Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/15/$12.00 http://dx.doi.org/10.1037/a0039429

Spatial Attention and Environmental Information

William Prinzmetal
University of California, Berkeley

Kelly L. Whiteford
University of Minnesota

Joseph L. Austerweil
Brown University

Ayelet N. Landau
The Hebrew University of Jerusalem

Navigating through our perceptual environment requires constant selection of behaviorally relevant information and irrelevant information. Spatial cues guide attention to information in the environment that is relevant to the current task. How does the amount of information provided by a location cue and irrelevant information influence the deployment of attention and what are the processes underlying this effect? To address these questions, we used a spatial cueing paradigm to measure the relationship between cue predictability (measured in bits of information) and the voluntary attention effect, the benefit in reaction time (RT) because of cueing a target. We found a linear relationship between cue predictability and the attention effect. To analyze the cognitive processes producing this effect, we used a simple RT model, the Linear Ballistic Accumulator model. We found that informative cues reduced the amount of evidence necessary to make a response (the threshold), regardless of the presence of irrelevant information (i.e., distractors). However, a change in the rate of evidence accumulation occurred when distractors were present in the display. Thus, the mechanisms underlying the deployment of attention are exquisitely tuned to the amount and behavioral relevancy of statistical information in the environment.

Keywords: voluntary attention, information theory, reaction time, linear ballistic accumulator model

Navigating through our perceptual environment requires constant selection of behaviorally relevant information and inhibition of irrelevant information. Attention entails selection of stimuli that are relevant to the current task. Information about where relevant and irrelevant stimuli are in the environment could guide the deployment of attention. In addition the presence of irrelevant information in the environment (i.e., distractors) should too have an effect on attentional selection. Contingencies between, for example, a cue stimulus and a target stimulus, could serve to guide attentional selection. In turn the amount of information a cue provides about where an upcoming target appears may affect the degree of performance benefits because of attention. In this article, we explore how the amount of information provided by a cue affects the deployment of attention to relevant and irrelevant stimuli by manipulating the amount of information provided by the environment about where to attend, and the presence of irrelevant information.

We measured attentional deployment using the spatial cueing paradigm (Posner, 1980; Posner, Snyder, & Davidson, 1980). In this paradigm, observers are instructed to perform a discrimination on target stimuli, which are preceded by a cue. The cue can be highly informative, in which case most targets will be preceded by a cue that validly indicates its location. In other instances, a cue might be completely uninformative, in which case the target will be randomly positioned with respect to the cue. In three experiments, we systematically varied the amount of information provided by the cue. As discussed below, the functional relation between information provided by cues and attention, measured by reaction time (RT), constrain theories of how attention is deployed and may provide a powerful tool for understanding and measuring attentional deployment.

In addition to varying degrees of information between different stimuli, our environment typically contains irrelevant stimuli that distract us from our target, despite cues guiding us to the target. To explore the effect of distractors on the deployment of attention and how it interacts with the amount of information provided by a cue, we also varied the presence of distractors in a display (irrelevant information). The presence of distractors should impact the utility of spatial cues. One might think that the presence of distractors would enhance the effect of a spatial cue (because the cue helps guide the deployment of attention to the relevant stimulus and away from the irrelevant stimuli). However, previous studies investigating *involuntary* attention using peripheral, noninformative spatial cues, have shown that including distractors decreases the effect of the cue (Mordkoff, Halterman & Chen, 2008; Prinzmetal et al., 2010, 2011). The present article, therefore, explores two related issues in the domain of *voluntary* attention. First, we

William Prinzmetal, Department of Psychology, University of California, Berkeley; Kelly L. Whiteford, Department of Psychology, University of Minnesota; Joseph L. Austerweil, Department of Cognitive, Linguistic, and Psychological Sciences, Brown University; Ayelet N. Landau, Department of Psychology and Department of Cognitive Sciences, The Hebrew University of Jerusalem.

Correspondence concerning this article should be addressed to Ayelet N. Landau, Department of Psychology and Department of Cognitive Sciences, The Hebrew University, Mt. Scopus, Jerusalem 9190501, Israel. E-mail: ayelet.landau@gmail.com

systematically investigated how the amount of information exhibited by the presence of an environmental cue (at fixation) affects performance. Second, we examined whether the pattern of information utility is affected the presence of distractors in the display. Finally, we analyzed the pattern of RT performance using a formal model (linear ballistic accumulator) to discern different psychological mechanisms that could account for different effects of voluntary spatial attention, with and without distractors, on human performance.

Since the Cognitive Revolution (Gardner, 1987; Miller, 1956, 2003), psychologists have explored ways to quantify the amount of information processed by a perceiver and how information is used in cognitive tasks. For example, Miller (1956) discussed "the magical number seven, plus or minus two," as a potential capacity limit on the number of structures that can be simultaneously processed by the mind, and the mathematics of information theory is often used for understanding cognitive processes (Cover & Thomas, 1991; Shannon, 1948). Researchers have found a strong effect of the amount of statistical information provided by a cue on various cognitive and perceptual processes, such as judgment and decision making (McKenzie, 2004; Oaksford & Chater, 1994, 2007), category learning (Nelson, 2005; Nelson, McKenzie, Cottrell, & Sejnowski, 2010), choice (Hick, 1952; Hyman, 1953), visual search and memory search (Wolfe, 2012), eye movements during search (Najemnik & Geisler, 2005; Renninger, Verghese, & Coughlan, 2007), and reading (Legge, Klitz, & Tjan, 1997). For example, the time that people take to make a choice increases linearly with the number of alternatives in discrimination tasks, when plotted in bits of information; hence, a 2-alternative forced choice task is 1 bit of information, 4-alternatives is 2 bits, 8-alternatives is 3-bits, and so forth. This relation is called the Hick-Hyman Law (Hick, 1952; Hyman, 1953). Fitts Law is a similar relation applied to reaching and other voluntary motor movements (Fitts & Posner, 1967, chapt. 6). Recently, Wolfe (2012) found an analogous relation for memory search within large memory sets. In all cases, RT was linear when plotted in bits of information. Alternatively, results in traditional visual search (Wolfe, 1998) and memory search (Sternberg, 1966) tasks found that search time is linear in the number of options, not in the bits of information (the logarithm of the number of options) as would be expected if these tasks also were governed by the amount of information.

Our version of the spatial cueing task is illustrated in Figure 1. Each trial began and ended with a fixation screen that included a central dot and marked six possible target positions (with a rectangle per position). Participants were instructed to maintain fixation on a central fixation point, and eye movements were monitored. A cue (either an arrow or a barbell, see Figure 1) indicated one of the locations, followed by a target, which contained either the letter "F" or "T." On *valid* trials, the target was in the cued location. On *invalid* trials, the target was in an uncued location. Participants responded with a button press indicating whether the target was an "F" or a "T"; RT was the dependent variable.

Across blocks, we varied how often the cue correctly predicted the location of the target. For example, in the uninformative condition, the cue provided no information as to the target location. Hence, the target was equally likely to be in any location. For example, if there are six locations that the target can appear, it appeared in the cued location on 1/6 of the trials (16.7% valid) and in an uncued location on 5/6 of the trials (83.3% invalid). In the most informative condition, we reversed the proportions of valid and invalid trials (i.e., the cue was valid on 5/6 or 83.3% of the trials). There were four predictability conditions, listed in Table 1. Note that the present investigation of attention effects under different probability manipulations is concerned with voluntary, or endogenous, spatial attention rather than effects of involuntary attentional capture. Our experimental procedures and stimulus choice reflect this focus (cue to target interval, cue types; see below).

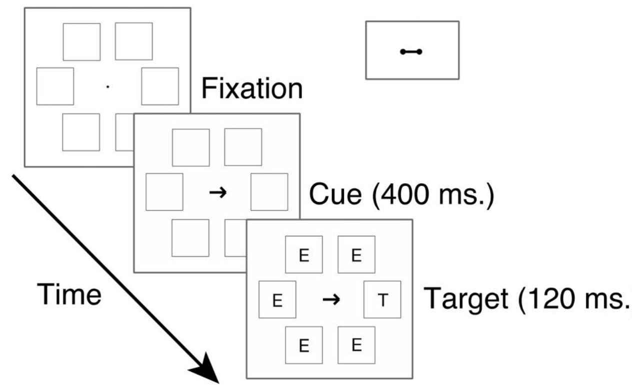

Figure 1. The sequence of events for each trial. The distance from the fixation to the center of each square subtended 4.29 degrees of visual angle. In Experiment 1, the nontarget locations contained distractors. The figure is drawn to scale. The barbell cue is from Experiment 2 shown in the box. One end of the barbell cue was red, and the other blue.

To gain an intuition for how the amount of information varies with the probability of a cue correctly predicting the target, consider the following thought experiments. Note that this is not the design in our present experiments (but see Close, d'Avossa, Sapir, & Parkinson, 2010). Suppose that there were only two possible display locations and two cue predictability conditions (related to information, as seen below): A random cue (no information, equally valid and invalid trials) or a 100% predictive cue (fully informative, valid trials only). The 100% predictive cue would provide 1 bit of information, whereas the random cue would provide 0 bits of information. If there were four locations, the 100% informative cue would provide 2 bits of information. If there were eight locations, the 100% informative cues would provide 3 bits of information, and so forth. Intuitively, the number of bits of information is the minimum number of binary questions one would need to ask to determine the target location. The problem with varying the number of nontarget items in the display is that increasing display-size affects many visual properties such as crowding and not just the information provided by the cue.

In the present experiments, with a fixed number of locations and different cue probabilities, information gained from observing the cue (Info Gain) is calculated as follows:

$$InfoGain = H_{No\ information} - H_{Cue} \quad (1)$$

H is Shannon's entropy (Shannon, 1948), which measures the amount of uncertainty inherent in a probability distribution. It is at its maximum when the target occurs in each location with equal probability (no information), and is reduced when the cue provides

Table 1
The Cue Predictabilities Used in Experiments 1–3

Condition	% Valid	% Invalid	Info (bits)
No information			
1/6 cued	16.7	83.3	0
5/6 uncued	53.6	46.4	0.51
5/6 cued	70.2	29.8	1.02
Most information			
1/6 uncued	83.3	16.7	1.55

information as to the location of the target. H is defined in the standard manner:

$$H = -\sum_i p_i \ \log_2 \ p_i \tag{2}$$

where p_i is the probability that the target appears at location i. For example, with 6 locations and the target occurring at the cued location on 5/6 of the trials:

$$H_{Cue} = -\left(\frac{5}{6}\log\frac{5}{6}\right) - \left(\frac{1}{30}\log\frac{1}{30}\right) - \left(\frac{1}{30}\log\frac{1}{30}\right) - \left(\frac{1}{30}\log\frac{1}{30}\right) - \left(\frac{1}{30}\log\frac{1}{30}\right) - \left(\frac{1}{30}\log\frac{1}{30}\right) \approx 1.04$$

$H_{No\ information}$ is approximately 2.59 bits, and so the information gained by this cue is $2.59 - 1.04 \approx 1.51$ bits.

We created the four cue predictability conditions so that they were equal steps in terms of the amount of information gained from observing the cue (see Table 1). We operationalize the effect of attention by quantifying the effects of different cue types on RTs as follows:

$$\text{Attention} = \text{Invalid trials RT} - \text{Valid trials RT}. \tag{3}$$

The difference between invalid and valid trial RTs (i.e., the attention effect) is commonly used to operationalize selective spatial attention (Posner, 1980; Prinzmetal et al., 2005). For example, a participant allocating no attention toward the cue would have RTs that are the same regardless of cue validity (i.e., attention effect = 0).

What is the quantitative relation between amount of information gain and the attention effect? The form of the function will depend on the units of the independent variable (e.g., proportion valid trials or bits of information), but for the purposes of this discussion, assume that the attention effect is plotted against bits of information (see Table 1). There are three possible outcomes. First, as predictability increases, attending more to the cued location might have diminishing returns in terms of performance (Shaw & Shaw, 1977, p. 203) because one is already attending to the cued location and it is unclear how much it would help to attend to it "more." The idea is called the marginal-gain hypothesis. This predicts that the attention effect would be a negatively accelerating function of cue predictability. In other words, once the cue is at all predictive (over chance), observers fully attend to the cued location and increasing cue predictability has no effect (i.e., the attention effect is already at its "ceiling"). Another hypothesis that makes a similar general prediction is the probability-matching hypothesis, where attention effects are a linear function of cue predictability measured as cue probability rather than information.[1] Support for the probability-matching hypothesis is found in the visual search paradigm. There, the search RT scales linearly with the number of possible target locations (set size). In a display where target location is random, such a linear effect over set size essentially corresponds to a linear scaling with a given location's probability of containing the target (Wolfe, 1998).

Second, one might predict that people attend to a cued location to the degree that it provides information about where the target is located. Thus, this third hypothesis, the information-gain hypothesis, predicts that the attention effect is a linear function of cue predictability measured in bits of information. If the cueing data follow the information-gain hypothesis, they will be consistent with memory set size effects in hybrid visual and memory search performance, which also increase linearly when measured in bits of information (Wolfe, 2012).

Finally, one might predict that people would be insensitive to the cue unless it is clearly predictive. This information-threshold hypothesis predicts that the cue would be ignored unless the cue is highly predictive of the target location. Accordingly, the attention effect would be a positively accelerating function of bits of information. These hypotheses form a continuum of how environmental information might affect attention.

The form of the relation between cue probability and the attention effect might also depend on whether the display contains irrelevant information, that is, distractors (as in Figure 1), or not. If there are no distractors in the display (or the target is unique in some attribute) it may "pop-out" of the display and in a sense be self-cueing. In these circumstances, it might not be worth orienting to the cue unless the cue is highly predictive (information-threshold). In terms of information (bits), the cueing effect would be positively accelerating. Experiment 3 examines the effect of cue probability without distractors. With distractors, finding the target is substantially more difficult, as the target location may no longer pop-out of the display. Hence, any information provided by the cue is going to be helpful. It might behoove participants to attend to the cue even if it is only slightly predictive. The results might correspond to the marginal-gain or information-gain hypothesis.

Previous investigators have found that the effect of attention increases with the amount of information provided by the cue. For example, Eriksen and Yeh (1985) found that cue predictability influences the effect of attention in a task with a peripheral cue. However, a peripheral cue affects both voluntary and involuntary attention, so their effect cannot be attributed solely to voluntary attention. Close et al. (2010) as well as Mordkoff et al. (2008) varied the cue information by varying the number of behaviorally relevant locations from trial to trial. Comparing different conditions of distributed attention (e.g., to one vs. two locations) might

[1] The logic for why the probability-matching hypothesis predicts a linear attention effect is as follows. Assume that t_0 is the time to respond after observing the target, t_s is the amount of time it takes to switch locations, the cue validity is p, and there are K-1 otherwise equally likely uncued locations; then, the expected RT for a valid trial is $p\ t_0$ and the expected RT for an invalid trial is $p\ t_0 + (1\text{-}p)\ K/2\ t_s$ (the probability of finding the target on the n-th try in an uncued location is $1/(K\text{-}1)$). The difference of these is the expected attention effect and is clearly linear in p.

involve a number of processes in addition to voluntary attention (e.g., attentional sampling, Landau & Fries, 2012). Hence, previous research has not investigated the function relating cue information to the attention effect with purely voluntary attention.

This article is divided into two parts. In the first part, we examine how manipulating the amount of information gained from observing a cue influences the attention effect in the presence and in the absence of distractors. To anticipate our results, in Experiments 1–3 the attention effect increased in a linear manner with the amount of information, in bits, provided by the cue, consistent with the information-gain hypothesis, but not the other hypotheses. We found this pattern of results both with and without distractors.

In the second part of this article, we investigated the mechanisms underlying the relation between information gain and attention. In particular, we analyze the RTs by modeling the attention effect in the results of Experiments 1–3 using a standard RT model, the Linear Ballistic Accumulator (LBA; Brown & Heathcote, 2008). The LBA model provides a method for examining the entire distribution of RT, not just the means. The LBA model extracts processing threshold and rate parameters from the RT distribution, which reflect the amount of evidence needed to make a decision and the rate of evidence accumulation. We found fundamental differences in processing depending on the amount of information provided by the cue and, importantly, whether there were distractors in the display. These fundamental differences were not apparent when examining only the mean RT. Alternatives to the LBA model, including a variety of serial models, are discussed in the General Discussion.

Experiment 1

The purpose of Experiment 1 was to ascertain the relation between cue predictability and the attention effect, as measured by Equation 3. Each participant was tested with four different cue probabilities (proportion valid trials) described in Table 1 in separate sessions. The sequence of stimulus events during a trial is shown in Figure 1.

Method

Procedure. The experiment took place on two separate days, with each day containing two sessions of five blocks. The probability of the cue correctly predicting the target location (cue predictability) varied in equal steps of information gain across the sessions: 16.7% (uninformative condition), 53.6%, 70.2%, or 83.3% (most informative condition). The cue predictability was manipulated within-subjects over different sessions. At the start of each session, participants were informed of the percent of trials that the cue correctly predicted the target location. In the uninformative cue condition, participants were told to ignore the cue, as it was random with respect to the target location and would not help them predict the target location. For all other cue probabilities, participants were instructed to attend to the cue. More important, all conditions included the same number of stimulus locations (cf. Mordkoff et al., 2008).

Participants were given a practice block of 32 trials at the beginning of each session. They continued to receive practice blocks until performance was at or above 90%. Each experimental block consisted of 84 trials, and the end of the first session of each day was followed by a short break (~5 min). Throughout the experiment, participants were asked to keep their eyes fixated on the fixation point at the center of the screen, surrounded by six square boxes. Eye movements were monitored with a video camera as described in Prinzmetal et al. (2005). As shown in Figure 1, a central arrow cue appeared for 400 ms. At the offset of the cue, the target letter ("F" or "T") appeared in one of the six possible locations for 120 ms, while the distractor letter "E" simultaneously appeared in the other boxes. The participants' task was to detect whether the target was an "F" or a "T" by pressing one of two keys. The target letter and the target location were determined randomly on each trial as was the cued location in invalid trials.

Participants were instructed to respond as quickly as possible while maintaining their accuracy. If they responded incorrectly, the computer-generated voice responded, "Incorrect." Responses longer than 1,500 ms were followed with the computer-generated response, "Too slow." When eye movements were detected, the computer-generated voice responded, "Eye movement." Block order was counterbalanced between participants such that every condition was equally often first, second, and so forth, and every condition followed every other condition once.

Stimuli. The task was displayed on a 17-in. iMac computer monitor with a 48 cm viewing distance. A chinrest held the viewing distance constant throughout the experiment. A fixation point was located at the center of the screen, surrounded by six square boxes. The target letters, "F" and "T," and distractor letters, "E," were black, size 36 Helvetica font. The screen resolution was 1,024 × 768 pixels. The central arrow cue, boxes, and fixation point were also black. The arrowhead shaft was 1.13 cm in length, and all stimuli were presented on a white background. Figure 1 is drawn to scale and the distance from the fixation to the center of each square subtended 4.29 degrees of visual angle.

Participants. There were eight participants, all of whom were undergraduate volunteers recruited through the University of California, Berkeley's Research and Participation Program. All participants gave informed consent and reported normal or corrected to normal vision. They were given course credit for their participation.

Results and Discussion

Before analysis, we removed incorrect trials (3.2%), trials with eye movements (1%), and trials with RTs exceeding 1.5 s (0.04%). Mean RTs for correct valid trials (trials where the target was validly cued) and correct invalid trials (trials where a location that did not contain the target was cued) are shown in Table 2 (*SD*s,

Table 2
The Valid and Invalid Correct RTs in Milliseconds for Experiments 1–3

Condition (in bits)	Experiment 1		Experiment 2		Experiment 3	
	Valid	Invalid	Valid	Invalid	Valid	Invalid
0	373 (91)	382 (83)	391 (85)	401 (89)	337 (103)	348 (97)
0.51	362 (81)	397 (95)	401 (94)	422 (102)	333 (96)	355 (91)
1.02	354 (83)	409 (99)	394 (95)	424 (101)	338 (97)	367 (93)
1.55	350 (83)	430 (109)	373 (92)	437 (107)	331 (105)	372 (95)

Note. The *SD*, averaged over participants, is shown in parenthesis.

averaged over participant, are in parentheses). As expected, participants were significantly faster on valid than on invalid trials, $F(1, 7) = 41.04$, $p < .01$, and this effect significantly interacted with the cue predictability, $F(3, 21) = 8.47$, $p < .01$. To illustrate the interaction, the attention effect (invalid trials RT – valid trials RT, as defined by Equation 3) is plotted in Figure 2 as open circles connected by a solid line. Error bars represent $\pm SE$ calculated for within subject designs using the method outlined by Morey (2008) for within-subjects designs.

In addition, the attention effect (see Figure 2) was significantly linear (trend test, linear component $F(1, 7) = 23.34$, $p < .01$). Neither the quadratic nor cubic trends approached significance (Fs <1.0). The linear trend accounted for 99.7% of the attention effect variance. (The *MSE* of the cue predictability by the attention effect interaction = 412.8.) To the extent that this generality holds, the relation between the amount of information supplied by the cue and the attention effect in RT is quite simple.

It is possible that the linear trend was the result of averaging over observers. Each observer's mean attention effect only has four values, so it is difficult to do an individual participant analysis equivalent to the trends analysis of variance (ANOVA) above. To overcome this problem, we devised a bootstrap analysis that allowed us to test for linearity of the attention effect within each observer's data. Instead of taking the means of valid and invalid trials and calculating an overall "attention effect" by subtracting the latter from the former, we created "individual-trial attention effects" by randomly pairing valid and invalid trials from each subject. For each of these pairs, we calculated an attention effect (invalid RT – valid RT). The number of pairings was limited to the condition with fewer trials and as a result the data were noisier than the actual data. Nonetheless, six of eight participants had a significant linear fit (making the linear effect significant at the group-level as well; Binomial test, $p < .05$); another participant had a linear fit approaching significance ($p = .09$). In addition, the linear fit was magnitudes better than quadratic or cubic fits for all but one participant, for which all fits were far from significant ($p = .75$, 0.55, and 0.46 for linear, quadratic, and cubic).

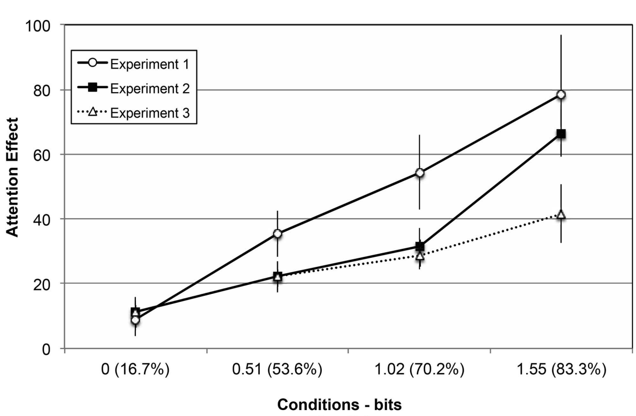

Figure 2. Attention effect for Experiments 1 to 3 plotted as a function of cue predictability for Experiments 1–3. Note that in all experiments the number of locations was held constant and the different attention conditions were created by manipulating the amount of information provided by the cue within a block carried.

If the results are considered in terms of the percent of valid trials, it is clear to see that they contradict the marginal-gain hypothesis. The marginal-gain hypothesis predicts that the attention effect will be a negatively accelerating function of cue predictability. However, going from 16.7% to 53.5% valid trials (a change of 36.8% validity) caused a 26 ms increase in the attention effect. Going from 72.2% to 83.3% valid trials (a change of 11.1%) caused almost the same increase in the attention effect (24 ms). The results are consistent with the information-gain hypothesis, which predicts our observed results, a linear relation between the information provided by a cue and the attention effect. As there is a clear attention effect at all levels of predictability, the results also contradict the information-threshold hypothesis. We discuss the differences between the information-gain hypothesis and the probability-matching hypothesis in the discussion.

Note that even with the uninformative cue condition (i.e., 16.7% valid), there was a small, yet reliable cueing effect (9 ms; $t(7) = 2.20$, $p < .05$). We explore two possible accounts for this effect. First, previous work has shown that a nonpredictive arrow cue can cause an involuntary cueing effect, because the pointiness of arrows implies directionality (e.g., Ristic & Kingstone, 2006; Tipples, 2002). If this is the case, then our effect is due in part to involuntary attention. This interpretation is tested explicitly in Experiment 2.

A second possible cause for this small attention effect at the uninformative (0 bit) condition is a carry-over effect from previous sessions, even though the order of probability conditions was counterbalanced across participants. To analyze this possibility, we divided the participants into those that had the uninformative condition the first day (Session 1 or 2) and those that had it the second day (Session 3 or 4). Those who had the uninformative condition on the first day had either one or zero predictive sessions before the uninformative condition. Those who had the uninformative condition on the second day had either two or three predictive sessions before the uninformative condition. The attention effect for the uninformative condition was a negligible 1.7 ms for those who had the uninformative condition on the first day. However, it was 17.4 ms for those who had the uninformative condition on the second day. This analysis suggests that the small cueing effect at the uninformative condition was a carry-over effect from previous sessions that used a predictive cue. In other words, participants who learned the task with a predictive cue still allocated attention to the random cue condition, even when instructed by the experimenter to ignore it. In addition to this post hoc analysis exploring the carry over possibility, Experiment 2 directly tests whether the attention effects found when the cue is uninformative were because of the involuntary attention engendered with an arrow.

Experiment 2

In Experiment 2, we used a barbell cue, which unlike an arrow, has no prior associations for its directionality. This experiment was identical to Experiment 1 except we replaced the arrow with a barbell that had a red solid circle at one end and a blue solid circle at the other (a method first used by Lambert, Roser, Wells, & Heffer, 2006). For one half of the participants, the red end was the "valid" end (indicated the target position on valid trials) and for the other half of participants, the blue end was the valid end. Participants were told which end of the barbell was the informative end at the beginning of the experiment and were reminded again at the start of Day 2. The same end of the barbell (red or blue) was informative during both days. When the target was not in the location indicated by the valid end, it appeared equally often in the other five locations (as in Experiment 1). Cue predictability condition was counterbalanced analogously to Experiment 1. Except for the previously mentioned differences (barbell cue vs. arrow cue), the experiment was identical to Experiment 1.

Method

Stimuli. The barbell shaft was the same length and width as the arrowhead shaft in Experiment 1 (see Figure 1). Each barbell ball was .35 cm in diameter and subtended 1.1 degree visual angle (the shaft was 2 px wide and the balls 10 px in diameter). One ball was blue while the other was red.

Participants. There were 16 participants, drawn for the same subject pool as in Experiment 1, who all received course credit for participating.

Results and Discussion

As in Experiment 1, we removed incorrect trials (4.5%), trials with eye movements (1%), and trials with RTs exceeding 1.5 s (0.13%) from analysis. Correct RTs for valid and invalid trials are shown in Table 2. As before, participants were significantly faster on valid than invalid trials, $F(1, 15) = 36.97$, $p < .01$, and this effect significantly interacted with the cue predictability effect, $F(3, 45) = 22.90$, $p < .01$. To illustrate the interaction, the attention effect (invalid RT – valid RT) is plotted in Figure 2 as filled squares connected by a solid line.

As in Experiment 1, the linear component of the attention effect was significant, $F(1, 15) = 50.77$, $p < .01$. However, unlike Experiment 1, there was also a small but significant quadratic component, $F(1, 15) = 8.09$, $p < .05$. Nevertheless, the linear component accounted for the major share of the attention effect variance (90.0%) while the quadratic and cubic trends accounted for less of the variance (8.0% and 2.0%, respectively). (The *MSE* of the cue predictability by the attention effect interaction = 388.2.)

With respect to the four hypotheses about how cue predictability and the attention effect might relate, the results provide strong support for the information-gain hypothesis. As we just discussed, we essentially found the same simple linear relation between *information gain* and attention. The influence of cue predictability on the attention effect was even more inconsistent with the marginal-gain hypothesis than Experiment 1. Going from 16.7% to 53.5% valid (a change of 36.8%) caused an 11 ms increase in the attention effect. Going from 72.2% to 83.3% (a change of 11.1%) caused a 34 ms increase in the attention effect. Again, there is a cueing effect at all levels of predictability, contradicting the information-threshold hypothesis.

As in Experiment 1, there was a small attention effect at 0 bits of information (11.22 ms; see Figure 2) because of a small carryover effect. When participants had the 0 bits condition in the first or second session, the attention effect was 3.7 ms. When participants had the 0 bits condition in the third or fourth session, the attention effect was 16.6 ms.

The results were more variable than Experiment 1, which is reasonable given that participants had to remember which color they were supposed to attend to as well as the cue predictability. As weak probabilistic associations are much harder to learn and remember than strong probabilistic associations, the nonlinearity could be because of learning and remembering which end of the barbell was the cue. Some participants may have learned the association better than others. In any case, the effect was not because of any pre-experiment association of the cue (for either the dumbbell or arrow cues), and the results are well fit by a linear function.

Experiment 3

In Experiments 1 and 2, the relation between the attention effect and the information provided by the cue in bits was well fit by a linear function. The purpose of Experiment 3 was to begin to explore the generality of this relation and explore whether there are conditions for which the linear relation might break down.

With involuntary attention, using an uninformative peripheral cue, Prinzmetal and his colleagues (Prinzmetal, Ha, & Khani, 2010; Prinzmetal, Taylor, Myers, & Nguyen-Espino, 2011) found two substantial differences in performance between experiments with distractors in nontarget locations (like the letter "E" in Figure 1) and those without distractors (empty placeholder squares). They compared displays with either two or six possible target locations. As the cue was uninformative, the cue was valid on 50% of the trials when there were two locations and 16.7% of the trials when there were six locations. With distractors, the attention effect was larger with six locations than with two locations (also see Eriksen & Yeh, 1985). Without distractors, the attention effect was the opposite; it was larger when there were two locations (also see Mordkoff et al., 2008).

As discussed above, the function relating cue information to the attention effect might be different with and without distractors. With distractors (Experiments 1 and 2), participants must find the target and an informative cue generally should be useful. However, without distractors, the target might pop-out, making the cue less useful unless it is highly predictive of the target location. Thus without distractors, the function relating cue information to the attention effect would be positively accelerating when compared with the function with distractors.

Finally, with involuntary attention, Prinzmetal et al. (2011) found that divergent mechanisms produced the attention effect with and without distractors by analyzing the entire RT distributions using a simple RT model, the Linear Ballistic Model (LBA). Later, in the Model section, we examine the mechanisms underlying the attention effect in our experiments with the LBA model.

Method

Stimuli and procedure. Experiment 3 was identical to Experiment 1 except there were no distractors in the display (i.e., no letter Es). Block order was counterbalanced identical to Experiment 1. All four cue-predictability sessions were held in a single day, and the experiment took approximately 2 hr.

Participants. There were eight participants, drawn from the same subject pool as in Experiment 1, who all received course credit for participating.

Results and Discussion

We removed incorrect trials (3.3%), trials with eye movements (1%), and trials with RTs exceeding 1.5 s (0.1%) from the analysis. Correct RTs for trials varying whether the cue was valid or invalid are shown in Table 2. As expected, participants were significantly faster on valid than invalid trials, $F(1, 7) = 33.73$, $p < .01$, and this effect significantly interacted with cue predictability, $F(3, 21) = 7.18$, $p < .01$. To illustrate the interaction, the attention effect (invalid RT − valid RT) is plotted in Figure 2 as open triangles connected by a dashed line.

Finally, it is possible to compare directly the goodness of linear fit between the attention effects scaled with bits (information-gain hypothesis) compared to percent valid (probability-matching hypothesis). For all three experiments, we found that a linear trend analysis with the independent variable bits accounted for a substantial proportion of effect variance. We used standard linear coefficients for those analyses (coefficients: −3, −1, 1, and 3). It is possible to create linear coefficients that reflect the spacing of percent valid as the independent variable (coefficients: −39.725, −2.925, 15.775, and 26.875). Across Experiments 1 to 3, the linear trend, scaled as bits, accounted for a higher proportion of variance than when scaled as percent valid: Experiment 1, 99.7% versus 94.3%; Experiment 2, 90.0% versus 72.3%; Experiment 3, 98.3% versus 89.1%. In each case, the linear trend scaled with bits accounted for more variance than the trend scaled with percent correct, providing support for the information-gain hypothesis as a better description of human performance than the probability-matching hypothesis.

As Prinzmetal et al. (2010, 2011) found different attention effects depending on the absence or presence of distractors, we thought that without distractors, the attention effect in this experiment might take a different form than it did Experiments 1 and 2. However, we found the same linear trend in this experiment, $F(1, 7) = 14.82$, $p < .01$. The linear trend accounted for 98.3% of the attention effect variance. The quadratic and cubic trends accounted for negligible components of the variance (both Fs < 1.0; the *MSE* of the cue predictability by the attention effect interaction = 200.7). Once again, we found the same simple linear relation providing strong support for the information-gain hypothesis and evidence against alternative hypotheses.

As in the previous experiments, there was a small attention effect at 0 bits, which was because of a carryover effect. The participants who had the 0 bit condition as the first or second session had an attention effect of 6 ms. The participants who had the 0 bit condition as the third or fourth session had an attention effect of 17 ms.

Overall, the attention effect without distractors (Experiment 3) was about one half as strong as the attention effect with distractors (Experiment 1). Nevertheless, the pattern of means was similar (see Figure 2). We found this result surprising. Although the differences in the attention effect because of the absence or presence of distractors found by Prinzmetal et al. (2011) were not reflected in the mean RTs in Experiment 3, there may be some signature of different mechanisms engaged when examining the full RT distribution. Thus, in the next section, we analyze the RT distributions of Experiments 1–3 using the LBA in a similar manner to Prinzmetal et al. (2011).

Modeling Cognitive Mechanisms Using the Linear Ballistic Accumulator

Analyzing the results in terms of mean RT revealed that there is a linear relationship between the amount of information provided by a cue and the attention effect, which did not depend on cue type (arrow or barbell) or on the presence of distractors in the displays (whether it contains irrelevant information). This latter result was surprising in light of previous work, which found that distractors influence involuntary attention (Prinzmetal et al., 2010, 2011). There were two main reasons that Prinzmetal et al. (2010, 2011) concluded that distractors influence involuntary attention. First, without distractors, as the number of possible target locations increased (and hence, the amount of information diminished), the cueing effect decreased (also see Mordkoff et al., 2008). With distractors, as the number of target locations increased, the cueing effects increased. In the initial analysis of the present data, there was no *qualitative* difference with and without distractors in the attention effect with and with distractors (see Figure 2). However, the attention effect was larger with distractors than without distractors.

Second, Prinzmetal et al. (2011) analyzed the full RT distributions of their results with the LBA (described below). Regardless of the presence or absence of distractors, the amount of evidence participants needed to respond (the *threshold*) was lower for valid than for invalid trials. However, the rate of information accrual was faster for valid trials only when distractors were in the display. Thus, the processes that lead to the involuntary attention effect are partly different with and without distractors. Although we did not find an effect of the presence or absence of distractors on the mean attention effect, perhaps examining the data through the lens of a RT distribution model could reveal effects of distractors on attention.

To investigate the processes underlying how the information provided by a cue influences voluntary attention, we applied the LBA to Experiments. The LBA model provides a method for analyzing the entire RT distribution, not just the means (Brown & Heathcote, 2008). According to the LBA, information accumulates until a threshold is reached, at which time a response is made. Given that our data contain a negligible error rate, we modeled the RTs of correct responses only. On a given trial, the accumulator races to its associated threshold in a linear, deterministic manner. The rate of processing and the threshold are the critical parameters that were assessed for valid trials and for invalid trials (in each cue predictability condition separately). The modeling, hence, focuses on the comparison of these parameters across cue conditions (valid vs. invalid cue conditions) in the separate predictability conditions. On each trial, the rate is drawn from a normal distribution (with the same *SD* for valid and invalid trials) and threshold is drawn from a uniform distribution.

The LBA model appealed to us because (a) it provides an explicit method for analyzing the RT distributions; (b) it is theoretically motivated; and (c) it is relatively simple to implement. Nevertheless, there are alternatives to the LBA model (e.g., serial models) and these are discussed later. Following Prinzmetal et al. (2011; also see Ludwig et al., 2009), we parameterized the rate and threshold in the following way. There are two accumulators racing in a trial, one for the correct response and one for the incorrect response. Disregarding the cue condition, accumulators for valid and invalid trials have the same parameters, with a rate of v and threshold of 1. Cueing a location has the effect of multiplying the associated accumulators' rate by α and threshold by β. Therefore, when the cue is valid, the accumulator of the correct response has a rate of αv and threshold of β, and when the cue is invalid, the accumulator has a rate of v and threshold of 1. Thus, $\alpha > 1.0$ indicates the rate for valid is higher than invalid. A $\beta < 1.0$ indicates the threshold for valid is less than invalid. Either $\alpha > 1.0$ or $\beta < 1.0$ (or both) would cause valid RTs to be faster than invalid RTs.

By parameterizing rate and threshold in terms of multipliers α and β we can easily compare rates and thresholds across experiments in an intuitive way. To the extent that attention increases the rate of processing on valid trials, α should be greater than 1.0. To the extent that attention decreases the threshold for valid trials, β should be less than 1.0.

In addition to rate and threshold related parameters, the LBA has three other parameters: the *SD* of the rate of an accumulator for a given probability condition, σ, the boundary for the starting point, A, and the "nondecision" time, T_0. The first of the additional parameters, σ is the *SD* of the normal distribution of rates that generates the rate of an accumulator on a given probability condition. The second of the additional parameters, A, provides the upper bound for the random amount of evidence that starts in the accumulator—the initial evidence is distributed Uniform (0, A). The third of the additional parameters, T_0 gives the nondecision time, which is a fixed amount of time before evidence starts accumulating in the accumulator. All three of these additional parameters were fixed to be equal for valid and invalid trials, although they varied between conditions and participants.

The LBA generates the entire distribution of RTs, expressed as cumulative distribution function (CDF) of RTs (e.g., Figure 3A–C). This distribution is the cumulative RT distributions (separate for valid and invalid RTs), which is the probability that a participant responds by time t. One advantage of the LBA, and other RT models, is that it uses the entire distribution of RT, not just the means. RTs for trials in the with and without distractors experiments might show the same trend in means (as they did in Experiments 1 and 3) and yet an examination of the entire distribution might reveal differences that can be tied to differences in processing.

In Figure 3, we provide examples of how different parameters affect the CDF. Figure 3A was generated with the assumption that the rates for valid and invalid trials are the same but the threshold for valid is less than invalid. The CDFs differ and this difference is constant for most of the distribution (it has the effect of shifting the distribution horizontally). Figure 3B was generated with the assumption that the rate was faster for valid than invalid trials, but they had the same threshold. The difference between valid and invalid is greater for longer RTs. The difference between CDFs will not be as clear at fast RTs, but will grow with longer RTs. Finally, Figure 3C is an example where both rate and threshold are different for valid and invalid.

Using the accumulator specification discussed above, we fit the parameters of the accumulators per subject, per cue predictability condition (so, each subject for each cue predictability condition had five free parameters) following the broad guidelines outlined in Donkin, Brown, and Heathcote (2011). To fit the model parameters to the RT distributions of the participants, we minimized the G^2 statistic (that χ^2 approximates, but G^2 is a more robust estimate; Cressie & Read, 1989) for the quantile maximum probability estimates (Heathcote & Brown, 2004; Heathcote, Brown, & Mewhort, 2002) of [0.1, 0.23, 0.37, 0.5, 0.63, 0.77, 0.9] of the participants' and the LBA's CDFs. We optimized this quantity using a bounded form of the SIMPLEX algorithm (Nelder & Mead, 1965).[2] Because of the small number of incorrect responses (some participants did not make any errors in at least one cue predictability condition), the RT distributions of individual subjects for incorrect responses were noisy. Thus, we only modeled correct responses. To ensure the global optimality of the reported fits, we report the best parameter values after running the optimization 15 times, each time from a randomly generated starting point.

Table 3 reports the parameter fits for each cue predictability condition of each experiment averaged over participants. The most important parameters, α and β, are in Columns 1 and 2. These are the only parameters with separate fits for valid and invalid trials. α and β encode the change in rate and threshold because of attending to the cue, respectively. Figure 4 plots only the mean α and β for the most predictive cue condition of each experiment. Whether or not there are distractors (all experiments), attending to an informative cue lowers the threshold for making the correct response. Alternatively, attending to an informative cue increases the rate that evidence is accumulated for making the correct response only when distractors are present in the scene (Experiments 1 and 2).

The results are consistent with the involuntary attention results of Prinzmetal et al. (2011). The differences between performance with or without distractors might stem from a difference in task difficulty rather than different mechanisms being recruited in the presence of distractors. For example, it is plausible that beyond a certain level of difficulty, different mechanisms and more resources are engaged. According to this explanation, increasing the task difficulty in any manner should correspond to a rate change. In Experiment 4, we dissociate these two possibilities by increasing task difficulty (by decreasing the size of the letters), without including distractors in the scene.

Experiment 4

Experiment 4 examined how task difficulty affects attentional processing with partially informative cues by manipulating the font size of the target. We wanted to know whether the rate change we found in the experiments with distractors (Experiments 1 and 2) was because of increased difficulty compared with the experiment without distractors (Experiment 3) or whether it was because of some other change in processing.

Method

Stimuli and procedure. To vary difficulty, we varied the font size. In the easy condition, the stimuli were identical to Experiment 3 (no distractors, arrow cue); the font was Helvetica 36 point. In the difficult condition, the font size was reduced to Helvetica 18 point. The difficulty was changed every block, with one half of the participants beginning with the difficult condition.

We used only the highest cue predictability condition (1.55 bits) and the no distractor condition (like Experiment 3). For practice, each participant had two blocks of 32 trials, one block of the easy condition and one block of the difficult condition. The reported data was then collected for 10 blocks of 81 trials each.

[2] We used fminsearchbnd, which is available at http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd.

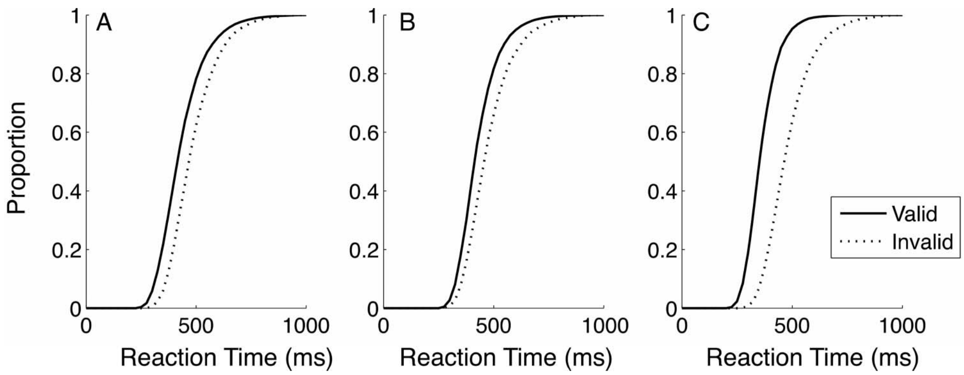

Figure 3. Simulated cumulative distribution functions: (A) valid and invalid differ in threshold, not rate; (B) valid and invalid differ in rate, not threshold; and (C) valid and invalid differ in both rate and threshold.

Participants. Twelve undergraduates from University of California, Berkeley participated in the experiment in exchange for course credit.

Results and Discussion

As before, we removed incorrect trials (3.3%), trials with eye movements (1.1%), and trials with RTs exceeding 1.5 s (0.07%) before analysis. Participants were significantly faster on valid trials than invalid trials, $F(1, 11) = 39.48$, $p < .01$. The mean RTs for trials where the cue was valid and invalid were 330 and 369 ms, respectively. Furthermore, participants were significantly faster on the easy (large font) then the difficult condition (small font; $F(1, 11) = 21.87$, $p < .01$). The mean RTs for the trials with large and small font were 339 and 360 ms, respectively. There was no interaction between task difficulty and cue predictability, $F(1, 11) = 0.028$. The mean RT for the large font conditions for valid and invalid trials was 320 and 358 ms, respectively (*SD*s, averaged across participants, were 83 ms and 95 ms, respectively). For the small font condition, the mean RT for valid and invalid trials was 341 and 379 ms, respectively (*SD*s were 85 ms and 88 ms, respectively). Thus the attention effect was approximately the same for the easy and difficult conditions (39 and 38 ms, respectively).

The mean parameter fits are in Table 3, and the rate and threshold parameters are plotted in Figure 4. For both the easy and difficult conditions, attending to a partially valid cue lowered the threshold for the correct response, but did not affect the rate.

In Experiment 4, increasing difficulty by reducing font size did not induce a difference in rate between valid an invalid trials. However, in the previous experiments, increasing difficulty by adding distractors did induce a difference in rate. Hence not all methods of manipulating task difficulty have the same effect on rate. Adding distractors to the display makes it difficult to find the target. Using a small font size makes it difficult to identify the target. However, in our experiments, only adding distractors had the effect of modulating the rate parameter. Thus, the presence of distractors, not task difficulty, may account for the observed increased rate in the previous experiment.

Table 3
Estimated Parameters

	α rate multiplier	β threshold multiplier	σ rate *SD* ($\times 10^{-4}$)	A starting point boundary	T_0 nondecision time	*V* rate ($\times 10^{-3}$)
Experiment 1 condition						
0	1.11	1.03	7.73	0.46	130.67	3.29
0.51	1.07	0.95	6.22	0.49	111.64	2.78
1.02	1.13	0.94	5.98	0.47	103.39	2.62
1.55	1.20	0.98	7.46	0.58	141.11	2.67
Experiment 2 condition						
0	1.01	0.98	8.31	0.44	121.77	3.12
0.51	1.01	.94	9.31	0.50	149.56	3.09
1.02	1.03	0.93	6.78	0.44	124.79	2.77
1.55	1.07	0.86	8.22	0.47	133.81	2.74
Experiment 3 condition						
0	0.89	0.90	16.57	0.78	185.67	4.91
0.51	.856	0.83	8.59	0.56	129.21	3.55
1.02	0.88	0.82	9.71	0.62	152.52	3.68
1.55	0.88	0.76	10.40	0.66	173.64	3.68
Experiment 4 condition						
1.55 easy	1.02	.91	4.73	.46	50	2.60
1.55 difficult	.96	0.83	3.98	.44	50	2.43

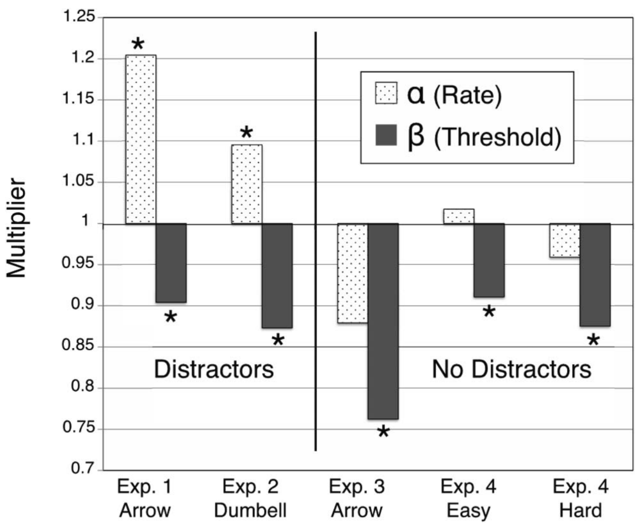

Figure 4. Rate (α) and threshold (β) for most predictive cue condition (1.55 bits). Higher rate for valid than invalid trials is reflected in $\alpha > 1.0$. Lower threshold for valid than invalid trials is reflected by $\beta < 1.0$. An asterisk (*) indicates significantly different from 1.0 ($p < .05$) by a paired *t* test (two-tailed).

There are still other possible explanations for why our manipulation of task difficulty with font size did not affect the rate of evidence accrual for the cued target. It is possible that the difficulty manipulation was not strong enough to be reflected in a difference in rate. The difference in size of the attention effect depending on difficulty in Experiment 4 (large font vs. small font) was 28 ms. The difference in RT for the most predictive cue condition between Experiment 1 and Experiment 3 was 21 ms. Thus, our difficulty manipulation in Experiment 4 was in the same range as the effect of distractors. Although this provides some evidence against the possibility that our difficulty manipulation was just not large enough, it clearly is not definitive. We leave further exploration of this question for future research.

Nevertheless, in the present experiments, attending to a partially valid cue only increases the rate of accumulation for the correct response when there are distractors in the scene. In other words, the presence of distractors engages different mechanisms that affect how attention is allocated to a partially valid cue.

General Discussion

We examined how voluntary attention, measured by RT, varies as a function of cue predictability. We also investigated the cognitive processes whereby attention affects performance by analyzing full RT distributions with the LBA model. We found that the effect of attention increased linearly with the amount of information provided by the cue quantified in bits of information. In all of the experiments, attention affected the threshold for responding. Additionally, attention affected the rate of responding only when there were distractors in the display. This effect was only uncovered by analyzing the full RT distribution.

The linear increase in the attention effect with information gain might be considered surprising for several reasons. Although memory search for visual displays increases linearly with the amount of information held in memory (in bits; Wolfe, 2012), visual search does not increase linearly with the amount of information in the display. Rather, visual search is notoriously linear with the number of items in the display (Wolfe, 1998). Of interest to the authors, visual attention manipulated with cue predictability follows a different rule than visual search (cf. Palmer, Ames, & Lindsey, 1993). The relation between visual search varying set size and voluntary attention is discussed further below.

Our results are also inconsistent with alternative hypotheses for how cue predictability and the attention effect might relate. In addition to the previously discussed marginal-gain and information-threshold hypotheses, our results are inconsistent with a simple *probability-matching* strategy (e.g., Vulkan, 2000) whereby participants allocate attention according to the percent of valid trials. For example, one might suppose if there are 80% valid trials, participants might allocate 80% of their attention to the cued location and divide the remaining 20% to the invalid locations. If this were the case, the attention effect would be linear in terms of the percent of valid trials, not the amount of information (in bits). In this respect, while our findings are not consistent with visual search findings, which do follow a linear trend as a function of target location probability, our findings are consistent with many other domains where information scales behavior rather than probabilities (e.g., memory search, decision processes with multiple alternatives and motor movement).

Beyond constraining theories of attentional deployment, quantifying the relation between the amount of information in a cue and the effect of attention is important for several reasons. First, at a pragmatic level, knowing the relation between cue predictability and attention is useful when designing experiments. For example, if one believed that the marginal-gain hypothesis was true, one would have erroneously believed that attention would have diminishing returns, so for example, there would not be much of a difference between a 70% and 80% percent valid cue (given six target locations). In cases with a small cueing effect, it might be well worth increasing cue predictability.

Second, knowing the *general* relation between the information provided by the cue and the attention effect makes it possible to vary set size in a meaningful way. Display-size (i.e., number of possible target locations) has been one of the main independent variables in the search literature and may prove important in spatial orienting of attention as well. For example, Prinzmetal and colleagues (Prinzmetal et al., 2010, 2011) found that with involuntary attention (random peripheral cue) when distractors were absent, the attention effect was larger when there were two possible locations rather than six, but when distractors were present, the attention effect was larger with six locations than two. It is easy to control information with an uninformative cue (0 bits of information). For example, the probability that the target was in the cued location is one-half when there were two possible locations for the target, but only one-sixth, when there were six possible locations.

Although varying the display-size under the paradigm of involuntary attention is straightforward, it is not straightforward for voluntary attention. Consider an experiment by Geng and Behrmann (2005). They varied display-size (either four or eight locations), while controlling for crowding and other sensory factors. Attention was manipulated with location probability (instead of cue predictability). There was a 75% probability that a target would appear in a particular location, and a 25% probability that the target would appear in one of the other locations. The high probability location was varied between blocks. RT from these blocks were compared with blocks where the target location was determined randomly (equal probability for each location). Obviously, location probability provides more information with eight locations than four locations (with, e.g., the target appearing 75% of the time in one location). That is, knowledge of the location probabilities reduces uncertainty more with eight than four locations. Equations 1 and 2 can be applied to the stimulus conditions used by Geng and Behrmann (2005). When compared with the target having an equal probability of occurring in any location, a condition where the target occurred in a given position on 75% of the trials provides 0.79 bits of information with a display size of four and 1.49 bits of information for a display size of eight. The location probability effect in RT was about twice as large with display size of eight than display size of four. Unfortunately, there were only two different display sizes and two location probability conditions (75% high to 25% other; vs. random). By knowing the relationship between cue predictability and cue information, we can vary the display size, while keeping the information provided by the cue constant. Because display-size is an important variable in the visual search literature, cueing research that varies display-size with voluntary attention may help unify these important, yet different, visual attention bodies of work (for a discussion of models for accuracy data, as opposed to RTs see Kinchla et al., 1995, and Shimozaki et al., 2003).

Probabilities are a parameter manipulated also outside spatial attention. For example, Moore and Egeth (1998) examined feature-based attention. They had participants search for a digit among letters. The display characters were presented in different colors. They varied the probability that the target was a particular color. Search RT varied with the probability that the target was a particular color. Unfortunately, there were not enough data points to determine if feature probability affects search in the same manner as cue predictability. However, if this feature probability effect follows the same linear relation as cue predictability, it may be a way to parameterize one aspect of guided search (e.g., Wolfe, 2007).

Future research will determine the generality of the linear relation between attention and cue predictability, as in the case of voluntary attention. There may be cases where the effect of attention is not linearly related to cue predictability. Although the relation held in all of our experiments, there are many factors that were constant across our experiments that could influence the relation. For example, participants were nearly 100% correct and only made a few errors. The relation between information and RT might be more complex when error rates are high and such effect might well be mediated by the presence of distractors. If we included errors in the experimental design, we could more accurately characterize the RT distribution when people make errors using the accumulator model. However, our primary interest was in understanding the relation between information and RT performance, and so, we used a design where errors were rare and used RT as our dependent variable. It could be the case that other measures related to attentional mechanisms would not be linear in information. Further, it is possible that the actual "attention mechanisms" are not linearly related to RT, but rather the linear relation reflects some other process (or interaction between processes). Regardless of the various possible underlying mechanisms it is remarkable that the relation was consistent and robust across the reported experiments and distractor conditions.

We used the LBA model to investigate the processes that might underlie the measured attention effects. The model allows us to analyze the entire RT distribution, not just the central tendency (i.e., mean). We allowed two parameters to vary between valid and invalid trials: the rate of information accrual and the threshold. In each experiment, the threshold was lower for valid than invalid trials. However, the rate of information accrual was only different when there were distractors in the display.

The results with respect to rate and threshold are identical to those we found previously for involuntary attention with a peripheral nonpredictive cue (Prinzmetal, Taylor, Myers, & Nguyen-Espino, 2011). In each experiment, we found a threshold difference between valid and invalid trials, but only a difference in rate when there were distractors in the display. To test the generality of this finding, we reanalyzed the RT distributions of a similar experiment by Prinzmetal, Ha, and Khani (2010, Experiment 3) using the same methodology used in this article. We examined display Size 2 without distractors and display Size 6 with distractors because these were the conditions that showed the largest cueing effects. In all conditions, the cue affected the threshold but only affected rate when there were distractors in the display. In Prinzmetal et al. (2010, 2011) and the present work, the presence or absence of distractors seems to change attentional processing in the same qualitative manner. When the distractors are present, attention affects the rate of processing and the threshold. Without distractors, attention only affects the threshold. This generality seems to hold for involuntary attention and voluntary attention. Previous work has shown that performance with or without distractors can completely alter the type of selection modulation because of attention (e.g., Yiğit-Elliott, Palmer & Moore, 2011). In our setup, given that there are always stimuli that appear in unattended locations, all uncued locations are in a position to introduce perceptual interference to performance on the target location. Hence in the presence of perceptual interference, we demonstrate that attention can alleviate interference effects brought about by distractors.

The LBA model revealed differences in processing for experiments with and without distractors. In a complementary analysis we also explored whether the LBA model fits could also produce the linear increase in the attention effect with bits of information in Experiments 1 to 3 (see Figure 2). To test this, we began with the predicted CDF for each experiment, subject, probability condition, and cue-type. We quantized the predicted CDF in 1 ms bins, and then calculated the mean for each fitted CDF. From the means of the fits, we calculated an attention effect. These values, derived from the model fit, were then subjected to the same ANOVA trends analysis as in Experiments 1 to 3 and replicated the findings.

There was a significant linear trend in the model fit means for Experiments 1 to 3. The linear trends accounted for 96.8%, 86.2%, and 98.9% of the between condition variance for Experiments 1 to 3, respectively. Therefore, the model generates attention effects that are similar to the empirically obtained means, above. The LBA fits the entire RT distribution and in that sense provides a fuller description of the data. The means derived from the fits demonstrate that the model also captures the linear increase of the attention effect with information.

The LBA model is a parallel model; however, it is plausible that serial models can account for our results. It is somewhat surprising that serial models are not developed as fully as parallel models in terms of fitting the entire RT distribution (e.g., Palmer, Horowitz, Torralba, & Wolfe, 2011), but we suspect that new serial models can be developed to account for the RT distributions found in the literature. Prinzmetal et al. (2011) pointed out that rate differences in a parallel model could be mimicked by a serial model. There are also hybrid serial-parallel models that could be considered (e.g., Harris, Shaw, & Bates, 1979). One hybrid approach is that items would be grouped, and the search within a group would be in parallel, but between groups, serial (Treisman, 1982). Perhaps an array of, for example, 8 items is parsed into 2 groups, and these groups are attended serially but the items within each group are processed in parallel. Then each of the groups is parsed into two groups of two items, and so forth. This idea is similar to the Zoom Lens model proposed by Eriksen & Yeh (1985). Such approaches may explain our attention effect patterns, but there are certainly other possibilities. Further, our study proposes a framework for thinking about the relationship between information and attention, as it is measured in behavior (RTs). Other studies would have to investigate the linking assumptions between these measures and physiological models for spatial deployment of attention.

Regardless of the "real" model that accounts for our results, the LBA provides a convenient method for analyzing the RT distributions of the participants in our and others experiments. The serial-parallel-hybrid issue is a thorny issue and one that has been difficult to solve. In fact, it may be impossible to solve without particular assumptions about the representations used by the visual system (Anderson, 1978). What we believe is more critical is whether alternative models concur with a fundamental difference in locating the target (with distractors similar to the target) and identifying the target.

The distinction between locating the target and identifying the target may clarify other results in the literature. For example, in several experiments designed to use accuracy as the primary dependent variable,[3] Prinzmetal, McCool, and Park (2005) found that involuntary attention did not affect accuracy (also see, e.g., Prinzmetal, Zvinyatskovskiy, Gutierrez, & Dilem, 2009; Stevens, West, Al-Aidroos, Weger, & Pratt, 2008). Those experiments were scrupulously designed so that there was no location uncertainty, and hence, finding the target contributed little to performance. We predict that if these experiments were run as RT experiments (see Footnote 3) the distribution of RTs would resemble Figure 3A (little rate differences between valid and invalid trials). In experiments that found an effect of involuntary attention on accuracy, the results may have been because of an increased difficulty of locating the target (e.g., Gould, Wolfgang, & Smith, 2007). The target may be rendered difficult to find by adding distractors, masking all stimulus locations, or using low contrast targets (e.g., Gould, Wolfgang & Smith, 2007; Kerzel, Gauch, & Buetti, 2010). We would expect RT distributions in these studies to resemble Figure 3C (rate and threshold). However, as in Experiment 4, task difficulty alone may not be enough to produce a difference in rate, and there may be something special about the presence of distractors in increasing the perceptual difficulty of the display.

In summary, orienting of voluntary attention is sensitive to the amount of statistical information in the environment, as quantified by information theory. Specifically, voluntary attention increases to environmental cues with increasing information in a linear manner. The processing underlying voluntary attention when distractors are in the display may be fundamentally different than when no distractors are in the display.

[3] By pure RT experiment, we mean that participants were nearly 100% correct, as in the present experiment. By accuracy experiment, we mean that participants were under no speed pressure, but the task was more difficult so that even without speed pressure, participants made substantial errors and accuracy was the dependent variable.

References

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review, 98,* 249–277. http://dx.doi.org/10.1037/0033-295X.85.4.249

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178. http://dx.doi.org/10.1016/j.cogpsych.2007.12.002

Close, A., d'Avossa, G., Sapir, A., & Parkinson, J. (2010). Effects of central cue reliabilities on discrimination accuracy and detection speed. *Journal of Vision, 10,* 153. (Presented at the Vision Science Society annual meeting, May, 2010).

Cover, T., & Thomas, J. (1991). *Elements of information theory*. New York, NY: Wiley. http://dx.doi.org/10.1002/0471200611

Cressie, N., & Read, T. R. C. (1989). Pearson's ×2 and the Loglikelihood Ratio Statistic G2: A comparative review. *International Statistical Review, 57,* 19–43. http://dx.doi.org/10.2307/1403582

Donkin, C., Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology, 55,* 140–151. http://dx.doi.org/10.1016/j.jmp.2010.10.001

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 583–597. http://dx.doi.org/10.1037/0096-1523.11.5.583

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Oxford, England: Brooks/Cole.

Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution*. New York, NY: Basic Books.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics, 67,* 1252–1268. http://dx.doi.org/10.3758/BF03193557

Gould, I. C., Wolfgang, B. J., & Smith, P. L. (2007). Spatial uncertainty explains exogenous and endogenous attentional cuing effects in visual signal detection. *Journal of Vision, 7,* 1–17.

Harris, J. R., Shaw, M. L., & Bates, M. (1979). Visual search in multicharacter arrays with and without gaps. *Perception & Psychophysics, 26,* 69–84. http://dx.doi.org/10.3758/BF03199862

Heathcote, A., & Brown, S. D. (2004). Reply to Speckman and Rouder: A theoretical basis for QML. *Psychonomic Bulletin & Review, 11,* 577–578. http://dx.doi.org/10.3758/BF03196614

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394–401. http://dx.doi.org/10.3758/BF03196299

Hick, W. E. (1952). On the rate of gain of information. *The Quarterly Journal of Experimental Psychology, 4,* 11–26. http://dx.doi.org/10.1080/17470215208416600

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45,* 188–196. http://dx.doi.org/10.1037/h0056940

Kerzel, D., Gauch, A., & Buetti, S. (2010). Involuntary attention with uncertainty: Peripheral cues improve perception of masked letters, but may impair perception of low-contrast letters. *Journal of Vision, 10,* 12. http://dx.doi.org/10.1167/10.12.12

Kinchla, R. A., Chen, Z., & Evert, D. (1995). Precue effects in visual search: Data or resource limited? *Perception & Psychophysics, 57,* 441–450. http://dx.doi.org/10.3758/BF03213070

Lambert, A., Roser, M., Wells, I., & Heffer, C. (2006). The spatial correspondence hypothesis and orienting in response to central and peripheral spatial cues. *Visual Cognition, 13,* 65–88. http://dx.doi.org/10.1080/13506280500264460

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology, 22,* 1000–1004. http://dx.doi.org/10.1016/j.cub.2012.03.054

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review, 104,* 524–553. http://dx.doi.org/10.1037/0033-295X.104.3.524

Ludwig, C. J. H., Farrell, S., Ellis, L. A., & Gilchrist, I. D. (2009). The mechanism underlying inhibition of saccadic return. *Cognitive Psychology, 59,* 180–202. http://dx.doi.org/10.1016/j.cogpsych.2009.04.002

McKenzie, C. R. M. (2004). Hypothesis testing and evaluation. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 200–219). Oxford, England: Blackwell. http://dx.doi.org/10.1002/9780470752937.ch10

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97. http://dx.doi.org/10.1037/h0043158

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences, 7,* 141–144. http://dx.doi.org/10.1016/S1364-6613(03)00029-9

Moore, C. M., & Egeth, H. (1998). How does feature-based attention affect visual processing? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1296–1310. http://dx.doi.org/10.1037/0096-1523.24.4.1296

Mordkoff, J. T., Halterman, R., & Chen, P. (2008). Why does the effect of short-SOA exogenous cuing on simple RT depend on the number of display locations? *Psychonomic Bulletin and Review, 15,* 819–824.

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4,* 61–64.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature, 434,* 387–391. http://dx.doi.org/10.1038/nature03390

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313. http://dx.doi.org/10.1093/comjnl/7.4.308

Nelson, J. D. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review, 112,* 979–999.

Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience matters: Information acquisition optimizes probability gain. *Psychological Science, 21,* 960–969. http://dx.doi.org/10.1177/0956797610372637

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101,* 608–631. http://dx.doi.org/10.1037/0033-295X.101.4.608

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198524496.001.0001

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology: Human Perception and Performance, 37,* 58–71. http://dx.doi.org/10.1037/a0020747

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 108–130. http://dx.doi.org/10.1037/0096-1523.19.1.108

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32,* 3–25. http://dx.doi.org/10.1080/00335558008248231

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160–174. http://dx.doi.org/10.1037/0096-3445.109.2.160

Prinzmetal, W., Ha, R., & Khani, A. (2010). The mechanisms of involuntary attention. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 255–267. http://dx.doi.org/10.1037/a0017600

Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General, 134,* 73–92. http://dx.doi.org/10.1037/0096-3445.134.1.73

Prinzmetal, W., Taylor, J. A., Myers, L. B., & Nguyen-Espino, J. (2011). Contingent capture and inhibition of return: A comparison of mechanisms. *Experimental Brain Research, 214,* 47–60. http://dx.doi.org/10.1007/s00221-011-2805-x

Prinzmetal, W., Zvinyatskovskiy, A., Gutierrez, P., & Dilem, L. (2009). Voluntary and involuntary attention have different consequences: The effect of perceptual difficult. *Quarterly Journal of Experimental Psychology, 62,* 352–369.

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision, 7,* 6. http://dx.doi.org/10.1167/7.3.6

Ristic, J., & Kingstone, A. (2006). Attention to arrows: Pointing to a new direction. *Quarterly Journal of Experimental Psychology, 59,* 1921–1930.

Shannon, C. E. (1948). The mathematical theory of communication. *The Bell System Technical Journal, 27,* 379–423, 623–656. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shaw, M., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 201–211.

Shimozaki, S. S., Eckstein, M. P., & Abbey, C. K. (2003). Comparison of two weighted integration models for the cueing task: Linear and likelihood. *Journal of Vision, 3,* 3. http://dx.doi.org/10.1167/3.3.3

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654. http://dx.doi.org/10.1126/science.153.3736.652

Stevens, S. A., West, G. L., Al-Aidroos, N., Weger, U. W., & Pratt, J. (2008). Testing whether gaze cues and arrow cues produce reflexive or volitional shifts of attention. *Psychonomic Bulletin & Review, 15,* 1148–1153.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 194–214. http://dx.doi.org/10.1037/0096-1523.8.2.194

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys, 14,* 101–118. http://dx.doi.org/10.1111/1467-6419.00106

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9,* 33–39. http://dx.doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford.

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science, 23,* 698–703. http://dx.doi.org/10.1177/0956797612443968

Yiğit-Elliott, S., Palmer, J., & Moore, C. M. (2011). Distinguishing blocking from attenuation in visual selective attention. *Psychological Science, 22,* 771–780.

Received October 28, 2013
Revision received March 4, 2015
Accepted March 20, 2015 ■