Taking stock: The role of environmental appraisal in the strategic use of attentional control

Heather A. Hansen, Jessica L. Irons, and Andrew B. Leber

The Ohio State University

Please address correspondence to:

Heather A. Hansen, Department of Psychology, The Ohio State University, 225 Psychology

Building, 1835 Neil Avenue, Columbus, OH, 43210. E-mail: hansen.508@osu.edu

and

Andrew B. Leber, Department of Psychology, The Ohio State University, 225 Psychology

Building, 1835 Neil Avenue, Columbus, OH, 43210. E-mail: leber.30@osu.edu

Word Count: 8800, including main body and references

Abstract

There are many strategies we can use to control attention when approaching a visual search task, but some are more effective than others. How do we choose the most optimal strategy? We have proposed that individuals must appraise the stimulus environment, taking in relevant statistical information about task-relevant features. In the present experiment, we examined whether interfering with the appraisal process via a secondary task decreases participants' use of the optimal strategy. We used a modified version of the Adaptive Choice Visual Search paradigm (Irons & Leber, 2016; 2018), whereby individuals can freely search for either of two targets on every trial. Each search display was preceded by a colored environmental preview, offering participants time to appraise the display and determine which target would be more optimal to search for. On some blocks, participants also completed a secondary task – a central line-length judgment – either before or during this colored preview. We found that participants were significantly less likely to search optimally when the line task occurred during the colored preview than when it occurred beforehand or was absent. Insofar as the secondary task disrupts an individual's ability to engage in appraisal, these results support the need for such an appraisal mechanism in the optimal choice of attentional control settings.

Introduction

When searching for your red SUV in a parking lot, the most efficient strategy to employ

will depend on the other cars parked around you. For instance, if you parked amongst black or

white cars, you should bias your search by color; if you parked amongst small cars, you should

bias your search by size. Thus, the most optimal search strategy to use often changes each new

time you search and will depend on the particular features of your current environment.

Individuals are capable of allocating attention toward a particular feature through goal-directed

attentional control, and foreknowledge of the feature improves efficiency by selectively biasing

the processing of that feature (e.g., Green & Anderson, 1956; Folk, Remington, & Johnston,

1992). However, while the optimal strategy yields reliably better performance (i.e., in speed

and/or accuracy), people frequently fail to use it (Bacon & Egeth, 1994).

Why do people choose suboptimal search strategies? We have previously posited that

such choices may be due to a failure or unwillingness to update cognitive control settings as the

environment changes, because doing so is effortful (Irons & Leber, 2016; 2018). Optimizing

performance requires sustained proactive control and maintenance of task goals (Braver 2012;

Braver, Gray, & Burgess, 2007), which is cognitively demanding (Braver et al., 2007; Chatham,

Frank & Munakata, 2009; Lock & Braver, 2008). Additionally, optimality necessitates

performance monitoring, which requires individuals to metacognitively judge the effectiveness

of their current strategy relative to their goals (Cain, Vul, Clark, & Mitroff, 2012; O'Leary &

Sloutsky, 2016; Wolfe 2013), and update when there is a mismatch. This employment of

conflict-monitoring mechanisms, whereby conflicts in information processing are monitored

and adjustments in cognitive control are made accordingly, is inherently resource demanding

(Botvinick, Braver, Barch, Carter, & Cohen, 2001; Lorist, Boksem, & Ridderinkhof, 2005).

Further, switching between cognitive strategies or tasks takes effort and is often avoided (Kool,

McGuire, Rosen, & Botvinick, 2010; Arrington & Logan, 2004); it has been suggested that the

cognitive control system only adjusts if the expected subjective performance gain outweighs

the effort required (Botvinick & Braver, 2015; Botvinick & Rosen, 2009; Walton, Bannerman,

Alterescu, & Rushworth, 2003).

    In addition to proactive control and performance monitoring, in the case of visual

search, another key factor may underlie optimal choices: *environmental appraisal.* When you

are searching for your car in a parking lot, selecting the most efficient strategy might require

you to first scan your present environment and take in task-relevant information – whether

there are more black cars or small cars, for example. Often times real-world search tasks take

place in environments that change between searches, such as screening bags with different

shapes at the airport or locating your child amongst different sets of play equipment.

Effectively, optimal choices may rely on first taking a step to appraise the features of the

current search environment, so cognitive control settings can be adjusted accordingly. For the

purposes of this paper, we are using appraisal to mean the rapid processing of the stimuli in an

array, during which statistical summary representations, namely representations of statistical

information, such as mean set or stimulus size, are formed.

    Prior work suggests that extracting statistical information about the display is rapidly

and accurately accomplished with distributed attention to the entire environment (Ariely, 2001;

Chong & Triesman, 2003; Chong & Triesman, 2005a; Chong & Triesman, 2005b). Further,

multiple colored sets can also be enumerated and compared against each other swiftly and

efficiently (Chong & Triesman, 2005b), for up to three sets of colored items in parallel

(Halberda, Sires, & Feigenson, 2006). Thus, appraising an environment with a few sets of

colored stimuli should be a relatively effortless process. However, these tasks require

participants to specifically make a simple enumeration or comparison judgment – how does the

ability to enumerate integrate into more complex task environments, such as visual search?

How might individuals take the result of this rapid statistical computation and use it to inform

their search strategy?

        In this paper, we investigate how appraisal of the stimulus environment contributes to

visual search optimality. To this end, we used a modified version of the Adaptive Choice Visual

Search (ACVS; Irons & Leber, 2016; Irons & Leber, 2018), a paradigm we developed to explore

how individuals choose attentional control settings in a dynamic and relatively unconstrained

task environment. In the ACVS, participants search a colored array for either a blue or red

square containing a target digit; that is, two targets are present, but participants only have to

respond to one of them. Critically, one target color is typically faster to find (i.e., more optimal)

than the other because there are fewer squares of that color to search through. Additionally,

the more optimal color is switched unpredictably over the course of trials, requiring

participants who seek to optimize their performance to occasionally update their control

settings. This paradigm emphasizes target choice (i.e., optimal vs. non-optimal) as a key

dependent measure, and the method allows us to characterize group-wide tendencies as well

as individual variation in the choice of attentional control strategies, a trait that has been

shown to remain stable across experimental sessions conducted on separate days (Irons &

Leber, 2018). Results have shown that, overall, individuals are far below optimal in which target

they choose to search for; the sub-optimal target was chosen in about 40% of the trials in the original experiment (Irons & Leber, 2016).

Our present goal is to use the ACVS paradigm to understand the role of appraisal in the choice of attentional control strategies. To most effectively study this question, we sought to modify the paradigm to provide participants with a maximal opportunity to appraise the stimulus environment. Our chosen approach was motivated by recent experiments revealing that individuals make about 10% more optimal choices when they are a) presented a colored preview of the display beforehand (i.e., a display that reveals the ratio of colored subset sizes but not the target and distractor digits), and b) explicitly informed which target is more optimal to find (Irons, Hansen, & Leber, in preparation; Irons & Leber, 2018, JOV abstract). Given these findings, we used a combination of both preview and instruction in the present experiments as tools to maximally facilitate appraisal.

When given instructions of the optimal strategy, a colored preview increases optimal performance relative to no preview. However, it is unclear exactly why this preview period is beneficial. One potential explanation is that offering a colored preview of the display beforehand allows individuals time to appraise the color ratio and determine which subset they ultimately want to search through before digits appear and the search begins. If this appraisal period is crucial for optimality, then if we disrupt participants' ability to appraise the search environment by introducing a secondary task during this time, will it lead to suboptimal choices? To test this, we added an irrelevant line task before the search array, whereby participants needed to judge which line of a cross (horizontal or vertical) was longer. Critically, this line task could occur either *before* the colored preview or *during* the colored preview. We

hypothesized that this irrelevant task could have two alternative consequences: First, if the

appraisal period contributes to optimal performance, then the additional task might only be

detrimental if it occurs simultaneously with the preview display. Alternatively, if the appraisal

period itself is not critical or if the benefits of the preview period are attributable to some other

type of preparatory process that participants engage in, then the additional task might either

have no impact at all or produce a general interference cost in all conditions equally. This paper

presents two experiments aimed to test the effect of this irrelevant secondary task on the

appraisal period, and ultimately how the ability to appraise one's environment impacts optimal

performance.

Experiment 1

Method

     Participants. Twenty-five individuals (11 male, 14 female) aged 18 to 35 (M = 20.21)

completed Experiment 1. One participant in Experiment 1 was removed for low accuracy (more

than three standard deviations below the group mean). Since a prior version of this

manipulation had not been previously run, we could not use a power analysis to estimate

necessary sample size; as such, we elected to establish the effect size and analyze results with a

sample of twenty-four participants. All participants were recruited from the undergraduate

psychology research pool at The Ohio State University, and all self-reported normal or

corrected-to-normal visual acuity and normal color vision. Individuals came in for one one-hour

long session, and received course credit for their participation. All experimental methods were

approved by The Ohio State University Institutional Review Board, and all participants provided

informed written consent.

Stimuli. *Adaptive Choice Visual Search.* Each search array was composed of 54 colored squares (sized 1° x 1°) arranged in three concentric rings centering around a fixation dot (Figure 1). The inner ring contained 12 evenly-spaced squares (each centered at 6.3° eccentricity), the middle ring contained 18 squares (9.4° eccentricity), and the outer ring contained 24 squares (12.4° eccentricity). Of the 54 squares, 13 were colored red, 13 were colored blue, 14 were colored green, and 14 were a variable set. The variable set changed between all red and all blue, alternating in runs of one to six trials (see Figure 1). The spatial arrangement of the colored squares within each display was randomized on each trial.

A white digit between 2 and 9 (0.48°, font: Arial) was superimposed on the center of each square; digits were small enough that gaze had to be fixated in their vicinity to be identified. Every search array contained two targets: one red square and one blue square, each with a digit between 2 and 5 (i.e., 2, 3, 4, or 5). All other red and blue squares contained digits between 6 and 9. Green squares could contain any digit between 2 and 9, so as to prevent participants from ignoring color and biasing their search using digit value only. All non-target digits were assigned pseudorandomly across the squares, such that within each color subset, each digit appeared as close to equally often as possible. Red and blue target digits were chosen pseudorandomly, such that, across all trials, each digit appeared equally often in each color; also, the two targets always contained different digits on each trial, in order to discern which target was reported.
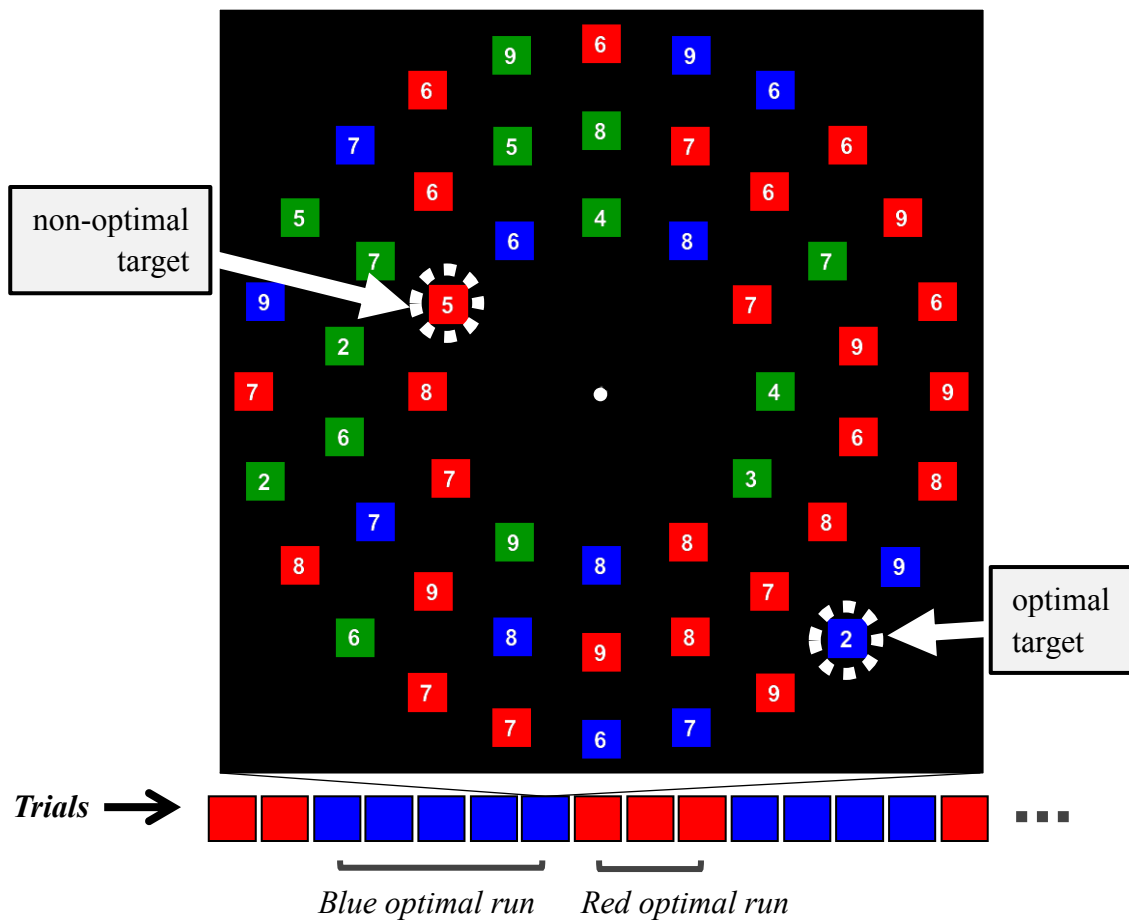
Figure 1. An example search array, with targets highlighted. In this example, the blue target containing a two is considered more optimal to find than the red target containing a five because there are half as many blue squares as red to search through. Below, a sample trial order is shown, in which the optimal color varies in runs of one to six.

Line task. The line task consisted of a white cross presented at the center of the

screen, comprised of a longer line segment (0.57°) bisecting a shorter line segment (0.29°). The

presentation order of the crosses was randomized in each block such that 50% of the trials

contained a longer vertical line segment and 50% contained a longer horizontal line segment,

with equal numbers of each orientation at each run position in the search task.

Design & Procedure. The experiment was run in a dimly lit, sound attenuated testing

room using a Mac Mini computer and 24in LCD monitor. Participants were seated at a viewing

distance of approximately 60cm (head position was not fixed). Stimuli were presented using

Matlab (Mathworks, Natick, MA) with Psychophysics Toolbox extensions (Brainard, 1997; Pelli,

1997; Kleiner et al., 2007).

The experiment consisted of a blocked design with three conditions: 1) search-only

trials, 2) line-before trials, and 3) line-during trials.

Search-only trials (Figure 2A): Each trial began with a fixation dot for 2s, followed by a

preview of the colored squares (but excluding digits) for 1s. After this brief preview, the digits in

each square appeared and participants had unlimited time to search for a target. Participants

indicated their target selection on each trial by responding with the digit inside the target

square they located, using the V, B, N, and M keys on the keyboard (corresponding to 2, 3, 4,

and 5, respectively). Once a keypress was made, the array was removed, and a new trial began

after a 2s inter-trial interval (ITI). If participants responded incorrectly (i.e., they made a

response that did not match either of the two targets), they heard a 400Hz auditory tone for

150ms during the ITI.

Line-before trials (Figure 2B): Each trial began with a fixation dot for 1s, then the line

task for 1s. The line task required participants to judge whether the horizontal or vertical line

segment was longer in a cross presented at fixation, using the 1 and 2 keys on an attached

number keypad (corresponding to 'horizontal' and 'vertical', respectively) to indicate their

response. The cross appeared on the screen for 1s, and participants had to make their response

within this time frame; if they responded incorrectly or timed out, they heard a 400Hz auditory

tone for 150ms either immediately following their incorrect response, or at the onset of the

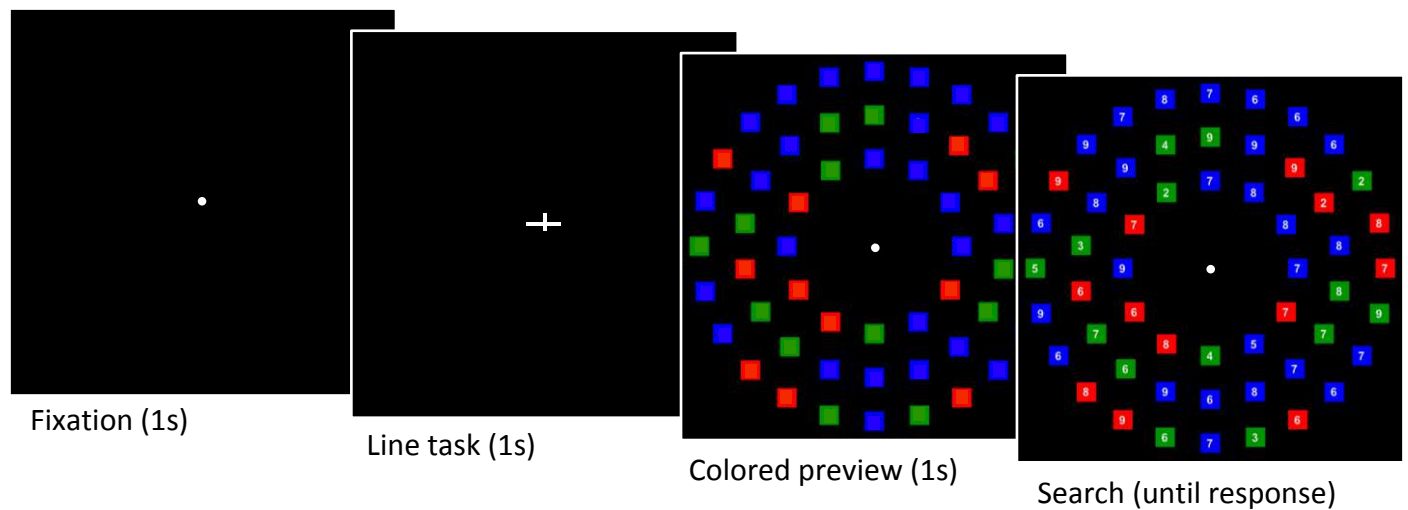search array, respectively. The cross of the line task was replaced with the fixation dot again
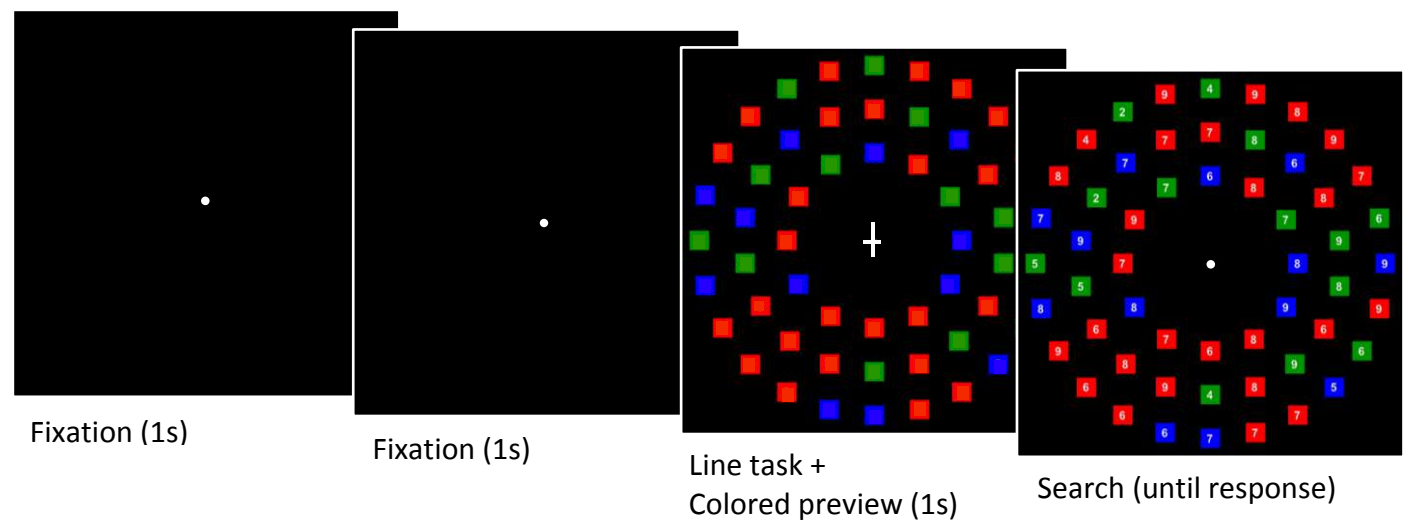
immediately following a participant's response, or at the end of the 1s if participants did not

respond. Following the line task, participants were shown the preview of the colored squares

(excluding digits) for 1s, and finally the full search array until a response was made and a new

trial started. Participants advanced to and completed the search task, regardless of line task

performance.

Line-during trials (Figure 2C): Each trial began with a fixation dot for 2s, as in the search-

only blocks. Then, the line task appeared concurrently with the preview of the colored squares

for 1s, followed by the full search array.

Participants began the experiment by practicing the search task. They were told they

would see search arrays of colored squares containing two potential targets, but would only

need to locate one of the targets on each trial and were free to choose which target they

searched for each time. They were then given 10 practice trials of the search task. Following

these practice trials, participants were informed of the optimal strategy to complete the search

task: "The fastest way to do the task is to look for whichever color has the *fewest* squares. For

example, if there are fewer red than blue squares, it will usually be faster to look for the red

target." After, participants were instructed how to do the line task. Participants were given five

practice trials of each of the two line conditions before starting on the main experiment.

Participants were instructed to place their dominant hand on the four search task keys and

their other hand on the two line task keys for the duration of the experiment; the placement of

the keypad was arranged to facilitate this.

The experiment consisted of six blocks of 84 trials (504 total trials), with self-paced

breaks in between each block. Participants completed two blocks each of the search-only, line-

before, and line-during conditions. Block order of the three conditions was completely

counterbalanced across participants, and then the order repeated to create six blocks (e.g.,

ABCABC, ACBACB, etc.) . After completion of the experiment, participants filled out a short self-

report questionnaire that probed the percentage of time they utilized different search

strategies as well as how they utilized the 1s colored preview.

A) Search-only condition: no line task

Fixation (1s)   Fixation (1s)   Colored preview (1s)   Search (until response)
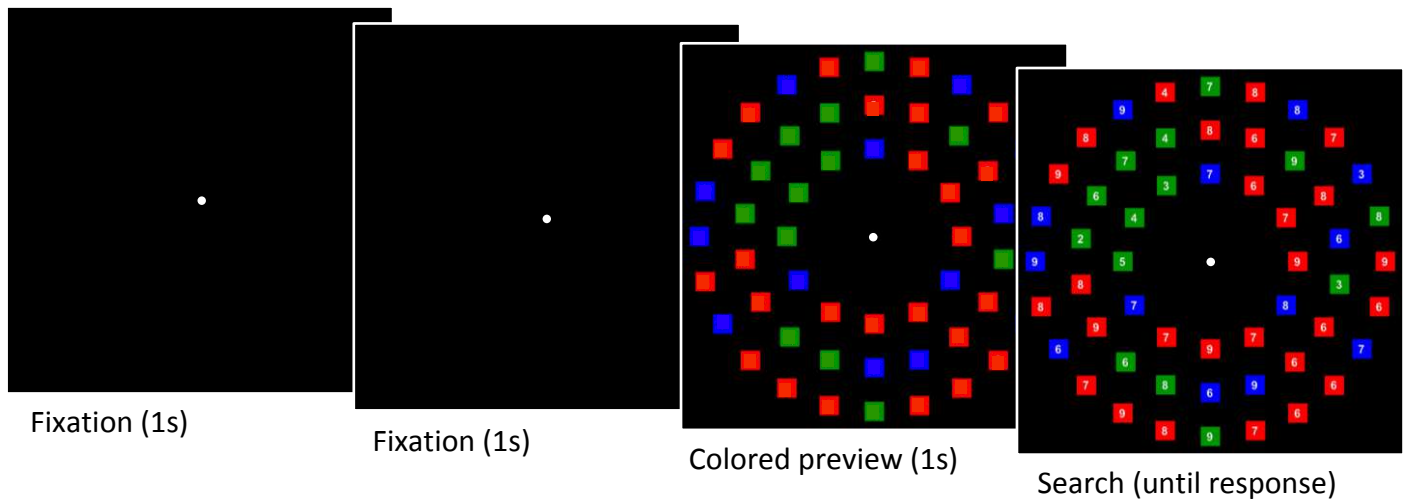
B) Line-before condition: Line task appears *before* colored preview

Fixation (1s)   Line task (1s)   Colored preview (1s)   Search (until response)

C) Line-during condition: Line task appears *during* colored preview

Fixation (1s)   Fixation (1s)   Line task + Colored preview (1s)   Search (until response)

Figure 2. Schematic of the three task conditions.

Results and Discussion

Search accuracy was close to ceiling for all three conditions (search-only M = 96.53%, line-before M = 96.03%, line-during M = 95.86%), and was not significantly different between the conditions, $F(2,46)$ = 0.325, p = 0.724, η_p^2 = 0.014. The following analyses exclude trials in which participants were unable to correctly find or report a search target, as well as trials with search RTs less than 300ms or more than three standard deviations above the mean (0.99% of search-only trials, 1.58% of line-before trials, 2.05% of line-during trials). For analyses in which multiple comparisons were conducted, we used the Holm-Bonferroni method (Holm, 1979) to control the familywise Type I error rate (corrected p-values are denoted by p_{HB}).

Search: Optimality. First, we sought to explore the main strategy measure reported by Irons and Leber (2016, 2018): Optimal choice. Optimal choice was defined as the percentage of correct search trials in which the target from the smaller subset (i.e., the more optimal target) was chosen. Note that while we refer to choices as optimal or not, we accept that target choice on each individual trial is not a perfect reflection of strategy usage; even participants searching optimally may sometimes opportunistically choose a non-optimal color (e.g., if it is located near fixation; see Irons & Leber, 2016; 2018). Nevertheless, we have repeatedly found that choosing the optimal color is overall the most efficient search strategy, as it is reliably associated with faster RTs (including in this paper; see RT results below).

Other data from our lab have shown that, with both a colored preview and instruction of the optimal strategy, participants made approximately 80% optimal choices (Irons, Hansen & Leber, in prep). Similarly in this experiment, percentage of optimal choices was overall relatively high and negatively skewed, with individual participants ranging from 50-100% optimal on any

given block. Thus, even when explicitly told the strategy to search most optimally and given

appraisal time to execute the strategy, some participants still choose to search suboptimally.

Optimal performance varied significantly between the conditions, $F(2,46) = 5.236$, $p =$

0.009, $\eta_p^2 = 0.185$ (Figure 3). There was no difference in optimality between the search-only

condition ($M = 89.16\%$, $SD = 10.87\%$) and the line-before condition ($M = 89.27\%$, $SD = 10.46\%$),

$t(23) = -0.072$, $p_{HB} = 0.943$, $d = 0.010$. However, participants were significantly less optimal in

the line-during condition ($M = 85.44\%$, $SD = 10.80\%$) compared to both the search-only

condition ($t(23) = 2.733$, $p_{HB} = 0.024$, $d = 0.343$) and the line-before condition ($t(23) = -3.174$, p_{HB}

$= 0.012$, $d = 0.360$). This result shows that there was a tendency for participants to search less

optimally when an irrelevant secondary task disrupted the appraisal preview period but not

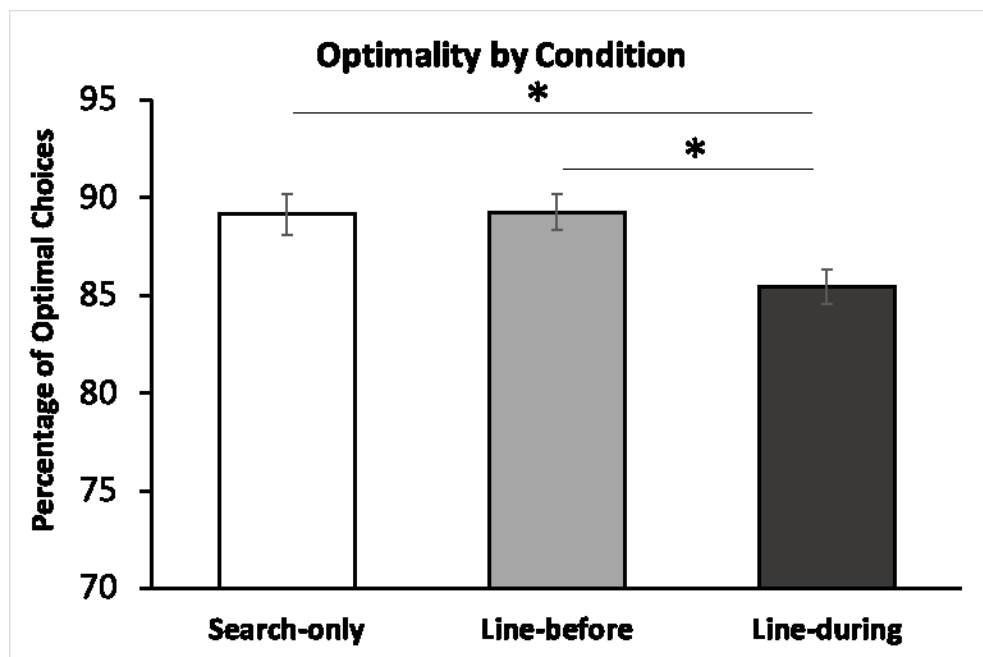when it occurred before appraisal could be carried out.

Figure 3. Percentage of optimal choices made in each of the three conditions, on average. Error bars depict within-subjects standard error calculated using the Cousineau method (Cousineau, 2005) with a Morey correction (Morey, 2008). * $p < 0.05$.

Search: Response Time. Although the optimal search strategy is effortful in that it requires continuous monitoring and target switching, it is associated with quicker performance; as in previous versions of the task, individuals who made a higher percentage of optimal choices also achieved faster RTs (r = -0.341, p = 0.003). Like percentage of optimal choices, search response time (RT) varied significantly between the conditions, $F(2,46)$ = 6.751, p = 0.003, η_p^2 = 0.227 (Figure 4). Again, there was no difference in search RTs during the search-only condition (M = 2413ms, SD = 441ms) and the line-before condition (M = 2446ms, SD = 518ms), $t(23)$ = -0.532, p_{HB} = 0.600, d = 0.069. However, participants were significantly slower in the line-during condition (M = 2634ms, SD = 547ms) compared to both the search-only condition ($t(23)$ = -3.410, p_{HB} = 0.006, d = 0.445) and the line-before condition ($t(23)$ = 2.803, p_{HB} = 0.020, d = 0.353). Thus, not only are participants making fewer optimal choices when the line task is presented during the colored preview, but their performance is slower. It is possible this speed deficit may be attributable in part to a switch cost, since the line task in the line-during condition immediately precedes the start of the search trial. However, participants are also switching between tasks in the line-before condition and performance is no different there than when the search is performed alone. Further, the slowing is not explainable by participants using search time to appraise the display and merely delaying the commencement of search; if so, we would expect slower RTs but equal optimality in the line-during condition, but optimal performance was significantly lower. Therefore, the slowing might be more readily attributable to the disruption of the appraisal period, whereby participants actually adopted less efficient strategies on line-during trials.
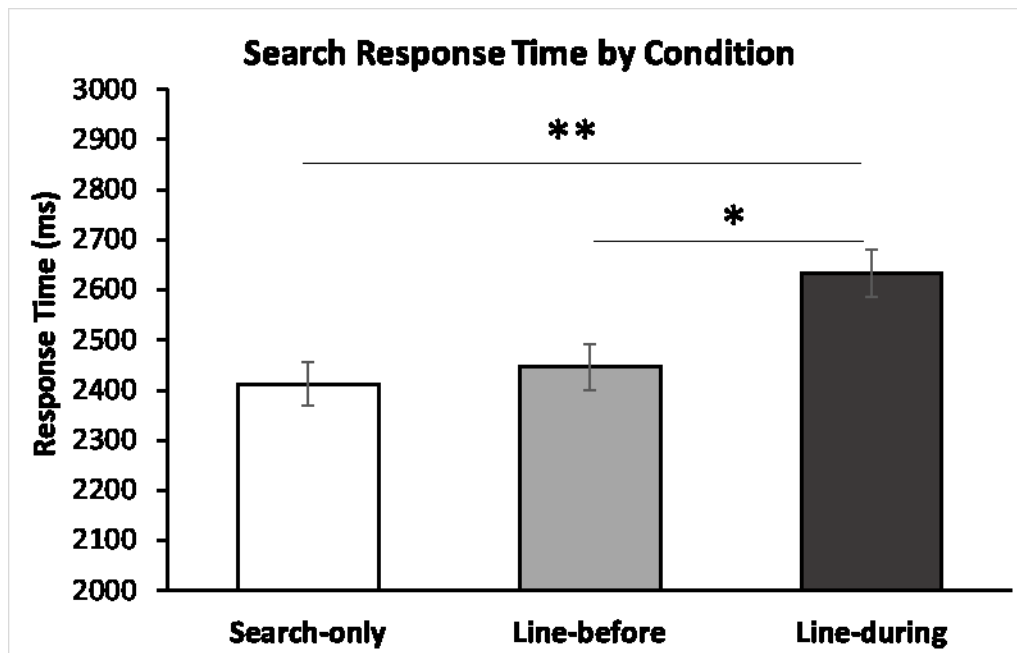
Figure 4. Average response times to find a target in the search task, in milliseconds, in each of the three conditions. Error bars depict within-subjects standard error calculated using the Cousineau method with a Morey correction. ** $p < 0.01$, * $p < 0.05$.

Line Task. Accuracy at the line task was marginally better when the line task was presented before the colored preview display ($M = 93.65\%$, $SD = 4.98\%$) than when it was presented during the colored display ($M = 92.14\%$, $SD = 5.62\%$), $t(23) = -1.843$, $p = 0.078$, $d = 0.284$. Further, participants were significantly faster at completing the line task in the line-before condition ($M = 546$ms, $SD = 42$ms) than in the line-during condition ($M = 578$ms, $SD = 51$ms), $t(23) = 5.742$, $p < 0.001$, $d = 0.685$. This might be due in part to there being more visual input on the screen in the line-during condition, such that the onset of the color preview captured attention and impaired processing of the cross, compared to a cross appearing on a blank background; relatedly, the cross might have increased perceptual load (Lavie, 2005) or diluted processing of the color information (Tsal & Benoni, 2010).

Relationship between Line Task and Search Task. As optimality is our main measure of interest, we sought to explore the relationship between line task performance and subsequent search performance. First, we looked at whether line task performance was associated with optimal choices across individuals. Individuals who find the line task more difficult and respond more slowly may have less time to appraise the preview, and therefore make fewer optimal choices overall. However, average line task speed was not correlated with percent optimal search choices in either the line-before condition ($r < 0.001$, $p = 0.999$) or the line-during condition ($r = 0.017$, $p = 0.963$), which supports the interpretation that the degree to which an individual found the line task difficult had no bearing on their likelihood to make optimal choices. However, assessing variation *within* individuals by dividing each individual's line task RTs into fast trials and slow trials, using a median split, *does* shed light on resulting optimal choices: whether the line task was accomplished quickly or slowly did not predict subsequent search optimality in the line-before condition (fast $M = 89.38\%$, slow $M = 89.53\%$, $t(23) = 0.105$, $p = 0.917$, $d = 0.014$), but it *did* significantly predict optimality in the line-during condition (fast $M = 87.69\%$, slow $M = 84.41\%$, $t(23) = -2.302$, $p = 0.015$, $d = 0.296$). The speed x condition interaction was marginally significant, $F(1,23) = 4.036$, $p = 0.056$, $\eta_p^2 = 0.149$ (Figure 5). One possible explanation for these results is that when participants made the judgment quickly in the line-during condition, they could then use the remainder of the 1s preview time to appraise the display. In contrast, when participants were slower and needed the majority of the 1s preview to make the line judgment, they had no leftover time to appraise and were thus significantly less optimal than on fast trials. Therefore, it could be that the degree to which an

interruption impairs task optimality depends on the extent to which the interruption prevents

the participant from appraising the display.
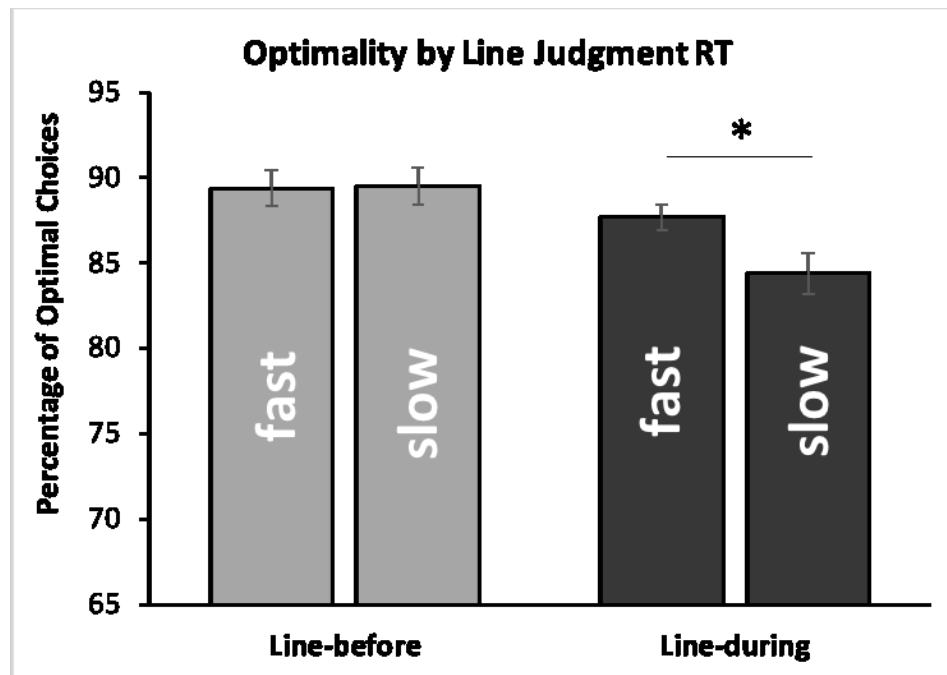
Figure 5. Average percentage of optimal choices made on the search task for trials in which the line task preceding that search array was completed quickly (i.e., fastest 50% of trials, for each individual) versus slowly (i.e., slowest 50% of trials, for each individual). Error bars depict within-subjects standard error calculated using the Cousineau method with a Morey correction. * $p < 0.05$.

Experiment 2

From Experiment 1, we might conclude that the appraisal process is necessary for

optimal choices, as evidenced by significantly lower optimality and slower response times in the

line-during condition, compared to the search-only baseline. This result contrasted with the

line-before condition, which produced no reliable differences in performance, compared to the

search-only baseline. However, Experiment 1 presented conditions in a blocked format,

wherein participants knew when to expect the line task disruption for the entirety of the 84

trials in each block. In the real world, we are often not aware of when disruptions will occur.

Critically, the line-before and line-during conditions may have unintentionally differed in more

than just the timing of the disruption with the preview period. Specifically, because we used a

blocked design, participants could have formed different expectations of effort for each line

task condition, which may have affected task preparation. If they perceived the blocks of line-

during trials as more cognitively demanding, then they may have adopted a strategy of *effort*

avoidance (Kool et al., 2010), which would lead them to abandon the optimal strategy more

often. Importantly, this effort avoidance account is distinct from our original interpretation that

the denial of time to appraise the displays is what drove the observed effects.

One approach to address the alternative effort avoidance explanation is to endow

participants with similar expectations of cognitive demand prior to each trial. In particular, we

mixed the line-before and line-during conditions within blocks, so that participants would enter

both the line-before and line-during trials with similar expectations and thus similar degrees of

effort avoidance. If effort avoidance drove the effects in Experiment 1, then we should now see

reduced optimality in both the line-before and line-during conditions, compared to the search-

only baseline. However, if the Experiment 1 results were driven by the denial of access to

display appraisal in the line-during condition – but not the line-before condition – then we

should essentially reproduce the Experiment 1 results: reduction of optimality in the line-during

condition but not in the line-before condition.

Method

Participants. Twenty-five individuals (13 male, 12 female) aged 18 to 35 ($M = 22.33$)

completed Experiment 2. One participant in Experiment 2 was removed for low accuracy (more

than three standard deviations below the group mean), yielding an analyzed sample of twenty-four. Given the effect on optimality by condition found in Experiment 1 (η_p^2 = 0.185), a sample of this size should yield a significant result with power greater than 0.999. Participants were recruited from both the undergraduate psychology research pool and general population at The Ohio State University. All individuals self-reported normal or corrected-to-normal visual acuity and normal color vision. Individuals came in for one one-hour long session, and received either course credit or $10 compensation for their participation. All experimental methods were approved by The Ohio State University Institutional Review Board, and all participants provided informed written consent.

Procedure. Experiment 2 used the same stimuli and procedure as Experiment 1, except with a mixed design. As in Experiment 1, we presented the search-only condition in pure blocks. However, now the line task conditions were mixed: Participants completed four blocks with both line-before and line-during trials. Line blocks were configured such that there were equal amounts of each line task condition within a block, as well as within each position in a search run before the optimal color switched. The line task conditions were randomized, with the constraint that no more than five trials of one condition could appear in a row. As with Experiment 1, block order was counterbalanced across participants.

Results and Discussion

As in Experiment 1, search accuracy was at ceiling for all three conditions (search-only *M* = 97.64%, line-before *M* = 97.52%, line-during *M* = 97.74%), and was not significantly different between the conditions, $F(2,46)$ = 0.223, p = 0.801, η_p^2 = 0.010. Analyses for Experiment 2 likewise excluded inaccurate trials and trials with search RTs less than 300ms or more than

three standard deviations above the mean (0.93% of search-only trials, 1.32% of line-before

trials, 1.54% of line-during trials).

Search: Optimality. As before, we first explored the percentage of optimal choices that

participants made. Participants ranged from 41% to 99% optimal on any given block, and

overall optimal performance varied significantly between the conditions, $F(2,46) = 8.701$, $p =$

0.001, $\eta_p^2 = 0.274$ (Figure 6). The difference in optimality between the search-only condition (M

$= 80.49\%$, $SD = 16.79\%$) and the line-before condition ($M = 78.21\%$, $SD = 17.74\%$) was marginal

but not significant ($t(23) = 1.884$, $p_{HB} = 0.072$, $d = 0.132$), which may reflect a marginal

contribution of effort avoidance when completing an additional secondary task. However,

participants were significantly less optimal with their choices in the line-during condition ($M =$

75.85%, $SD = 16.71\%$) compared to both the search-only condition ($t(23) = 4.040$, $p_{HB} = 0.003$, d

$= 0.277$) and the line-before condition ($t(23) = -2.447$, $p_{HB} = 0.044$, $d = 0.137$). Thus, this provides

evidence that results from Experiment 1 were not solely driven by effort avoidance; as before,

only when participants were denied access to the appraisal preview period was optimal
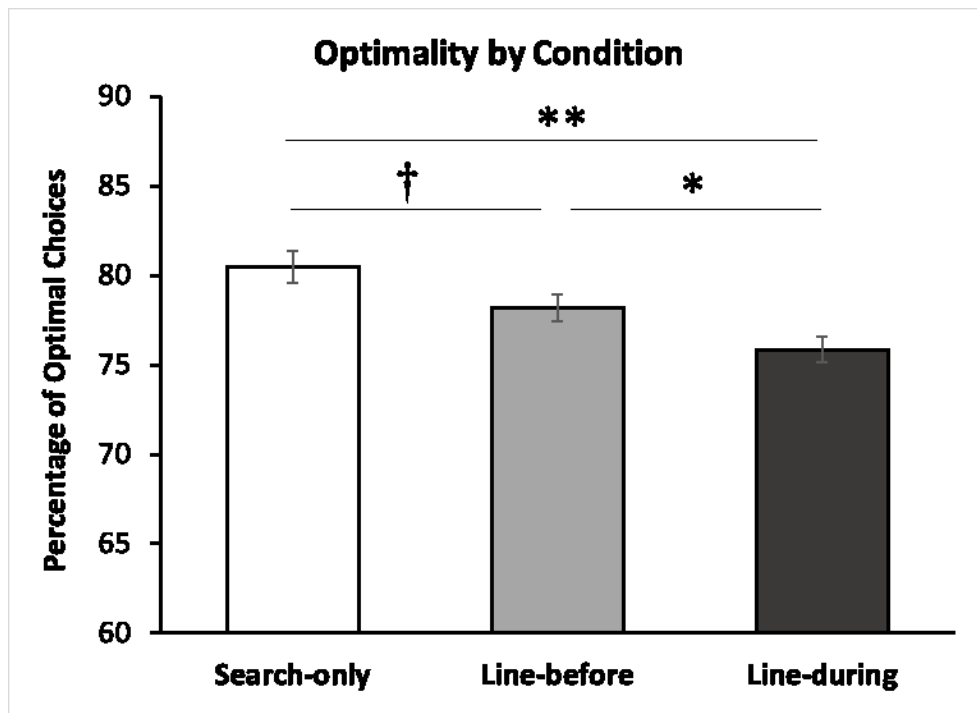
performance significantly lower.

Figure 6. Percentage of optimal choices made in each of the three conditions, on average. Error bars depict *within-subjects* standard error calculated using the Cousineau method with a Morey correction.
** *p* < 0.01, * *p* < 0.05, †*p* < 0.10.

Search: Response Time. As before, across participants, optimal choice percentages were significantly correlated with faster RTs (r = -0.593, $p < 0.001$). Additionally, mean search RTs varied significantly between the conditions, $F(2,46) = 10.508$, $p < 0.001$, $\eta_p^2 = 0.314$ (Figure 7). Again, there was no difference in search RTs between the search-only condition (*M* = 2685ms, *SD* = 768ms) and the line-before condition (*M* = 2749ms, *SD* = 681ms), $t(23) = -0.690$, $p_{HB} = 0.497$, $d = 0.088$. However, participants were significantly slower in the line-during condition (*M* = 2992ms, *SD* = 718ms) compared to both the search-only condition ($t(23) = -3.729$, $p_{HB} = 0.002$, $d = 0.413$) and the line-before condition ($t(23) = 4.896$, $p_{HB} < 0.001$, $d = 0.347$).
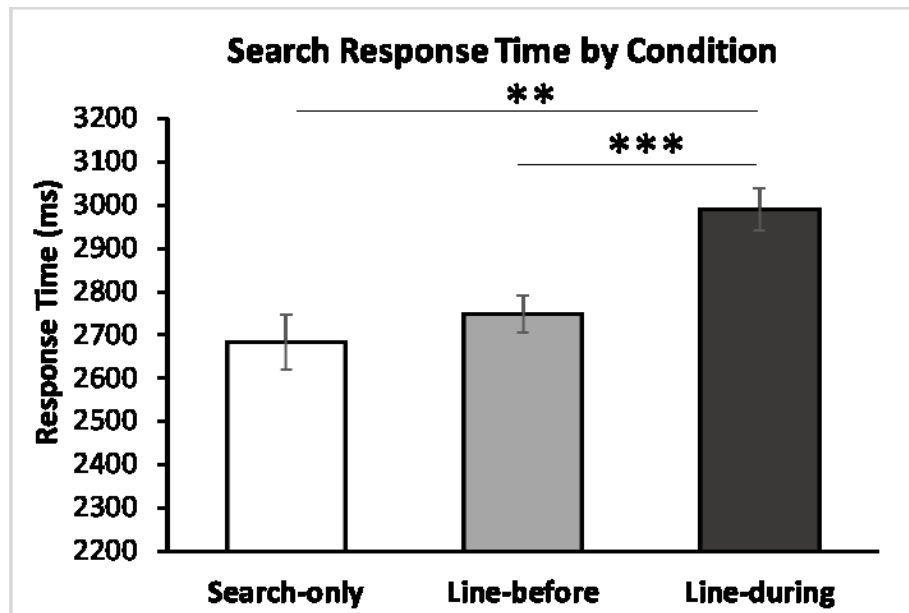
Figure 7. Average response times to find a target in the search task, in milliseconds, in each of the three conditions. Error bars depict within-subjects standard error calculated using the Cousineau method with a Morey correction. *** $p < 0.001$, ** $p < 0.01$.

Line Task. With unpredictable line judgments, there was no significant difference between the line-before condition (accuracy: $M = 90.23\%$, $SD = 6.98\%$; response time: $M = 595ms$, $SD = 52ms$) and the line-during condition (accuracy: $M = 91.69\%$, $SD = 5.81\%$; response time: $M = 599ms$, $SD = 53ms$) in either accuracy ($t(23) = 1.711$, $p = 0.101$, $d = 0.227$) or RTs ($t(23) = 0.897$, $p = 0.379$, $d = 0.076$).

Relationship between Line Task and Search Task. Like in Experiment 1, participants' average line task speed was not correlated with optimal search choices in either the line-before condition ($r = 0.158$, $p = 0.461$) or line-during condition ($r = 0.166$, $p = 0.438$), but marginal differences appear when taking a median split of individual trial RTs. Whether the line task was accomplished quickly or slowly did not vary significantly with search optimality in the line-before condition (fast $M = 78.59\%$, slow $M = 78.82\%$, $t(23) = 0.286$, $p = 0.778$, $d = 0.013$), but it *did* marginally vary with optimality in the line-during condition (fast $M = 77.74\%$, slow $M =$

75.15%, $t(23)$ = -1.977, p = 0.060, d = 0.150). The speed x condition interaction was marginally

signficant, $F(1,23)$ = 3.223, p = 0.086, η_p^2 = 0.123. Although not significant, this suggests that

tying up attentional resources for more time during the appraisal period, even when the line

task could not be predicted, may likewise lead to lower optimality.

   Cross-Experiment Analyses. Since Experiment 1 and Experiment 2 differed only in the

design of the blocks, cross-experiment analyses were conducted to determine how the

experimental results differed, if at all. Using a 2 (Experiment: 1 vs. 2, between-subjects) x 3

(Condition: search-only vs. line-before vs. line-during, within-subjects) ANOVA, the experiment

x condition interactions were assessed for the two search task dependent measures of interest

(optimality and search RT). There was no significant interaction in percentage of optimal

choices ($F(2,45)$ = 0.849, p = 0.435, η_p^2 = 0.036) or search RT ($F(2,45)$ = 0.374, p = 0.690, η_p^2 =

0.016), but a significant main effect of experiment was found for percentage of optimal choices

($F(1,46)$ = 5.997, p = 0.018, η_p^2 = 0.115), such that participants in Experiment 1 were overall

more optimal than participants in Experiment 2. This may have been related to the

predictability of trials in Experiment 1. We also assessed line task RT in a 2 (Experiment: 1 vs. 2,

between-subjects) x 2 (Line task condition: before vs. during, within-subjects) ANOVA. As might

be expected, RT for the line judgment was significantly affected by experimental design ($F(1,46)$

= 6.424, p = 0.015, η_p^2 = 0.123), and there was a significant interaction between experiment

and line task condition, $F(1,46)$ = 14.945, p < 0.001, η_p^2 = 0.245. More specifically, participants

were overall slower at completing the line judgment in Experiment 2 when the line judgment

was unpredictable than in Experiment 1 when the line judgment could be anticipated, and a

significant difference in line task RT between conditions was only found in Experiment 1.

Additionally, a 2 (Experiment: 1 vs. 2, between-subjects) x 2 (Line task condition: before vs.

during, within-subjects) x 2 (Line task speed: fast vs. slow, within-subjects) ANOVA was

conducted on optimality, to assess changes in strategy over fast and slow line speed trials.

Significant within-subject effects were found for optimality based on line judgment speed and

the condition x speed interaction, as was discussed in the individual experiments, but

experiment did not enter into any significant between-group interactions or main effects.

Overall, the cross-experiment analyses did not reveal noticeably different patterns of results

between the two experiments.

<center>**General Discussion**</center>

When approaching a dynamic visual search task in the real world, people are not always

optimal in the strategy they choose. There may be many contributing factors that influence

how an individual configures their attentional control settings in these situations, such as desire

to minimize effort or willingness to monitor performance and update when there is a conflict.

We postulated that one underlying factor might be the ability to appraise the search

environment for changes and use the obtained information effectively, and we surmised that

making it more difficult for appraisal to occur would lead to fewer optimal choices. Previous

work has shown that appraising statistical information from a display, such as enumerating

mean set size, is a rapid and highly accurate process (Ariely, 2001; Chong & Triesman, 2003;

Chong & Triesman, 2005a; Chong & Triesman, 2005b), even when multiple colored subsets are

present (Chong & Triesman, 2005b; Halberda, Sires, & Feigenson, 2006). However, to the best

of our knowledge, experiments probing statistical representation and enumeration have used

tasks that only require participants to judge the numerosity or statistical representation of

interest. Here, we apply these concepts to a more complex dynamic visual search task: now, participants must use the information they gain from these set representations to update their cognitive control settings and choose which target they want to search for on each trial.

To test this, this paper presented two experiments that used a modified version of the Adaptive Choice Visual Search task, adding a colored preview of the display to provide a window of time in which participants could appraise the display prior to commencing search. The experiments also introduced an irrelevant secondary line task either *before* or *during* this preview period to interfere with the appraisal process and the ability to form statistical representations of the colored subset sizes before the onset of the search. Experiment 1 explored the relationship between disruption of appraisal by the line task and subsequent search optimality, and Experiment 2 addressed an alternative *effort avoidance* explanation of the results.

Both Experiments 1 and 2 converged on similar findings: There was no significant difference in optimal choices when the line task occurred before the preview period relative to a search-only baseline, but participants made significantly fewer optimal choices when the line task occurred during the preview period, suggesting that the ability to appraise and form statistical representations during the preview period contributes to optimal performance. Along with this, search RTs were significantly slower in both experiments when the line task was presented during the preview period, but not significantly different from baseline when the line task was presented before the preview period.

Although RTs in Experiment 2 were generally slower and performance less optimal than in Experiment 1, making the line task onset unpredictable in Experiment 2 did not significantly

alter strategies or use of the appraisal period. There was a marginally-significant trend towards

fewer optimal choices in the line-before condition compared with baseline, which may suggest

that effort avoidance had some impact on the block-wide strategy that participants selected.

However, effort could not account for the entire effect, as optimality was still significantly

worse in the line-during condition than in the line-before condition.

As mentioned in the discussion of Experiment 1, one might argue that the performance

deficits in the line-during condition may be a result of a switch cost between two temporally

sequential tasks, because the line task immediately precedes the onset of the search in the line-

during condition but not in the line-before condition. Switching to a new task has been shown

to increase RTs relative to performing a repeated task, with the RT difference constituting the

switch cost (e.g., Monsell, 2003). A switch cost has been reliably shown even when an

individual knows when to expect the switch – i.e., as in the blocked design of Experiment 1 –

and when the two tasks are spaced sufficiently apart (Meiran, 1996; Rogers & Monsell, 1995).

Rogers and Monsell (1995) observed no reduction in switch cost when response-stimulus

intervals varied unpredictably up to 1200ms, and only about a one-third reduction when the

intervals were predictably blocked – a substantial "residual" switch cost still remained. Thus,

although increasing the delay between tasks has been shown to reduce the associated switch

cost, it is by no means eliminated (Allport et al., 1994; Rogers & Monsell, 1995). Given that

participants are switching between the same cognitive tasks in both the line-before and line-

during conditions, we would expect to see evidence of switch costs in both conditions. More

specifically, task-switching costs should imply an RT increase for line-before trials relative to

search-only trials, and this was not the case in these experiments. Since the presence of the line

task was only detrimental when it occurred during the preview period, not before, then the

opportunity to appraise must be at least somewhat contributing to an individual's ability to be

optimal; although switch cost may play some role, it is unlikely to fully account for these results.

An intriguing finding mentioned in the discussion of Experiment 1 was that the extent to

which the line task disrupted the preview period (i.e., whether the line task was completed

quickly or slowly) seemed to correspond with optimal performance – but only in the line-during

condition. This finding suggests that participants did not completely abandon the optimal

strategy when presented with a disruption during the appraisal period, and still attempted to

use the strategy on a trial-by-trial basis whenever time afforded them to do so. However, when

participants had no leftover time after the line task to appraise the display, they actually

adopted suboptimal strategies during the subsequent search. Thus, cognitive control is most

impaired when attentional resources are tied up for a longer duration of the preview period.

Nevertheless, secondary task difficulty and preview period duration were not directly

manipulated in these experiments, so these conclusions are speculative.

The present experiments sought to explore the role of appraisal in search optimality,

extending classic enumeration tasks by challenging participants to use acquired statistical

summary representations of the display to inform their search strategy. Given the performance

deficits in the line-during condition, our results support the idea that the ability to appraise the

environment and use resultant statistical representations effectively is a key factor underlying

why individuals make suboptimal choices when approaching visual search tasks. Just like when

you are tasked with finding your red SUV in a crowded parking lot, being able to quickly analyze

the particular features of the current environment is necessary in order to search most

efficiently and quickly locate your vehicle. If, for instance, your attention is disrupted by a text

or phone call while you are approaching the parking lot and you are unable to appraise your

environment, you may find yourself searching suboptimally. This finding, that engaging in

environmental appraisal is necessary to be optimal, seems intuitive; however, this issue has

been neglected in the literature, possibly because most visual search tasks do not afford

participants the freedom to decide between optimal and non-optimal search strategies or

assess their choices. Here, by using a paradigm that affords individuals a choice of what they

search for and then analyzing subsequent strategy usage, we offer evidence that the ability to

appraise contributes to an individual's choice of strategy selection during visual search.

One might wonder if appraisal during the preview period may differentially benefit

performance depending on the particular search parameters. For instance, since the current

version of ACVS alternated the optimal target color in unpredictable runs of one to six trials in

length, appraisal may be more helpful when the optimal target color switches from the

preceding trial compared to when the optimal target color repeats. If this is the case, then trials

in the line-during condition in which the optimal target color repeats may suffer less (or not at

all) from the presence of a secondary task during the preview period. In a post hoc analysis, we

looked at optimality on trials in which the optimal color did not match the preceding trial

("switch" trials, i.e., position 1 of a run) compared to optimality in which the optimal target

color repeated ("repeat" trials, i.e., positions 2-6 of a run), using a 3 (Condition: search-only vs.

line-before vs. line-during, within-subjects) x 2 (repetition type: switch vs. repeat, within-

subjects) ANOVA for each experiment. Experiment 1 (Figure 8A) and Experiment 2 (Figure 8B)

both showed significant main effects of condition and repetition type, but critically no condition

x trial type interaction (Experiment 1: $F(2,22) = 0.737$, $p = 0.490$, $\eta_p^2 = 0.063$; Experiment 2:

$F(2,22) = 0.639$, $p = 0.537$, $\eta_p^2 = 0.055$). Given that the performance deficit for switch trials in

the line-during condition appears numerically larger than the deficit in the other conditions, we

did a combined analysis using both Experiment 1 and Experiment 2, in case the lack of

interactions separately was due to being underpowered. However, when pooling data across

experiments, the interaction was still non-significant ($F(2,46) = 1.402$, $p = 0.256$, $\eta_p^2 = 0.057$).

With a larger sample and sufficient power, it is possible a small interaction between the

benefits of appraisal and repetition type might be observed; future work is needed to examine

this in more depth.
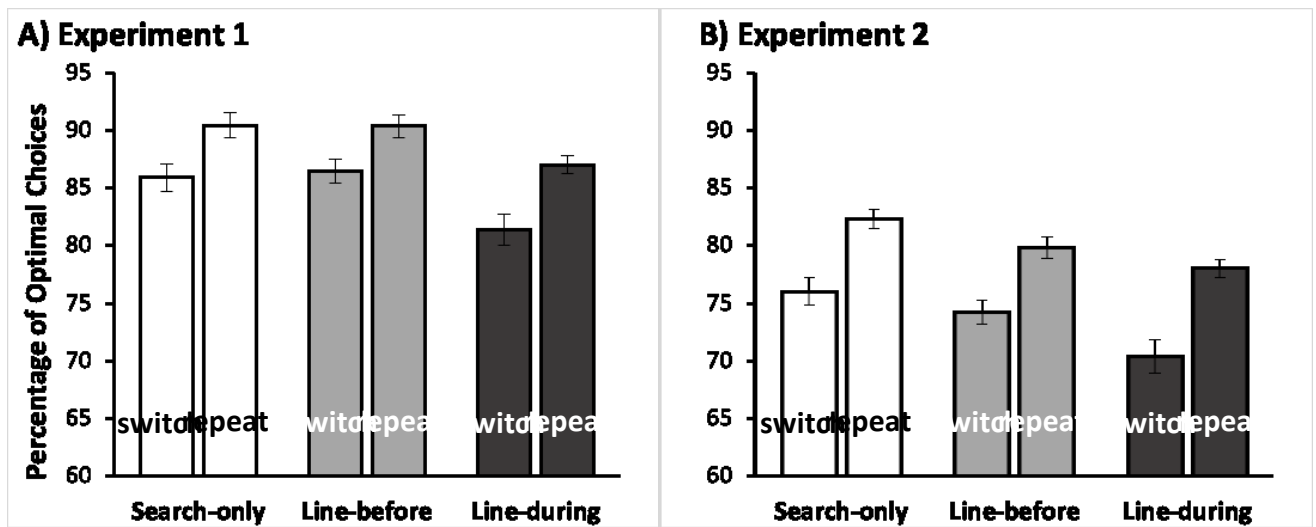
Figure 8. Percentage of optimal choices made in each of the three conditions, on average, based on whether the optimal target switched or repeated from the previous trial. Error bars depict within-subjects standard error calculated using the Cousineau method (Cousineau, 2005) with a Morey correction (Morey, 2008). A) Results from Expertiment 1. B) Results from Experiment 2.

Additionally, it is worth noting that visual search tasks might benefit from an appraisal

period differently. Our paradigm is constructed such that the color makeup of the search arrays

varies from trial to trial, and appraisal may play a different role in situations where the search

environments are more consistent. For example, a large body of work in radiology has sought

to explore strategies for improving identification of abnormalities in chest X-rays (e.g.,

Auffermann, Krupinski, & Tridandapani, 2018). Chest X-rays all contain the same general

physical structure and color makeup, so engaging in an appraisal period before viewing each X-

ray may not gain the radiologist any new meaningful information that would change their

search of that X-ray. Conversely, a real-world task such as airport baggage screening, in which

each instance of the search unpredictably varies in physical structure, may benefit more from

an appraisal period; perhaps the size or shape of a carry-on item directly informs where a

banned item is most likely to be found, and appraising the size or shape first can lead to a more

efficient search. Further, appraisal may uniquely benefit complex visual search tasks or tasks

with multiple targets (e.g., foraging: Cain et al., 2012; Wolfe 2013) moreso than singleton

detection paradigms. In the latter, attention may be deployed automatically in a bottom-up

fashion, and therefore an appraisal period may not have the benefit that it would when top-

down attentional mechanisms are deployed. Thus, engaging in environmental appraisal is

necessary to be optimal, but perhaps moreso in search tasks in which environments are

unpredictably changing and targets are not singletons.

   One limitation of this study is that we are unable to determine conclusively the

mechanism by which disrupting appraisal lowers optimality; future work will need to explore

whether a secondary task impairs the ability to acquire statistical summary representations

altogether, or merely the ability to use the information. We can offer some speculation,

however: given that individuals can enumerate different colored sets with relative set sizes

that are far closer in ratio than ACVS – which used a 2:1 ratio – with great speed and accuracy

(e.g., see Halberda, Sires, & Feigenson, 2006), statistical representations should be acquired easily and the mechanism is likely the latter.

Future work is necessary to further explore the characteristics of this appraisal period. How much appraisal time is required in order to reap the benefits and perform optimally? For instance, would secondary tasks that are elevated in difficulty or duration completely impair any residual ability to appraise, akin to the slow line task trials described above? Is this appraisal period only beneficial if immediately preceding onset of the search, or can appraisal of the environment be made in advance and statistical representations stored until applicable? How does the ability to appraise the environment for featural changes influence performance monitoring and strategy updating? There is still much left to explore regarding how environmental appraisal operates, but the findings presented here push us forward in our understanding that appraisal significantly contributes to configuration of optimal control settings.

Acknowledgements

Open Practices Statement

Data are available upon request. None of the experiments reported here were preregistered.

References

Allport, D.A., Styles, E.A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic

control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp.

421-452). Cambridge, MA: MIT Press.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,*

157-162.

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological*

Science, 15, 610–615.

Auffermann, W. F., Krupinski, E. A., & Tridandapani, S. (2018). Search pattern training for

evaluation of central venous catheter positioning on chest radiographs. Journal of

Medical Imaging, 5(3), 031407.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception*

& Psychophysics, 55, 485–496.

Botvinick, M. M., & Rosen, Z. B. (2009). Anticipation of cognitive demand during decision-

making. *Psychological Research, 73*, 835–842.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict

monitoring and cognitive control. *Psychological Review, 108*(3), 624–652.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural

mechanism. *Annual Review of Psychology, 66,* 83–113.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework.

Trends in Cognitive Sciences, 16, 106–113.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working

     memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold,

     M. C.Kane,A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76–106).

     New York: Oxford University Press.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A Bayesian optimal foraging model of

     human visual search. *Psychological Science, 23,* 1047–1054.

Chatham, C. H., Frank,M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a

     developmental shift in the temporal dynamics of cognitive control. *Proceedings of the*

     National Academy of Sciences, 106, 5529–5533.

Chong, S.C., & Triesman, A. (2003). Representation of statistical properties. *Vision Research, 43,*

     393-404.

Chong, S.C., & Triesman, A. (2005a). Attentional spread in the statistical processing of visual

     displays. *Perception & Psychophysics, 67,* 1-13.

Chong, S.C., & Triesman, A. (2005b). Statistical processing: Computing the average size in

     perceptual groups. *Vision Research, 45,* 891-900.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to

     Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1,* 42-

     45.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent

     on attentional control settings. Journal of Experimental Psychology: Human Perception

     and Performance, 18, 1030–1044.

Green, B. F., & Anderson, L. K. (1956). Color coding in a visual search task. Journal of

   Experimental Psychology, 51(1), 19-24.

Halberda, J., Sires, S.F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be

   enumerated in parallel. *Psychological Science, 17*(7), 572-576.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of*

   Statistics, 6, 65–70.

Irons, J. L., & Leber, A. B. (2016). Choosing attentional control settings in a dynamically changing

   environment. *Attention, Perception, & Psychophysics, 78*, 2031–2048.

Irons, J. L., & Leber, A. B. (2018). Characterizing individual variation in the strategic use of

   attentional control. *Journal of Experimental Psychology: Human Perception and*

   Performance, 44, 1637-1654.

Irons, J. L., Hansen, H.A., & Leber, A. B. (in prep). Why do people use suboptimal attentional

   control strategies?

Irons, J.L., & Leber, A.B., (2018, abstract) Ignorance vs. laziness: Why do people use suboptimal

   attentional control strategies? *Journal of Vision, 18,* 630.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in

   psychtoolbox-3. *Perception*, *36*(14), 1-16.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the

   avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139,* 665–

   682.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. Trends in Cognitive

   Sciences, 9(2), 75-82.

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: Behavior, brain

      activation, and individual differences. *Cognitive, Affective, & Behavioral Neuroscience, 8,*

      99–112.

Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and

      reduced cingulate activity during mental fatigue. *Cognitive Brain Research, 24,* 199–205.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of*

      Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1423-1442.

Monsell, S. (2003). Task Switching. *Trends in Cognitive Sciences, 7*(3), 134-140.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau

      (2005). *Tutorial in Quantitative Methods for Psychology, 4*(2), 61-64.

O'Leary, A., & Sloutsky, V. M. (2017) Carving metacognition at its joints: Protracted

      development of component processes. *Child Development, 88,* 1015-1032.

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

      into movies. *Spatial Vision, 10*, 447–466.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive

      tasks. *Journal of Experimental Psychology: General*, *124*, 207–231.

Tsal, Y., & Benoni, H. (2010). Diluting the burden of load: perceptual load effects are simply

      dilution effects. *Journal of Experimental Psychology: Human Perception and*

      Performance, 36(6), 1645-1656.

Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. (2003). Functional

      specialization within medial frontal cortex of the anterior cingulate for evaluating effort-

      related decisions. *The Journal of Neuroscience, 23*, 6475–6479.

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

human visual search. *Journal of Vision, 1*3(3):10.