrays of letters and shapes. Each display contained one red 'T' and three or seven red 'L's. These letters subtended 1.6 degrees of visual angle horizontally and 1.7 degrees of visual angle vertically. On each trial, subjects were asked to search for the T and respond to its orientation (tilted 90 degrees to the left or 90 degrees to the right). L's were oriented in one of four possible ways—upright, tilted 90 degrees to the left or right, and upside-down.

In addition to the letter items, some displays also contained thick orange cross shapes ('+') that were of the same dimensions as the letters in degrees of visual angle. The weight of the crosses' line segments was substantially thicker than the weight of the letter's line segments (see Figures 1 & 5). There were trials with no crosses present (letters only), or with 4, 8, 16, or 28 crosses intermixed with the letters. In this way, there were 10 possible display conditions (four letters with 0, 4, 8, 16, or 28 crosses, and eight letters with

0, 4, 8, 16, or 28 crosses). Examples of candidate-only and candidate+lure displays can be found in Figure 1. Participants viewed 30 displays per condition yielding a total of 300 trials.

On each trial, all items (letters and crosses) were assigned random locations on a 36-point grid (subtends 25.4 degrees of visual angle horizontally and 27 degrees of visual angle vertically). The only constraint on item placement was that each quadrant of the screen could only contain one letter item in the 4-letter conditions and only two letters in the 8-letter conditions. Forcing candidates to be spread throughout the display decreased the likelihood that the red candidate items would form a perceptual group that might guide participants' gaze to the target on a subset of trials.

All displays were presented on a 22inch CRT monitor with a refresh rate of 185Hz at a resolution of 1024 x 768 pixels. Stimuli were presented using the Psychophysics Toolbox extention for Matlab (Brainard, 1997; Pelli, 1997). Stimuli were viewed from a distance of 55cm. Eye-movements were measured during each trial with an EyeLink 1000 eye-tracking system (SR Research, Inc.) in the tower configuration. This eye-tracker configuration prevents head movements during the experiment. Eye-movements were sampled at a rate of 1000Hz.

Design and procedure. On each trial, a fixation cross appeared to signal the start of the trial. The search display onset after approximately 1 second and participants were allowed to freely move their eyes in search of the target. Once the target was found, participants pressed a key indicating that the T was tilted to the left or tilted to the right. This key press ended the trial. If the participant did not make a response within 5 seconds

of the onset of the display, the trial ended automatically with their response recorded as an error.

Data preparation. Before analysis, participants' accuracy scores were assessed. If a participant's accuracy fell below 90% for the entire experiment or for any of the trial types, their data was excluded from the analysis and a new, naive subject was run to replace their data. Five subjects' data were replaced using these criteria. For the remaining subjects' data, only correct trials were analyzed.

Results

Behavioral results. The behavioral data replicate the findings of Buetti and colleagues' (in revision) Experiment 3A: A two-way ANOVA on correct reaction times with number of candidates (4 or 8) and number of lures (0, 4, 8, 16, or 28) as within-subjects factors revealed a significant main effect for number of candidates, $F(1,19)=284.86$, $p<0.001$, $\eta^2=0.380$, and for number of lures, $F(4,76)=26.229$, $p<0.001$, $\eta^2=0.153$. The interaction was not significant, $F(4,76)=1.368$, $p=0.253$, $\eta^2=0.006$. These results indicate that RTs increased as the number of candidates and number of lures increased.

Further replicating Buetti and colleagues' Experiment 3A, when the RT data was plotted as a function of the number of non-target items, the best-fit line was a logarithmic function for both the four- ($R^2=0.959$) and eight- ($R^2=0.992$) candidate conditions (Figure 2). These results are, therefore, consistent with ITV's description of a screening function, suggesting that during this experiment, participants were able to 'screen' or filter out the orange non-targets (Buetti et al, in revision).

Eye-tracking results. Eye movement data was analyzed for the initial saccade latency, total scan path length, and the number of fixations on different item types.

Initial saccade latencies. Initial saccade latencies were calculated by subtracting the search display onset time from the time the first saccade started. Average values for each trial type are listed in Table 1. A two-way ANOVA on this data using within-subject factor of candidate number (4 or 8) and lure number (0, 4, 8, 16, or 28) yielded no effect of candidate number, $F_{(1,19)}=4.064$, $p=0.422$, lure number, $F_{(2.452,46.586)}=2.676$, $p=0.367$, or an interaction, $F_{(4,76)}=1.182$, $p=0.243$, (Mauchly's Test indicated that the assumption of sphericity had been violated ($\chi^2(9)=28.73$, $p=0.001$); Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.613$)). The number of lures and candidates in the display may have had no effect on the amount of time it took subjects to initiate an eye-movement away from fixation.

Scan path length. Each participants' saccade amplitudes for a given trial were summed, then averaged across conditions to obtain the total scan path length. A two-way ANOVA on the scan path length with number of candidates (4 or 8) and number of lures (0, 4, 8, 16, 28) as within-subject factors revealed significant main effects of candidate number, $F_{(1,19)}=167.29$, $p<0.001$, $\eta^2=0.382$, and lure number, $F_{(4,76)}=10.15$, $p<0.001$, $\eta^2=0.101$. The interaction was not significant, $F_{(4,76)}=1.46$, $p=0.22$. Unsurprisingly, participants moved their eyes further in displays containing more candidates and in displays containing more lures (Figure 3, see mean values in Table 1). However, the best-fit line to these data was logarithmic. While we initially predicted that scan-paths would increase with set-size, we did not anticipate logarithmically increasing scan-paths based on ITV. However, logarithmically increasing scan-paths are in no way inconsistent with ITV. To the contrary, this effect may reflect the same process driving the logarithmic RT functions: as the number of lures in the display increases, the number of items that can be handled by Stage

1 processing also increases, leading to a decelerating increase in the number of eye-movements required to process a display, and therefore a decelerating increase in scan path length. In the same vein, ITV would also predict that increasing the number of candidates should cause linear increases in scan path lengths, but this prediction was not systematically tested in the present study.

   Number of fixations. Fixation data was analyzed in terms of the total number of fixations per trial, the total number of times a candidate was fixated per trial, and the total number of lure fixations per trial. Participants were considered to have fixated an item if the item fell within a radius of 2.5° from the center of their fixation. Decreasing the size of this window to a 1° radius did not change the pattern of the effects.

   A two-way ANOVA on the total number of fixations with candidate number (4 vs 8) and lure number (0, 4, 8, 16, or 28) as within-subject factors resulted in main effects of candidate number, $F(1,19)=274.84$, $p<0.001$, $\eta^2=0.368$, and lure number, $F(4,76)=27.46$, $p<0.001$, $\eta^2=0.184$. The interaction approached significance, $F(4,76)=2.298$, $p=0.067$, $\eta^2=0.010$. These results are consistent with the results of the scan path analysis—more fixations were made in displays containing more candidates and in displays containing more lures (see mean values in Table 1). This finding is qualified by the presence of a marginal interaction. While this interaction is not statistically significant and may therefore be an anomaly, it is also possible that an interaction between lure and candidate number may be due to crowding or texture effects that occur only at large set-sizes of candidates and lures. Interestingly, there is no interaction present in the scan-path analyses, which are necessarily highly correlated with the number of fixations, or in the RT analyses, suggesting that a display totally filled by search items may require more, but shorter fixations to

resolve local features than a less-full display (or, again, that the marginal interaction is truly non-significant). Fitting with ITV, the best fit line to the fixation data was again logarithmic. The logarithmic relationship between number of fixations and set-size is consistent with our finding that scan path lengths increase logarithmically with set-size and is consistent with ITV for the same reasons.

A two-way ANOVA with candidate number (4 vs 8) and lure number (0, 4, 8, 16, or 28) as within-subjects factors was performed on the number of candidates fixated per trial. Mean values for the number of candidates fixated per trial type are listed in Table 1. For this analysis, candidates included the red L items and the target item. Participants made more candidate fixations in displays containing more candidates, $F(1,19)=627.4$, $p<0.001$, $\eta^2=0.884$, and more lures, $F(4,76)=13.423$, $p<0.001$, $\eta^2=0.150$. The interaction was also significant, $F(4,76)=6.065$, $p<0.001$ $p=0.902$, $\eta^2=0.61$.

A two-way ANOVA with candidate number (4 vs 8) and lure number (0, 4, 8, 16, 04 28) as within-subjects factors was also performed on the number of lures fixated per trial. Mean values for the number of lures fixated per trial type are listed in Table 1.[1] Mauchly's Test indicated that the assumption of sphericity had been violated for lure number ($\chi^2(5)=39.51$, $p=0.001$) and the interaction ($\chi^2(5)=23.074$, $p=0.001$). Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=0.661$, $\varepsilon=.471$, respectively). The ANOVA indicated significant main effects of candidate number, $F(1,19)=112.017$, $p<0.001$, $\eta^2=0.278$, and lure number, $F(1.41,26.83)=632.19$, $p<0.001$, $\eta^2= 0.923$. The interaction was also significant, $F(1.98, 37.66)=24.55$, p<0.001, $\eta^2= 0.161$.

[1] Note that the number of lures fixated will increase by necessity as the number of lures present in the display increases. The more lures in the display, the more likely lures are to be present near enough to candidates and to each other to be considered 'fixated' in this analysis. As such, the number of lures fixated often exceeds the total number of fixations.

Participants made more lure fixations in displays containing more candidates and more lures. When considering this result, it is important to remember that the density of the display at large set-sizes was such that there were often multiple items within the area we considered a participant's fixation. In this experiment, candidates were purposefully spread throughout the display, but lures were not, meaning that several lures could appear close enough to each other or near enough to a candidate to be counted as 'fixated' within the same fixation. This may artificially inflate the number of lures considered fixated. However, regardless of the absolute number of lures fixated, it is clear from this data that participants were fixating lures with some frequency.

Discussion

Experiment 1 successfully replicated the RT effects demonstrated by Buetti et al. (in revision). As the number of lures increased in the display, so did RTs, even as the number of candidates stayed constant. Furthermore, this RT effect was best fit with a logarithmic curve. The logarithmic contribution to RT by the presence of lures in a display is uniquely predicted by ITV.

In terms of eye-movements, we found no significant effects of candidate number, lure number, or an interaction on ISL, indicating that ISL may not be a good measure of processing time in this task. Other groups, including Pomplun, Garaas, & Carrasco (2013) have found increasing ISL with increasing set size during visual search tasks. However, Pomplun and colleagues also found an effect of task difficulty on ISL. In their experiments, a more difficult search task produced shorter ISL than an easier search task. In the present experiment, it is possible that effects of set-size and task-difficulty were confounded: If our large set-size displays were substantially more difficult to search through than our small

set-size displays, any effect of set-size may have been masked by a task difficulty effect. Search may have been more difficult in larger set-size displays for a variety of reasons, including crowding, increased uncertainty when choosing a saccade target, etc. At any rate, it is possible that ISL may be a good measure of Stage 1 processing time in some display types (such as the displays used by Pomplun and colleagues), but this was certainly not the case in the present study.

As anticipated, we found evidence that increasing the number of lures and the number of candidates in the display affected the total distance traveled by the eyes (scan path lengths) and the total number of fixations made. Through the lens of ITV, this increase in the number of fixations and length of scan paths is consistent with the increasing uncertainty present in a display as we increase the number of items. Clearly, more eye-movements and fixations are required to resolve the uncertainty of these displays. Furthermore, we found logarithmic best-fit lines for the scan-path data and the total-fixation data. Because eye-movements are an important component of reaction times, it is reassuring that the pattern of results found in RT is also present in the fixation and scan-path data. In combination with the null-result for ISL, these logarithmic effects of fixation number and scan-path length suggest that the discounting of distractors throughout the display (across eye-movements) is driving the logarithmic effects seen in RT. Taken with Buetti and colleagues' (submitted with revisions) evidence for Stage 1 processing being best described by a logarithmic function and Stage 2 processing being best described by a linear function, the present results suggest that with Stage 1 processing occurring across eye-movements.

The number of candidates fixated in this experiment did not fall in line with ITV's predictions. ITV predicts that the average number of candidates inspected should be equal to $(n+1)/2$, where n is the number of candidates. This would represent a random inspection of the items passed to Stage 2. In 4-candidate displays, ITV predicts an average of 2.5 candidate fixations, and 4.5 candidate fixations in 8-candidate displays. Instead, we see evidence of candidate fixations increasing with both candidate number (as anticipated) and lure number (unanticipated). Furthermore, participants inspected fewer than 2.5 candidates per trial when there were 4 candidates in the display, meaning participants were not inspecting candidates at random in the 4 candidate displays. At first pass, this suggests that some information about target-similarity is passed on to Stage 2, allowing participants to bias scrutiny towards some candidate items. This explanation would be consistent with predictions of Guided Search (Wolfe, 1994), but not with ITV (Buetti et al., submitted with revisions). However, in displays containing 8 candidates, the number of candidate fixations made were more consistent, though slightly larger than what would be predicted by ITV, and substantially larger than the number predicted by Guided Search (Table 1). Furthermore, in both Guided Search and ITV, the number of lures in the display should not affect the number of inspections of candidates—participants should be able to effectively discard lures from consideration before inspecting the candidates. Instead, the effect of lure number on candidate fixations in both display types suggests that participants made more candidate fixations in the presence of more lures (Table 2). Participants also fixated lures on all trial types. ITV predicts some false alarm inspections of candidates, but this pattern of effects could also be more evidence that Stage 1 processing is occurring across eye-movements: If not all lures are discarded before the first eye-movement is

15

made, or even on the first subsequent fixation, more fixations may be necessary to complete Stage 1 processing before Stage 2 can proceed full force. These fixations will likely focus on both candidates to-be-inspected and lures that have not yet been discarded as irrelevant.

Figures and Tables

Table 1

*Average values of ISL, Scan Path Length, and Number of Fixations in Experiment 1*

	Four Candidate Displays				
Means	No Lures	4 Lures	8 Lures	16 Lures	28 Lures
ISL in ms	254.32	234.94	239.91	236.34	233.53
(SD)	(31.62)	(26.18)	(28.54)	(34.53)	(30.13)
Scan Path Length	23.60°	26.18°	25.03°	26.43°	27.84°
(SD)	(4.44)	(6.06)	(6.03)	(4.81)	(6.37)
# Fixations	4.07	4.64	4.62	4.82	5.12
(SD)	(0.41)	(0.71)	(0.80)	(0.70)	(0.92)
# C. Fixations	1.43	1.71	1.64	1.77	1.79
(SD)	(0.31)	(0.51)	(0.35)	(0.40)	(0.42)
# L. Fixations		2.42	4.17	7.97	14.22
(SD)	N/A	(0.51)	(0.82)	(1.34)	(2.50)

|	Eight Candidate Displays				
	No Lures	4 Lures	8 Lures	16 Lures	28 Lures
ISL in ms	253.24	243.65	243.97	236.60	243.73
(SD)	(34.69)	(34.08)	(34.18)	(31.56)	(30.15)
Scan Path Length	30.54	32.95	33.12	36.12	37.30
(SD)	(5.44)	(7.32)	(8.10)	(7.33)	(6.93)
# Fixations	5.16	5.55	5.75	6.17	6.33
(SD)	(0.63)	(0.94)	(1.19)	(1.01)	(0.97)
# C. Fixations	4.48	4.68	4.87	5.49	5.40
(SD)	(0.63)	(0.98)	(1.03)	(1.23)	(0.89)
# L. Fixations		2.60	4.88	10.09	17.21
(SD)	N/A	(0.57)	(0.93)	(1.73)	(2.78)
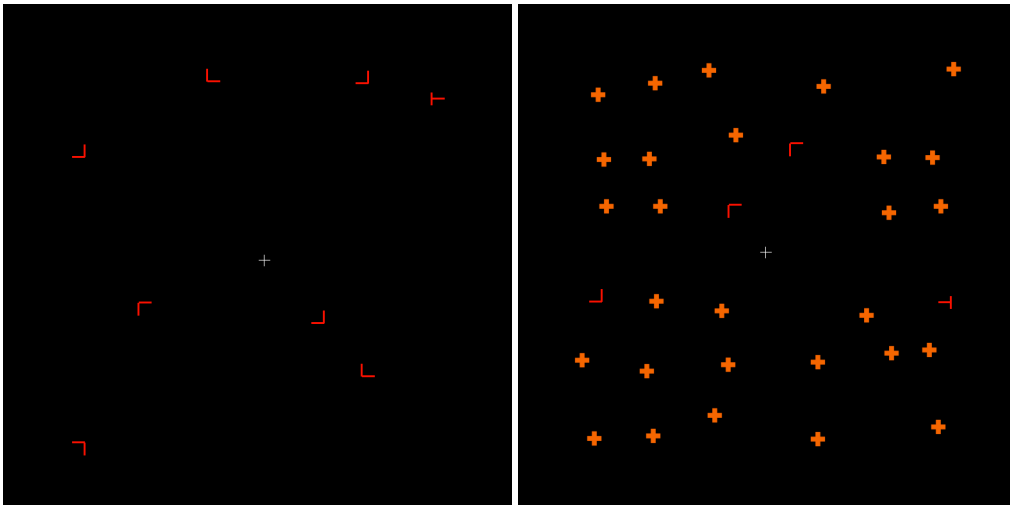
Figure 1. Sample stimuli used in Experiment 1. On the left is a display containing 8 candidate items. On the right is a display containing 4 candidates and 28 lure items. Participants were asked to respond to the orientation of the T.
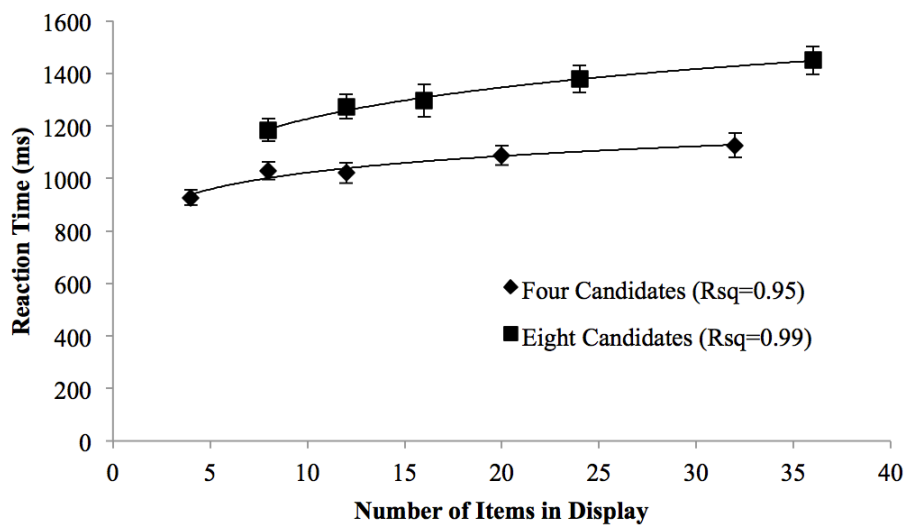
Figure 2. Reaction times increased as the number of candidate items and as the number of lure items in the display increased. The best-fit line was a logarithmic function of set-size.
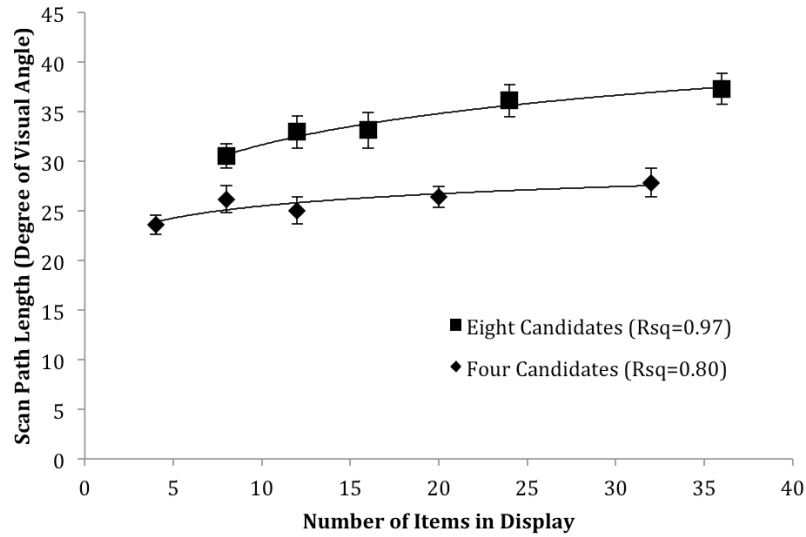
Figure 3. Scan path length increased as the number of candidate items and as the number of lure items in the display increased. The best-fit line for both candidate set-size functions was logarithmic.
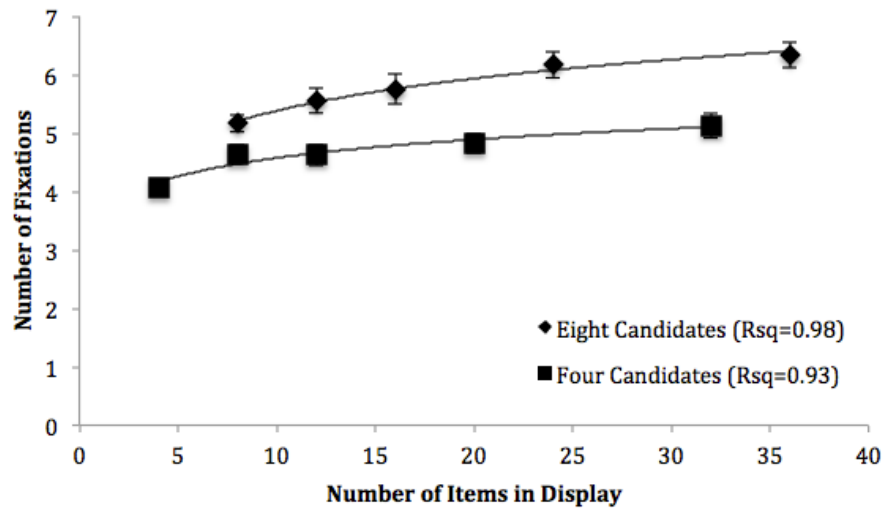
Figure 4. Total number of fixations increased as the number of candidate items and as the number of lure items in the display increased. The best-fit line for both candidate set-size functions was logarithmic.

Chapter 3: Experiment 2

ITV states that as lure-target similarity increases, so too do the lures' effects on performance. Therefore, we replicated Experiment 1 replacing the thick-orange cross lures with lures that were more similar to the target—thin orange crosses (Figure 5). These thinly-weighted lures are more target-similar in appearance than the thick-weighted lures used in Experiment 1 and therefore, should produce the same pattern of effects produced by the thick-crosses, but to a significantly greater degree. A between-subjects comparison was used to confirm this prediction. Experiment 2 also constitutes a replication of Buetti and colleagues (submitted with revisions) Experiment 3B with the addition of eye-tracking methods.

Method

Participants. Twenty naïve undergraduate students from the University of Illinois at Urbana-Champaign participated for course credit. All participants reported normal or corrected-to-normal vision.

Stimuli and apparatus. The stimuli and apparatus for this experiment were identical to those of Experiment 1, save for one change: the thick-cross distractor stimuli from Experiment 1 were replaced with thinly weighted crosses (Figure 5).

Design and procedure. The design and procedure were identical to that of Experiment 1.

Data preparation. No subjects in this experiment had accuracy below 90%. Only correct trials were analyzed.

Results

    Behavioral results. The behavioral data replicate the findings of Buetti and colleagues' (submitted with revisions) Expeirment 3B: A two-way ANOVA with number of candidates (4 or 8) and number of lures (0, 4, 8, 16, or 28) as within-subjects factors was performed on RT data. Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2(9)$=22.28, p=0.008). Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (ε=.615). Main effects of number of candidates, F(1,19)=212.53, p<0.001, η^2=0.300, and for number of lures, F(2.46,46.73)=81.34, p<0.001, η^2=0.459. The interaction approached significance, F(4,76)=2.45, p=0.054, η^2=0.013. These results indicate RTs increased as the number of candidates and number of lures increased (Figure 6). The presence of an interaction between candidate number and lure number may indicate that some of the thin-orange cross stimuli were passed on to Stage 2 processing as candidates, instead of being discounted during Stage 1. This hypothesis is supported by lure-fixation data (below), as more lures were fixated as the number of items in the display increased.

    Again replicating Buetti and colleagues' Experiment 3B, when the RT data was plotted as a function of the number of non-target items, the best-fit line was a logarithmic function for both the four- (R^2=0.968) and eight- (R^2=0.957) candidate conditions (Figure 6). These results are, therefore, consistent with ITV's description of a screening function, suggesting that during this experiment, participants were able to 'screen' or filter out the orange non-targets (Buetti et al).

    Eye-tracking results. Eye movement data was analyzed for the initial saccade latency, total scan path length, and the number of fixations on different item types.

Initial saccade latencies. Average initial saccade latency values per trial type are listed in Table 2. A two-way ANOVA with within-subjects factors of candidate number (4 or 8) and lure number (0,4,8,16, or 28) was performed on initial saccade latency data. There were no significant effects of candidate number, $F(1,19)=0.085$, $p=0.773$, lure number, $F(4,76)=1.023$, $p=0.401$, or an interaction, $F(4,76)=0.569$, $p=0.686$. Like in Experiment 1, the number of items in the display had no effect on the time it took subjects to initialize their first eye-movement.

Scan path length. Each participants' saccade amplitudes for a given trial were summed, then averaged across conditions to obtain the total scan path length. Mean values are listed in Table 2. A two-way ANOVA on the scan path length with number of candidates (4 or 8) and number of lures (0, 4, 8, 16, 28) as within-subject factors revealed significant main effects of candidate number, $F(1,19)=83.43$, $p<0.001$, $\eta^2=0.264$, and lure number, $F(4,76)=33.59$, $p<0.001$, $\eta^2=0.223$. The interaction was not significant, $F(4,76)=1.29$, $p=0.28$. Participants moved their eyes further in displays containing more candidates and in displays containing more lures (Figure 7). Consistent with Experiment 1 and ITV's predictions, the best-fit line was again logarithmic.

Number of fixations. Fixation data was analyzed in terms of the total number of fixations per trial, the total number of times a candidate was fixated per trial, and the total number of lure fixations per trial. Participants were considered to have fixated an item if the item fell within a radius of 2.5° from the center of their fixation. Decreasing the size of this window to a 1° radius did not change the pattern of the effect.

A two-way ANOVA with candidate number (4 vs 8) and lure number (0, 4, 8, 16, or 28) as within-subject factors was performed on the total fixations data. Mauchly's test

revealed that the assumption of sphericity was violated ($\chi^2(9)$=19.97, p=0.019). Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (ε=.604). Mean values for total number of fixations are listed in Table 2. The ANOVA indicated main effects of candidate number, $F(1,19)$=114.64, p<0.001, η^2=0.305, and lure number, $F(2.42,45.88)$=60.075, p<0.001, η^2=0.333. The interaction was not significant, $F(4,76)$=.968, p=0.430. These results are consistent with the results of the scan path analysis—more fixations were made in displays containing more candidates and in displays containing more lures. The best-fit line was logarithmic, consistent with Experiment 1 and ITV's predictions.

A two-way ANOVA with candidate number (4 vs 8) and lure number (0, 4, 8, 16, or 28) as within-subjects factors was performed on the number of candidates fixated per trial. Mean values for the number of candidates fixated per trial type are listed in Table 2. For this analysis, candidates included the red L items and the target item. Mauchly's Test indicated that the assumption of sphericity had been violated ($\chi^2(9)$=17.11, p=0.048). Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (ε=.659). Participants made more candidate fixations in displays containing more candidates, $F(1,19)$=526.29, p<0.001, η^2=0.842, and more lures, $F(2.637, 50.100)$=23.789, p<0.001, η^2=0.227. The interaction was also significant, $F(4,76)$=5.308, p=0.001, η^2=0.048, indicating that subjects made more candidate fixations in displays containing many lures and 8 candidates than in displays containing many lures and only 4 candidates, as in Experiment 1.

A two-way ANOVA with candidate number (4 vs 8) and lure number (0, 4, 8, 16, or 28) as within-subjects factors was also performed on the number of lures fixated per trial.

Mean values for the number of lures fixated per trial type are listed in Table 2. Mauchly's

Test indicated that the assumption of sphericity had been violated (Lure number:

$\chi^2(5)=64.05$, $p<0.001$; Interaction: $\chi^2(5)=21.30$, $p=0.001$). Degrees of freedom were

corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.395$ and $\varepsilon=.647$

respectively). Significant main effects of candidate number, $F(1,19)=41.721$, $p<0.001$,

$\eta^2=0.161$, lure number, $F(1.184, 22.495)=423.367$, $p<0.001$, $\eta^2=0.897$, and an interaction,

$F(1.941, 36.873)=10.910$, $p<0.001$, $\eta^2=0.068$ were found. Participants made more lure

fixations in displays containing more candidates and more lures.

Discussion

Experiment 2 again replicated the behavioral findings of Buetti et al. (submitted

with revisions). As in Experiment 1, RTs increased with increasing numbers of lures and

candidates in a display. The overall RT function was best fit by a log, again supporting ITV's

assertion that the presence of lures in a display affects reaction time in a logarithmic

manner.

The eye-movement results of Experiment 2 are similar to Experiment 1. We found

no effect of candidate or lure number on initial saccade latency, suggesting once again that

initial saccade latencies are not a good measure of Stage 1 processing times in this task.

Participants' scan path length and total fixations increased with increasing set-sizes of

candidates and lures. This suggests that more eye-movements were necessary to resolve

the uncertainty of displays with more items. Finally, participants made more candidate

fixations in displays containing more candidates and in displays containing more lures. The

average values for number of candidate fixations fall more in line with ITV's predictions

(2.5 candidate fixations for 4 candidate displays and 4.5 for 8 candidate displays) than

those of Experiment 1 (Table 2), which may suggest that candidates are inspected by Stage 2 in random order under some circumstances. One possibility is that, when the items surrounding a candidate are less target-similar, as in Experiment 1, rapid screening and scrutiny of items further from fixation may be possible due to low levels of visual-interference from nearby lures. In cases such as Experiment 2, where the lures shared a target feature (line weight), the presence of the lures may prevent such long-range screening and scrutiny, forcing the eyes to move around the display (Chang & Rosenholtz, 2014).

Between-Subjects Comparison of Experiments 1 and 2. A critical claim made by ITV is that target-similar lures require more processing by Stage 1 than target-dissimilar lures. Therefore, comparing the results of Experiments 1 and 2 should provide evidence that thin cross lures have a greater effect on RTs and the eye-tracking measures than thick cross lures. Indeed, participants exhibited significantly longer reaction times, $F(1,38)=5.652$, $p=0.023$, $\eta_p^2 =0.095$, longer scan paths, $F(1,38)=6.638$, $p=0.014$, $\eta_p^2 =0.149$, and made more candidate fixations, $F(1,38)=5.290$, $p=0.027$, $\eta_p^2= 0.122$, in thin-cross displays than in thick-cross displays. Participants were also more likely to fixate thin-cross lures than thick-cross lures, $F(1,38)=9.714$, $p=0.003$, $\eta_p^2= 0.122$, which may indicate that participants were more likely to pass thin crosses on to Stage 2 processing than thick crosses due to their increased target similarity. The between-subject effect of total number of fixations was marginally significant, $F(1,38)=3.606$, $p=0.065$, with more fixations occurring in displays containing thin crosses than displays containing thick crosses. There was no between-subjects effect of initial saccade latencies, $F(1,38)=1.142$, $p=0.290$.

Table 2

Average values of ISL, Scan Path Length, and Number of Fixations in Experiment 2

Means	Four Candidate Displays				
	No Lures	4 Lures	8 Lures	16 Lures	28 Lures
ISL in ms	242.25	232.69	231.65	237.03	238.84
(SD)	(42.74)	(43.10)	(39.64)	(41.76)	(41.80)
Scan Path Length	24.70°	31.70°	31.43°	31.95°	34.15°
(SD)	(6.56)	(7.63)	(6.85)	(5.71)	(6.01)
# Fixations	4.05	5.06	5.25	5.43	5.94
(SD)	(0.59)	(0.78)	(0.81)	(0.79)	(1.04)
# C. Fixations	1.01	1.34	1.50	1.51	1.66
(SD)	(0.22)	(0.40)	(0.40)	(0.46)	(0.49)
# L. Fixations		2.57	4.65	9.33	18.03
(SD)	N/A	(0.66)	(1.06)	(1.66)	(3.18)

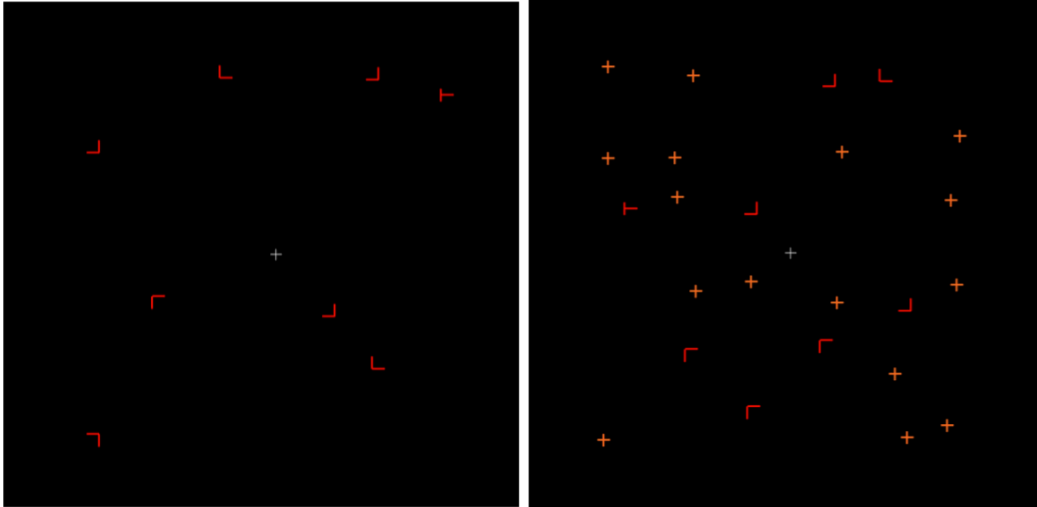|	Eight Candidate Displays				
	No Lures	4 Lures	8 Lures	16 Lures	28 Lures
ISL in ms	248.27	233.47	236.34	239.03	242.05
(SD)	(46.21)	(39.70)	(37.91)	(41.97)	(48.39)
Scan Path Length	31.52	37.55	33.90	41.36	42.74
(SD)	(5.29)	(8.66)	(6.54)	(9.50)	(9.72)
# Fixations	5.19	6.07	6.27	6.73	6.96
(SD)	(0.53)	(1.02)	(0.95)	(1.34)	(1.35)
# C. Fixations	3.50	4.24	4.60	4.80	4.88
(SD)	(0.86)	(0.86)	(0.90)	(1.25)	(1.27)
# L. Fixations	N/A	2.89	5.67	11.32	21.00
(SD)		(0.76)	(0.92)	(2.71)	(4.78)

Figure 5. Sample displays from Experiment 2 containing 8 candidates and no lures (left) and 8 candidates and 16 thin orange cross lures.
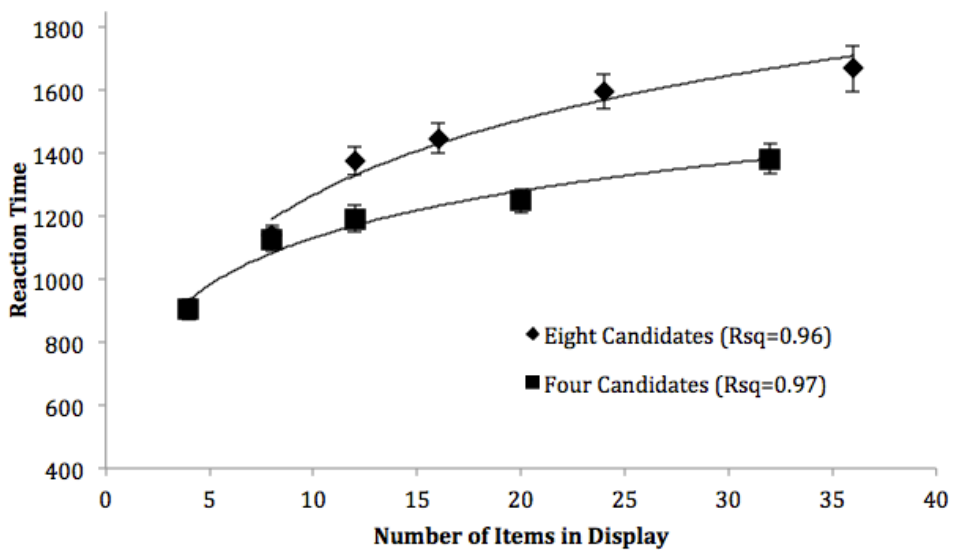

Figure 6. Participants responded slower in displays containing more candidates and more lures. The best-fit line to both candidate set-size functions was logarithmic.
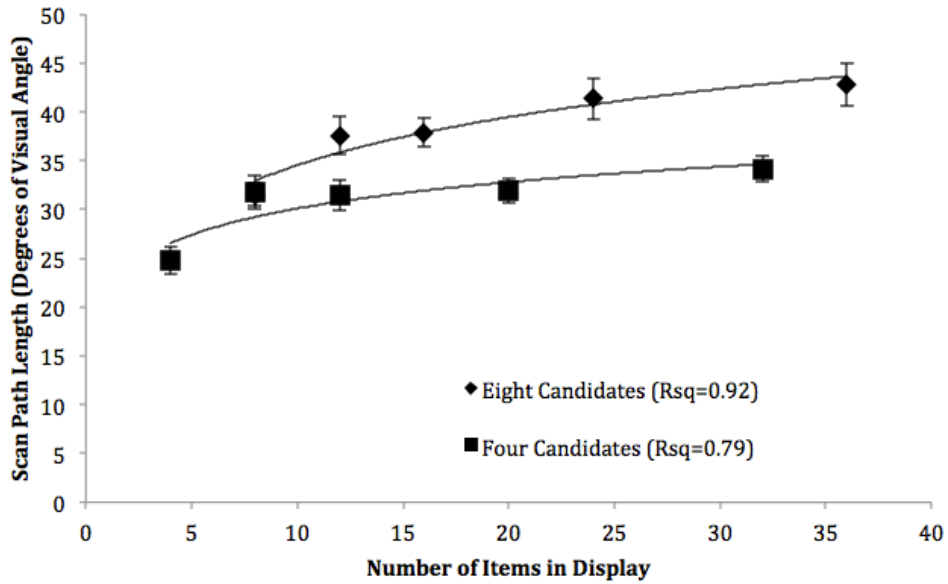
Figure 7. Scan path lengths were longer in displays containing more candidates and in displays containing more lures. The best fit lines to both candidate set-size functions was logarithmic.
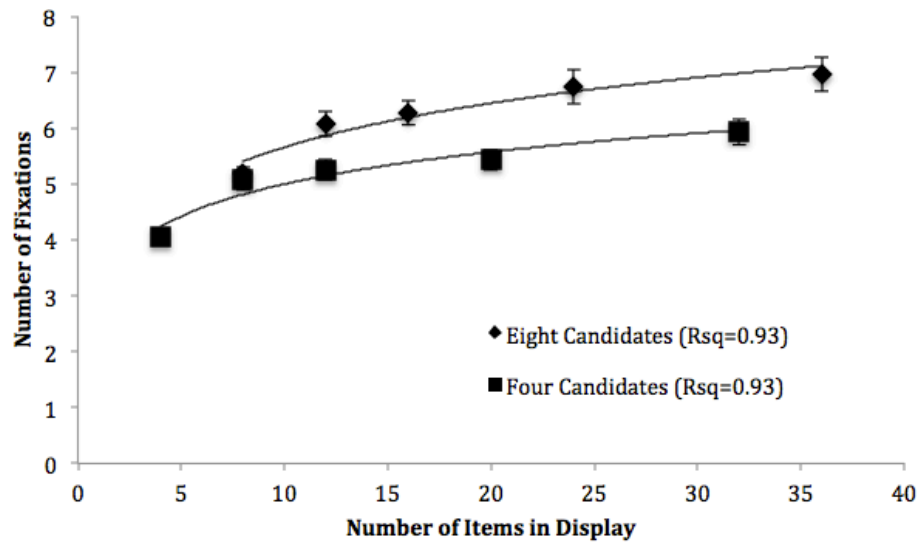

Figure 8. Participants made more fixations in displays containing more candidates and in displays containing more lures. The best fit lines to both candidate set-size functions was logarithmic.

Chapter 4: Experiment 3

Manipulating the lure-target similarity within-subjects provides a more powerful test of the effects found in our between-subjects comparison of Experiments 1 and 2. Thus, in Experiment 3 participants viewed both displays containing thick-cross lures (as in Experiment 1) and displays containing thin cross lures (as in Experiment 2). This experiment allows us to directly compare the impact of having target-similar lures and target-dissimilar lures in a within-subject design.

Method

Participants. Twenty undergraduates from the University of Illinois at Urbana-Champaign participated for course credit. All participants reported normal or corrected-to-normal vision. Three subjects were replaced due to low accuracy. Two determine the number of participants needed for this experiment an effect-size analysis was performed on several between-subjects comparisons of data from Experiments 1 and 2. An a priori power analysis for those effect sizes in within-subjects comparisons confirmed that 20 subjects were ample to detect the desired effects.

Stimuli and apparatus. As Experiment 2 was a within-subject replication of Experiments 1 and 2, stimuli were identical to those of Experiments 1 & 2 with two exceptions: (1) the 8-candidates conditions were eliminated and (2) each participant saw displays containing thick-weighted crosses and displays containing thinly-weighted crosses. The 8-candidate conditions were dropped to simplify the design, and as all of the effects found in Experiments 1 and 2 existed in the 4- and 8-candidate conditions. Crosses of different weights never appeared on the same trial. Therefore, in Experiment 2, there

were again 9 trial types each containing four letters and 0, 4, 8, 16, or 28 orange crosses that were either thick- or thin-weighted.

Design and procedure. The procedure of Experiment e did not differ from that of Experiments 1 & 2.

Data preparation. Subjects' data were checked for accuracy before further analysis. Any subject with accuracy below 90% was omitted from further analysis. This resulted in three subjects' data being replaced with data from new, naïve participants. Only correct trials were analyzed.

Results

Behavioral results. A two-way ANOVA with within-subjects factors of lure type (thin or thick crosses) and lure number (0,4,8,16,28) was performed on RTs. Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2(9)=34.33$, $p<0.001$). Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.537$). Main effects of lure type, $F(1,19)=63.47$, $p<0.001$, $\eta^2=0.045$, and lure number, $F(2.15,40.82)=26.67$, $p<0.001$, $\eta^2=0.087$. The interaction was also significant, $F(4,76)=10.353$, $p<0.001$, $\eta^2=0.029$, indicating that the thin crosses had a significantly larger effect on reaction times than thick crosses and the thin cross lures' effect on RTs was magnified at larger set-sizes.

Once again, the best-fit lines for both the thin ($R^2=0.98$) and thick ($R^2=0.95$) sets of RT data were logarithmic (Figure 9). In comparison to Experiments 1 & 2, the slope of this logarithm was slightly shallower: In Experiment 1, we found a log slope of 91.4ms/ln(unit set-size) compared to Experiment 3's 76.47 ms/ln(unit set-size) and in Experiment 2 we

found a log slope of 214.8 ms/ln(unit set-size) compared to Experiment 3's 175.0 ms/ln(unit set-size).

Eye-tracking results. Eye movement data was analyzed for the initial saccade latency, total scan path length, and the number of fixations on different item types.

Initial saccade latencies. Average initial saccade latencies are listed in Table 3. A two-way ANOVA with within subject factors of lure type (thin or thick crosses) and lure number (4, 8, 16, or 28) was performed on initial saccade latency data. There was a small main effect of lure type, $F(1,19)=5.845$, $p=0.026$, $\eta^2=0.003$, but no effect of lure number, $F(3,57)=0.428$, $p=0.734$, and no interaction, $F(3,57)=0.434$, $p=0.730$. Participants moved their eyes from fixation significantly faster in thick-lure displays than in thin-lure displays. This could be due to a number of factors. For example, thick lures may be more salient than thin lures and may have pulled participants' eyes away from fixation more quickly due to bottom-up processing (e.g., Itti & Koch, 2000). However, the effect is very small and thus may not be theoretically important.

Scan path length. A two-way ANOVA with within-subjects factors of lure type (thick or thin crosses) and lure number (4, 8, 16, or 28) was performed on the average summed saccade amplitudes (scan path lengths). The ANOVA revealed main effects of lure type, $F(1,19)=20.141$, $p<0.001$, $\eta^2=0.050$, and lure number, $F(4,76)=6.973$, $p<0.001$, $\eta^2=0.041$. The interaction was not significant, $F(3,57)=1.547$, $p=0.212$. These results indicate that the presence of thin cross lures in a display resulted in significantly longer scan paths than the presence of thick crosses and that at larger lure set-sizes, scan path lengths also increase (Figure 10).

Number of fixations. A two-way ANOVA on the total number of fixations with

within-subject factors of lure type (thin crosses or thick crosses) and lure number (4, 8, 16,

28) revealed main effects of lure type, $F(1,19)=37.13$, $p<0.001$, $\eta^2=0.0.45$, and lure number,

$F(3,57)=13.18$, $p<0.001$, $\eta^2=0.087$. The interaction was also significant, $F(3,57)=3.42$,

$p=0.023$, $\eta^2=0.006$. These results suggest that the presence of thin crosses in the display

resulted in an increase in the number of fixations participants used to inspect a display. The

more thin crosses were present, the more pronounced this effect was (Figure 11). This

effect was borderline-significant in our between-subjects comparison of Experiments 1 and

2.

Participants were significantly more likely to fixate thin-cross lures than thick-cross

lures, $F(1,19)=54.415$, $p<0.001$, $\eta^2=0.117$, and were more likely to fixate lures when more

lures were present in the display, $F(3,57)=387.275$, $p<0.001$, $\eta^2=0.867$. The interaction

between lure type and lure number also significantly affected the number of lures fixated,

$F(1.69,32.04)=10.36$, $p=0.001$, $\eta^2=0.117$(Table 3; assumption of sphericity violated:

$\chi^2(5)=27.54$, $p<0.001$; Greenhouse-Geisser correction: $\varepsilon=.562$). Participants also made

more candidate fixations when thin-crosses were present, $F(1,19)=5.376$, $p=0.032$,

$\eta^2=0.050$, and at larger lure set-sizes, $F(3,57)=7.435$, $p<0.001$, $\eta^2=0.050$. The interaction

was not significant, $F(3,57)=0.731$, $p=0.538$. Average values for number of candidates

fixated are presented in Table 3.

Discussion

The behavioral results of Experiment 3 replicate the findings of Buetti et al.

(submitted with revisions). Participants once again exhibited logarithmically increasing

reaction times with increasing numbers of lures, though the slope of this logarithm was

shallower than in previous Experiments for both lure-types, suggesting that a mixed- vs.

blocked-design may affect participants' search performance. The type of lure present in the

display was also significant—RTs were significantly longer and the slope of the logarithm

curve steeper for thin cross displays than thick-cross displays. This result is also consistent

with ITV—the more target-similar a lure is, the more time is spent screening out those

lures during Stage 1.

Once again, there was no significant effect of lure number on initial saccade

latencies, providing further evidence that the time it takes to make the first eye-movement

during search is not a good measure of the difficulty of the search task ahead and therefore

not a good measure of the time it takes to complete the Screening Stage of ITV..It is possible

that displays containing the maximum number of thick-cross lures created texture

segmentation effects—the small effect of ISL in Experiment 1 was driven by the difference

in ISL at lure set-size 28. At that set-size, the display is almost entirely filled with items, and

therefore texture segmentation may play a role in participants' faster eye-movements away

from fixation.

Scan paths were again shown to increase in length logarithmically as the number of

items in the display increased. Importantly, scan path lengths were longer in displays

containing thin cross lures compared to displays containing thick cross lures. Participants

also made more fixations in displays with more lures and in displays containing thin-cross

lures. In displays containing thin crosses, participants were more likely to fixate those thin

crosses and were more likely to make more candidate fixations than in displays containing

thick crosses. These two pieces of evidence (scan path length and fixations) provide more

evidence that the visual system is trying to resolve a larger amount of uncertainty in the

displays containing thin crosses, and that this process is driven by a logarithmic function, consistent with predictions made about ITV's screening stage.

Finally, participants made more candidate fixations in displays containing more lures and in displays containing thin-cross lures. The average number of candidate fixations once again did not fall directly in line with ITV's predictions about the number of candidates that should be fixated in 4-candidate displays (2.5). Participants made ~1 more fixation on average than this expected value (Table 3). This could be explained if Stage 1 processing has not run to completion before the first eye-movements are made, as suggested by the ISL data. Instead, Stage 1 may be occurring across eye-movements, requiring at least one extra fixation to complete screening before scrutiny begins.

Table 3

Average values of ISL, Scan Path Length, and Number of Fixations in Experiment 3

| Means	Initial Saccade Latencies				
	No Lures	4 Lures	8 Lures	16 Lures	28 Lures
Thick Crosses	195.74	198.02	194.74	202.01	185.37
(SD)	(11.52)	(14.59)	(14.49)	(13.00)	(15.91)
Thin Crosses	194.74	198.59	197.99	201.59	201.98
(SD)	(11.52)	(13.88)	(14.26)	(14.68)	(15.05)
	Scan Path Length				
Thick Crosses	21.22°	22.55°	22.92°	23.68°	24.52°
(SD)	(4.49)	(5.00)	(6.81)	(5.40)	(5.45)
Thin Crosses	21.22	23.90	26.11	26.58	28.12
(SD)	(4.49)	(5.68)	(4.98)	(5.56)	(5.35)
	Total Number of Fixations				
Thick Crosses	4.41	4.94	4.96	5.15	5.29
(SD)	(0.53)	(0.87)	(0.97)	(0.96)	(0.95)
Thin Crosses	4.41	5.24	5.54	5.74	6.12
(SD)	(0.53)	(0.90)	(1.05)	(1.20)	(1.32)
	Number of Candidate Fixations				
Thick Crosses	1.63	1.80	1.68	1.88	2.03
(SD)	(0.36)	(0.45)	(0.47)	(0.49)	(0.50)
Thin Crosses	1.63	1.92	1.96	2.06	2.25
(SD)	(0.36)	(0.49)	(0.39)	(0.67)	(0.54)
	Number of Lure Fixations				
Thick Crosses		2.87	4.42	8.43	14.48
(SD)	N/A	(0.66)	(1.11)	(1.96)	(2.86)
Thin Crosses		2.87	5.13	9.72	17.28
(SD)	N/A	(0.70)	(1.26)	(2.30)	(3.90)
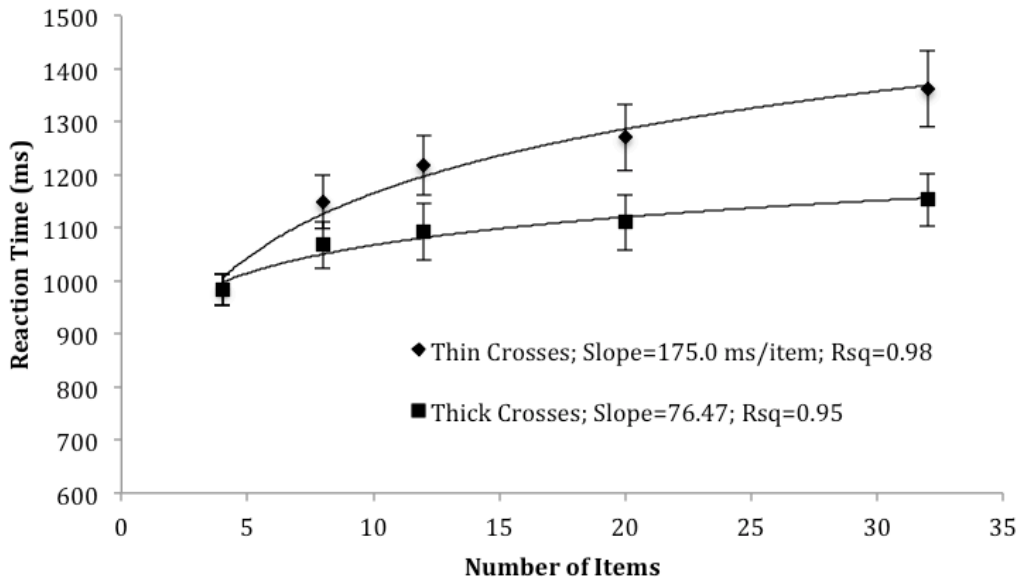
Figure 9. Reaction times for Experiment 3 indicate participants were significantly slower on trials containing thin cross lures than on trials containing thick cross lures. The best fit lines were logarithmic.
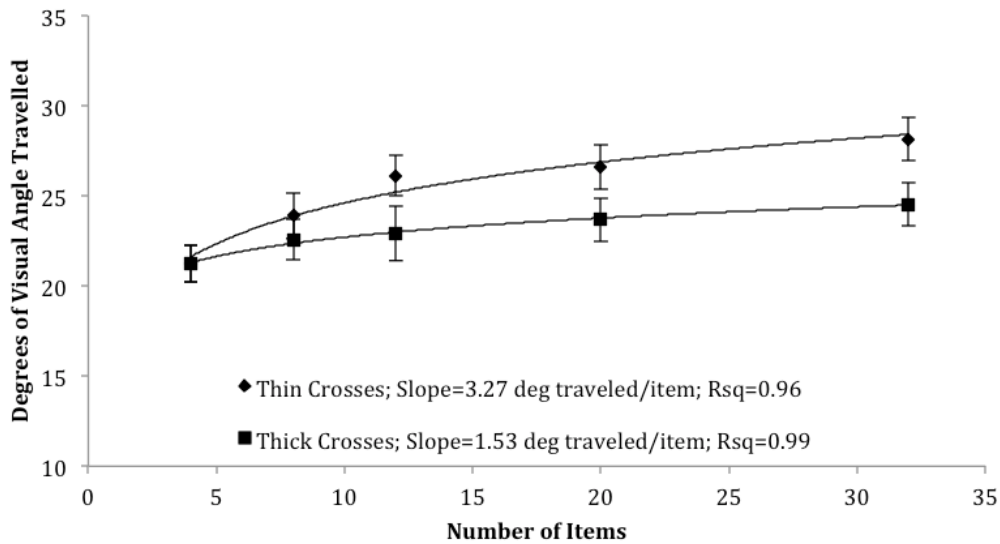

Figure 10. Scan path lengths during Experiment 3. Participants moved their eyes further in displays containing thin cross lures than in displays containing thick cross lures. The best fit lines were logarithmic.

Figure 11. Total fixations from Experiment 3 indicate that participants made significantly more fixations in displays containing thin cross lures than thick cross lures. The best fit lines were logarithmic.

Chapter 5: General Discussion

In three experiments we have presented evidence that the amount of information in a display and the characteristics of the items in a display affect reaction times and eye-movement patterns. Replicating Buetti et al. (submitted with revisions), we demonstrated that lure items do, in fact, affect reaction times in a logarithmic fashion as predicted by ITV. We further demonstrated that participants' scan path lengths and the number of fixations made were influenced by the number and characteristics of lure items present in the display. Lure items that have high target-similarity have a greater effect on RT, scan path length, and the number of fixations made in a display than target-dissimilar lures. Highly target-similar lure items are also more likely to change the focus of those fixations—the number of lure fixations in displays containing target-similar lures was greater than the number of lure fixations in displays containing target-dissimilar lures.

Implications for ITV

The results presented here offer support for several of ITV's claims. First, the RT data for all three experiments are consistent with ITV's predictions about the effects of the amount of information in a display (that is, the number of items) and the effects of lure-target similarity. ITV predicts that the Screening Stage (Stage 1) must gather more information about target-similar lures in order to screen them out compared to the amount of information gathered in displays containing target-dissimilar lures. Furthermore, this screening stage progress in a logarithmic fashion—as the number of lures increases, RTs increase as a logarithmic function of set-size. The slope of the RT logarithm is influenced by the target-similarity of the lures, with more target-similar lures yielding steeper log slopes. In the present study, the Screening Stage does not need to collect as much information to

determine that the lures are not the target when viewing the thick-cross lures in Experiment 1, as it does when viewing the thin-cross lures in Experiment 2. If we think about this in feature-space, the lures in Experiment 1 differ in color (orange) and the weight of their component lines (thick) from the thin-weighted, red target. In Experiment 2, the Screening Stage had to reject a lure that had a target-similar feature (line weight) based solely off of a color difference.[2] To be clear, ITV describes Stage 1 as evaluating overall similarity, not feature-similarity. Unlike other models of visual search (e.g., Guided Search; Wolfe, 1994), ITV does not rely on feature-maps to decompose an image before Stage 1 identifies likely targets.

The net prediction ITV makes is that RTs increase as a logarithmic function of the number of lure items in a display. The slope of that logarithm will be steeper for displays containing more target-similar lures (Experiment 2) and shallower for displays containing less-target similar lures (Experiment 1). We found evidence for these RT effects in every experiment presented here, including a within-subjects demonstration of the effects in Experiment 3.

In addition to these behavioral predictions, ITV's basis in Signal Detection Theory (Green & Swets, 1966) and Information Theory (Shannon, 1948) allowed us to also make predictions about eye-movements. As the amount of information in a display increases, so too does the amount of uncertainty (Shannon, 1948). Due to the structure of our visual

---

[2] Buetti et al. (in revision) demonstrate in their Experiment 2 that the resolution of Stage 1 processing is not sufficient to screen out L-shaped candidate items when searching for a T-shaped target. In that experiment, increasing the number of candidate items increased RTs in a linear fashion. Because the thin crosses in this experiment affected RTs in a logarithmic fashion, we can assume they are being eliminated from consideration by Stage 1. Together, this suggests that the resolution of Stage 1 may not be sufficient to distinguish the difference between the overall shapes of the thin-cross lure items and the target.

system, we typically cannot resolve the uncertainty of an entire image in one glance (e.g., Daniel & Whitteridge, 1961; Cowey & Rolls, 1974: Geisler and Chou, 1995). Therefore, ITV, like many of its predecessors, predicts that more eye-movements and fixations must be made in displays containing more items. This was experimentally confirmed in Experiments 1-3. ITV also states that increasing the lure-target similarity in a display also increases the uncertainty, again requiring more eye-movements and fixations in displays containing more target-similar lures, as was demonstrated in Experiment 3. Finally, ITV predicts that lures that are highly target similar have a higher probability of generating a false alarm—that is, a higher probability of being passed on to Stage 2 for scrutiny—than lures that are less target similar. We confirmed this prediction in Experiment 3: participants were significantly more likely to inspect (fixate) target-similar lures than target-dissimilar lures.

The Information Theory of Vision states that Stage 2 processing occurs in a random order and anticipates that the number of candidates inspected should equal $(n+1)/2$, where n is the number of candidates in the display. While we did not observe this result precisely, Experiments 2 & 3 offered near misses and, only the 4-candidate condition of Experiment 1 provided a wildly different number than was expected. In Experiments 2 & 3 and the 8-candidate condition of Experiment 1, we found slightly more candidate fixations than expected. The 4-candidate condition of Experiment 1 yielded many fewer candidate fixations than expected. Combined, this may suggest that participants are able to bias the processing of certain candidates under some conditions (low lure-target similarity, or a high ratio of candidates to lures), contrary to ITV's predictions. Future experimentation should explore this possibility. It is also important to note that the method used to

determine what a participant was 'inspecting' on any given fixation will change the precise number of items found to be inspected.

Finally, we had hoped to find an index for Stage 1 processing times in the initial saccade latency data. Our initial prediction was that initial saccade latencies would be longer in displays containing more lure or candidate items and in displays in which the lures were highly similar to the target. Instead, we found no consistent effect of candidate number or lure number. There was a very small effect of lure-type on initial saccade latencies in the predicted direction in Experiment 3, but in that experiment there was no effect of lure number. Possibilities for this null result were discussed in Experiments 1 and 3 and include crowding effects and effects of task difficulty on ISL.

Implications for other models of search

Other models of visual search make some similar predictions to ITV. In this section, we will explore how well two other models of visual search explain the data from this study. While there are many models of search, we focus here on the predictions of Guided Search (Wolfe, 1994) and Target Acquisition Model (TAM; Zelinsky, 2008). As described in the introduction, Guided Search represents an amalgamation of ideas from earlier theories such as FIT, Similarity Theory, Bundesen (1980)'s Theory of Visual Attention, and Koch and Ullman (1985)'s salience map, which makes it a good candidate for addressing many of the predictions those models of search might also make regarding this data.

While Guided Search is a purely behavioral model of visual search, TAM represents a theory grounded in explaining eye-movements during search for a target (as participants were asked to do in this experiment). Other models of search that have attempted to characterize eye-movements use free-viewing tasks or memory tasks (e.g., Peters, Iyer, Itti

& Koch, 2005; which was largely based on Itti, Koch, & Neibur, 1998 and Itti & Koch, 2000).
Because we asked participants to search for a target, TAM's predictions translate the most
readily to this set of experiments.

Guided Search 2.0 (Wolfe, 1994). As described in the introduction, Guided Search
is similar to ITV in that both models describe visual search as a two-stage process with
Stage 1 acting as a broad filter before Stage 2's careful inspection. However, there are
several differences between the two models and the predictions that they make. First,
Guided Search's Stage 1 relies on the development of a priority map based on several
feature filters. This map has peaks and troughs representing likely and unlikely target
locations based off of the boosting of _one_ target-relevant feature channel (e.g., the color
red). Ideally, the target represents the highest peak on the priority map, and that highest
peak is inspected first during Stage 2. Reaction times are driven primarily by the
inspections occurring during Stage 2, and are thought to be a linear function of set-size.
Stage 1 is thought to contribute no or minimal variation to RTs (Stage 1 processing time is
combined with response-related processing and together they are held constant at 400ms
in Guided Search 2.0). In ITV, Stage 1 gathers information from each location in a display,
rejecting non-targets as it goes based on overall target-similarity (as opposed to one target-
relevant feature channel), until it reaches a point where it cannot resolve the differences
between the remaining items. In this way, Stage 1 processing is not locked to the boosting
of one feature as in Guided Search, but to the relevance of multiple target features (e.g., red
and T-junctions, and line weight). This process occurs at a rate that is a logarithmic
function of the number of items in the display, with the slope of the logarithm dependent
on the target-similarity of the lure items. Because we find logarithmic effects in our eye-

movement data as well as in the RT data (i.e., scan-path lengths, number of fixations), it seems that Stage 1 is closely linked to eye-movements. This makes sense, given that Stage 1 and eye-movements both function to reduce the visual uncertainty of a scene. Once Stage 1 reaches the threshold at which it can no longer resolve differences between remaining items and the target, the location information of the remaining items is sent on to Stage 2 for scrutiny in a random order—unlike in Guided Search, there is no weighting of these items by their target similarity when they reach ITV's Stage 2. ITV assumes Stage 2's contribution to RTs is a linear function of the number of items inspected.

In all three experiments presented here we found logarithmic relationships between RTs and the number of items in the display, as predicted by ITV and in contrast to the linear functions predicted by Guided Search. Furthermore, we found evidence of RT modulations due to the similarity of the distractor (lure) items to the target. ITV, but not Guided Search, predicts that lure items may be processed in Stage 2 on occasion due to false alarms—this becomes more likely as the similarity of the lure to the target increases (e.g., a thin cross should be inspected more often than a thick cross, as we found). From the evidence presented here, Stage 1 is not a time-invariant process, as Guided Search assumes, but instead is affected by the target similarity of items in the display.

While Guided Search does not make explicit predictions about eye-movements, we can infer from the model that its predictions about what will be inspected might be similar to the predictions of ITV: items that have more features in common with the target produce larger peaks on the priority map and may end up being inspected. This is certainly true for targets and candidate items, and some target-similar lure items may occasionally produce a sufficiently large peak on the priority map with help from noise in the system. This is

consistent with our finding that participants were more likely to fixate the thin-cross lures than the thick-cross lures in Experiment 3.

Another major difference between Guided Search and ITV is that Guided Search asserts that Stage 1 is largely a pre-attentive process, while ITV asserts that Stage 1 is an attentive process. According to Guided Search, Stage 1 proceeds somewhat autonomously, with top-down input solely guiding the creation of the priority map by biasing the weight of specific features. ITV argues instead that attention is involved with the entire process—attention guides the rejection of non-targets based on their overall similarity to the target during Screening. Both Guided Search and ITV agree that Stage 2 is an attentive process. An interesting, though admittedly post-hoc, solution to this debate may lie in our initial saccade latency data. By definition of both Guided Search and ITV, participants did not complete Stage 1 processing before beginning to move their eyes at the start of a trial: Guided Search was simulated with Stage 1 processing times and response-related processing time fixed together at 400ms while ITV predicts Stage 1 processing times vary with lure number and lure-target similarity. In contrast to both, Experiments 1-3 yielded average initial saccade latencies of ~225ms that did not vary with lure number. So, in terms of both Guided Search and ITV, our results suggest that Stage 1 processing must be occurring across eye-movements. If we combine that knowledge with the well-established, tight linking between eye-movements and attention (e.g., Shepherd, Findlay, & Hockey, 1986; Hoffman & Subramaniam, 1995), we might presume that attention is involved in guiding Stage 1 processing, as ITV claims.

Finally, Guided Search predicts linear increases in RTs as a function of the number of items inspected by focused attention (Stage 2) and no RT effect of items that are not

inspected. ITV predicts the same linear impact of inspected items, but adds a logarithmic component of RT that is driven by the presence of lures. As described above, the RT evidence collected in this experiment clearly support the presence of a variable, logarithmic contribution to RT by lure elements, contrary to Guided Search's predictions.

Target Acquisition Model (Zelinsky, 2008). In contrast to Guided Search, which largely focuses on predicting reaction times, Zelinsky's (2008) Target Acquisition Model (TAM) focuses on predicting eye-movements, which it does remarkably well (within 95% confidence interval of participant data for most of the tasks reported in Zelinsky, 2008). TAM, while structurally different than ITV, makes similar predictions about what items will be fixated and how many saccades will be made (that is, how long scan paths will be).

Briefly, to model eye-movements during search, TAM is given two images, the image it must search through and an image of the target it is searching for. The search image is passed through a simulated retina, distorting it in much the same way that our retina would, with high fidelity representation at the foveated location and lower fidelity representation in the periphery. Then, the distorted image is passed through a set of filters to gather information about color and luminance and a target map is created. The model selects a bright spot on the target map for its next fixation and then the process repeats itself until the target is found.

In the present study, TAM would predict that the number of fixations should increase with the increase of lure-target similarity, as these target-similar lures would be more likely to show up on the target map and would be fixated. This is consistent with the increase in total fixations and lure fixations we saw in Experiment 3's thin-cross lure displays. For the same reason, TAM would also predict an increase in scan path length with

increasing lure-target similarity. This is also consistent with our results. However, TAM does not explicitly predict logarithmic increases in the number of fixations required or in the scan path lengths for displays containing more items, as was found in this study.

TAM also comes up short is in its inability to predict RTs. In fact, Zelinsky (2008) explicitly states that TAM is not intended to model RTs. TAM, therefore, while a brilliant predictor of eye-movements, does not explain all of the data presented here.

Conclusion

These results, in conjunction with the behavioral findings of Buetti et al. (submitted with revisions), provide unique evidence in support of ITV. While both Guided Search and TAM make some of the same predictions as ITV, no model save for ITV accounts for all aspects of this data.

It is exceedingly rare for a human to perform a visual search without making any eye-movements. To be ecologically valid, then, models of visual search should be able to account for both the eye-movements made during search and the amount of time it takes to perform a search. In this instance, ITV appears to be more successful than previous models on both counts.

REFERENCES

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433-436.

Buetti, S., Cronin, D.A., Madison, A.M., Wang, Z., & Lleras, A. (submitted with revisions). Evidence for parallel, unlimited capacity, exhaustive attentive processing in human vision.

Buetti, S., & Lleras, A. (in preparation). An information theory model of visual search.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4), 523-547.

Chang, H., & Rosenholtz, R. (2014). New exploration of classic search tasks. *Journal of Vision, 14*(10), 933.

Cowey, A. & Rolls, E.T. (1974). Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research, 21*, 447-454.

Daniel, P.M. & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology, 159*(2), 203-221.

Duncan, J. & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433-458.

Geisler, W.S. & Chou, K.-L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review, 102*, 356-378.

Green, D.M., & Swets, J.A. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Hick, W.E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*(1), 11-26.

Hoffman, J.E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*(6), 787-795.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid

      scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence,*

      20(11), 1254-1259.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of

      visual attention. *Vision Research, 40*, 1489-1506.

Koch, C. & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying

      neural circuitry. *Human Neurobiology, 4*, 219-227.

Neider, M. B. & Zielinsky, G.J. (2008). Exploring set size effects in scenes: Identifying the

      objects of search. *Visual Cognition, 16*(1), 1-10. doi: 10.1080/13506280701381691

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers

      into movies. *Spatial Vision, 10,* 437-442.

Pelli, D.G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in*

      Neurobiology, 18, 445-451. doi: 10.1016/j.conb.2008.09.008. PMC3624758

Pelli, D.G., Palomares, M., & Majaj, N.J. (2004). Crowding is unlike ordinary masking:

      Distinguishing feature integration from detection. *Journal of Vision, 4*(12), 1136-1169.

Peters, R.J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in

      natural images. *Vision Research, 45,* 2397-2416.

Pomplun, M., Garaas, T.W., & Carrasco, M. (2013). The effects of task difficulty on visual

      search strategy in virtual 3D displays. *Journal of Vision, 13*(3):24, 1-22.

Santhi, N., & Reeves, A. (2004). The roles of distractor noise and target-certainty in search: A

      signal-detection model. *Vision Research, 44*(12), 1235-1256.

Shannon, C.E. (1948). A mathematical theory of communication. *The Bell System Technical*

      Journal, 27, 379-423, 623-656.

Shepherd, M., Findlay, J.M., & Hockey, R.J. (1986). The relationship between eye-movements and spatial attention. *Quarterly Journal of Experimental Psychology, 38A,* 475-491.

Treisman, A.M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97-136.

VanRullen, R., & Thorpe, S.J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorization of natural and artifactual objects. *Perception, 30,* 655-668.

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

Wolfe, J.M., Alverez, G.A., Rosenholtz, R., Kuzmova, Y.I., & Sherman, A.M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception, and Psychophysics, 73,* 1650-1671. doi: 10.3758/s13414-011-0153-3

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115(4)*, 787-835.