

## How humans react to changing rewards during visual foraging

Jinxia Zhang<sup>1,2</sup> · Xue Gong<sup>3</sup> · Daryl Fougnie<sup>4</sup> · Jeremy M. Wolfe<sup>2,5</sup>

Published online: 30 August 2017 © The Psychonomic Society, Inc. 2017

**Abstract** Much is known about the speed and accuracy of search in single-target search tasks, but less attention has been devoted to understanding search in multiple-target foraging tasks. These tasks raise and answer important questions about how individuals decide to terminate searches in cases in which the number of targets in each display is unknown. Even when asked to find every target, individuals quit before exhaustively searching a display. Because a failure to notice targets can have profound effects (e.g., missing a malignant tumor in an X-ray), it is important to develop strategies that could limit such errors. Here, we explored the impact of different reward patterns on these failures. In the Neutral condition, reward for finding a target was constant over time. In the Increasing condition, reward increased for each successive target in a display, penalizing early departure from a display. In the Decreasing condition, reward decreased for each successive target in a display. The experimental results demonstrate that observers will forage for longer (and find more targets) when the value of successive targets increases (and the opposite when value decreases). The data indicate that observers were

☐ Jinxia Zhang jinxiazhang@seu.edu.cn learning to utilize knowledge of the reward pattern and to forage optimally over the course of the experiment. Simulation results further revealed that human behavior could be modeled with a variant of Charnov's Marginal Value Theorem (MVT) (Charnov, 1976) that includes roles for reward and learning.

**Keywords** Human foraging · Optimal foraging · Search termination · Reward pattern · Visual search

#### Introduction

In the last few decades, a substantial body of research has explored how humans search for a single target in displays of distractor items (Desimone & Duncan, 1995; Koch & Ullman, 1985; Treisman & Gelade, 1980; Tsotsos et al., 1995; Wolfe, 1994, 2012). Much less work has been directed at tasks where observers search for an indeterminate number of targets in each of many visual displays (visual foraging tasks). In our daily lives, numerous search tasks have a foraging structure, e.g. buying some fruit at a market or looking up information on a website. Important applied tasks can also involve foraging, e.g., finding metastases in an X-ray of a cancer patient. Unlike single target search tasks, the central question is less "Did I find the target" than "When do I quit searching in the current display and move to the next one?" Recently, search termination on foraging tasks of humans began to attract the attention of researchers (Cain, Vul, Clark, & Mitroff, 2012; Fougnie, Cormiea, Zhang, Alvarez, & Wolfe, 2015; Hutchinson, Wilke, & Todd, 2008; Wolfe, 2013; Zhang, Gong, Fougnie, & Wolfe, 2015).

When foraging, humans, like animals, tend not to search exhaustively even if explicitly asked to stop searching only when all targets were found (Wolfe, 2013). The satisfaction of



Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing, Jiangsu 210096, China

Visual Attention Laboratory, Brigham and Women's Hospital, Cambridge, MA 02139, USA

<sup>&</sup>lt;sup>3</sup> Psychiatry Neuroimaging Laboratory, Brigham and Women's Hospital, Boston, MA 02215, USA

Science Division, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

<sup>&</sup>lt;sup>5</sup> Harvard Medical School, Boston, MA 02115, USA

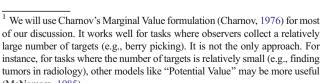
search (SOS) phenomenon, studied in radiology (Berbaum et al., 1990), can be seen as a troubling example. In SOS experiments, observers are more likely to miss a target if they have previously found a different target in the same display (Berbaum et al., 1990; Fleck, Samei, & Mitroff, 2010). Such SOS errors have been found in different kinds of medical images, e.g., abdominal radiography (Franken et al., 1994), chest radiography (Berbaum, Franken, Dorfman, Caldwell, & Krupinski, 2000; Samuel, Kundel, Nodine, & Toto, 1995) and osteoradiology (Ashman, Yu, & Wolfman, 2000). Efforts to reduce SOS have included computer-aided diagnosis (Berbaum, Caldwell, Schartz, Thompson, & Franken, 2007) and an informal checklist (Berbaum, Franken, Caldwell, & Schartz, 2006).

It can be helpful to think about SOS and related errors as the by-products of strongly engrained foraging behavior. If that is so, insights from the study of foraging might suggest new approaches to reducing those errors. Work on Optimal Foraging Theory in the animal literature provides a good theoretical foundation. Charnov's Marginal Value Theorem (MVT; Charnov, 1976) is a useful starting place for an understanding of why foragers do not collect everything from a patch in tasks like berry picking<sup>1</sup>.

The Marginal Value Theorem holds that an individual will quit searching the current patch when the current yield from searching (e.g., targets found per unit time) drops below the average rate. The rate of target acquisition slows when searching a display containing many targets because, as targets are found, the remaining items are less likely to be good quality targets and/or are harder to find. When current target acquisition becomes less valuable than average target acquisition, it is optimal to move on to the next patch. However, maximizing the rate of target acquisition is not always the ultimate goal of a foraging task. If the cost of missing a target is high, it may be preferable to search more exhaustively (i.e. avoid SOS-style errors). In such cases, it may be effective to make targets found in later periods of search more valuable than targets found first. If the value of a target increases after each target that is found, this could offset the reduced rate of target acquisition and, thus, could lead to more exhaustive

In this paper, we explored the effects of manipulating the value of targets over time in a laboratory analog of a berrypicking task as in Wolfe (2013) and Zhang et al. (2015). In this task, we generated small red squares in the display to represent "berries" and small green squares to represent "leaves" (see Fig. 1 for one example of these stimulus "patches"). Our

of our discussion. It works well for tasks where observers collect a relatively large number of targets (e.g., berry picking). It is not the only approach. For instance, for tasks where the number of targets is relatively small (e.g., finding tumors in radiology), other models like "Potential Value" may be more useful (McNamara, 1985).



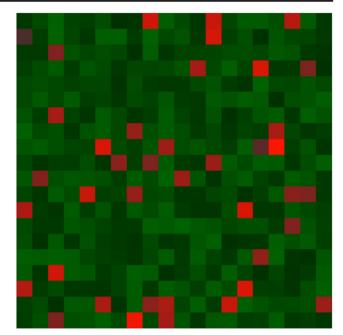


Fig. 1 Stimulus configuration: modified screenshot. Red items "Berries". Brightness is correlated with "goodness"

experiment had three reward conditions. In the Neutral condition, the value of a target was kept constant over time. This Neutral condition is similar to previous studies on foraging (Wolfe, 2013; Zhang et al., 2015) and to studies that have shown SOS (Ashman et al., 2000; Berbaum et al., 1990). In the Decreasing condition, the value of targets decreased with each selection in a patch (i.e., the Nth+1 target was worth less than the Nth target). In the Increasing condition, the value of targets increased with each selection in a patch. Each participant was tested in just one of the conditions.

Our experiment had two principle aims. First, to determine whether increasing the value of targets over time eliminates, or at least ameliorates SOS-style errors by making foraging more exhaustive; secondly, to determine if MVT continues to describe human foraging behavior when items have different / changing values.

To anticipate our results, we found evidence that humans would change their behavior in response to different patterns of reward, picking more berries when rewards increased from selection to selection, and picking fewer berries when rewards decreased. Moreover, our results suggest that observers were learning to forage more optimally over time. These results indicate that a good model to predict foraging behaviors under different patterns of reward should at least consider a role of reward and a role of learning. We simulate a foraging model based on Charnov's Marginal Value Theorem and compare its performance to human foraging behavior. The results indicate that the Marginal Value Theorem (MVT) is a useful starting place for an understanding of why foragers do not collect everything from a patch in tasks like berry picking, and how humans forage



under different patterns of reward. Note that, while our variant of a MVT model predicts human foraging behaviors well, other non-MVT models may also predict similar foraging behaviors so long as they consider the roles of reward and learning.

#### **Methods**

The experiments were conducted at the Harvard Decision Science Laboratory using Psychophysics Toolbox version 3.0.9 in MATLAB 7.10.0 (R2010a)(Brainard, 1997). The stimuli were presented on a 19-in Viewsonic NX1932w monitor with a resolution of  $1440 \times 900$  pixels. Each monitor was connected to a Dell computer running Windows 7.

#### Stimulus material

An example of the "berry bushes" used in our experiments is shown in Fig. 1. Each bush consisted of a  $20 \times 20$  grid of red or green squares. Each bush subtended a visual angle of  $15.6^{\circ}$  by  $15.6^{\circ}$  and each square subtended a visual angle of  $0.8^{\circ}$  by  $0.8^{\circ}$ . The stimuli were viewed at a 60 cm viewing distance. Observers' head positions were not fixed so the actual viewing distance was likely to vary somewhat.

Observers were instructed that the green squares were "leaves" and should be ignored when picking "berries". The color of the green squares varied from [0, 51, 0] to [0, 100, 0] in RGB color space with a uniform distribution of the values in the green color channel. Of the 400 squares in each bush, 40 were red shaded "berries". "Good" and "bad" berries had different average colors but were drawn from overlapping color distributions. "Good" berries that yielded points when picked while "bad" berries were worthless. The colors of these berries were defined in RGB color space by a triplet [R, (255-R)/4, (255-R)/4]. For good berries, redness (R) was assigned a value, drawn from a normal distribution with a mean of 190 and a standard deviation of 30. For bad berries, R was assigned a value, drawn from a normal distribution with a mean of 130 and a standard deviation of 30. Thus, the d' of the color signals between good berries and bad berries was 2.0, which means the color of a berry provided an informative but imperfect guide to its "goodness". Bush quality was fixed at 50%, i.e., 20 berries out of 40 were good. Observers were told that brighter redder berries were more likely to be good than the darker berries, but that the color information of the berries was not completely reliable. Observers were not informed about the bush quality.

### Procedure

Figure 2 shows the procedure of our experiments. Each experiment consisted of an unlimited number of berry bushes in which observers could forage for good berries. Within each

bush, observers picked a berry by moving the mouse and left-clicking on a red square in the display. Click time of each berry was self-paced. Once a berry was picked, it was replaced by a green square. Meanwhile, both visual and auditory feed-back signals were given. The feedback time was fixed as 0.6 s for each click. Visual feedback was presented onscreen near the picked berry. Bad berries earned 0 points and made a "bad" noise (a low frequency with 150 Hz) when picked. Good berries earned some reward and made a "good" noise (a higher frequency) when picked. The amount of reward changed over time to explore how different reward changing patterns would influence observers' foraging behavior. The pitch of the feedback tone changed with the changing point value (the higher the pitch, the larger the point values).

There were three conditions, each with a different pattern of rewards as shown in Fig. 3. In the *Increasing* condition, the first good berry on each new bush was worth 2 points. The value of each subsequent good berry was 2 points higher (2, 4, 6...points). In the *Decreasing* condition, the first good berry on each new bush was worth 40 points. The value of each subsequent good berry decreased by 2 points (40, 38, 36... points). In the *Neutral* condition, the points for each good berry remained constant at 20 points/berry. The frequency of the auditory feedback (F) was proportional to the points (P) earned for each good berry according to the following formula: F = 25P + 300 Hz.

The running total of points was displayed in the upper right corner of the screen (outside of the berry bush). The point total automatically updated after each click. Importantly, observers were allowed to leave current bush at any time by clicking on the "next patch" box, even if there were berries still remaining on the screen. The "next patch" box was located in the lower right corner. Participants were not penalized (or rewarded points) for any remaining good berries. When a participant quit searching a bush, a new bush would be presented after a constant 3 s delay plus recording and displaying time ("travel time", in the language of the foraging literature; Charnov, 1976).

Each participant completed three blocks of one of the three conditions. Completing a block required reaching a particular point total. Observers were asked to reach a total of 8000 points in the *Increasing* condition, 10,000 points in the *Decreasing* condition, and 8500 points in the *Neutral* condition. These different point goals were designed to roughly equate the time needed to complete each condition. Since the point goal determined the termination time of the experiment, observers were motivated to earn points as quickly as possible.

#### **Participants**

A total of 39 observers took part in our experiments, of which 13 observers (three males, mean age 25.2 years, range 19–39



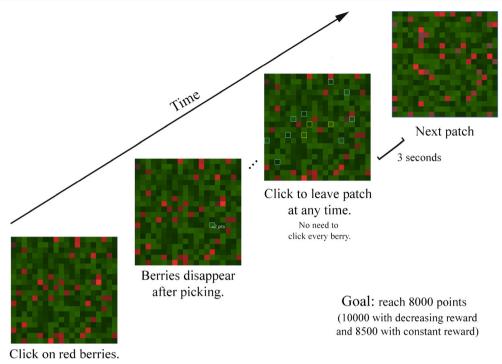


Fig. 2 The procedure of the experiment. Observers pick as many berries as they wish before moving to the next berry patch

years old) were tested in the *Increasing* condition, 13 observers (three males, mean age 30.4 years, 19–50 years old) were tested in the *Decreasing* condition, and 13 observers (six males, mean age 24.1 years, 18–42 years old) were tested in the *Neutral* condition. All observers gave informed consent that was approved by Brigham and Women's Hospital and was consistent with the Declaration of Helsinki. All observers were paid US \$10/hr for their time. All observers had vision corrected to at least 20/25 and passed the Ishihara color vision screen.

#### Results

In each of the three blocks of the *Increasing* condition, the average number of bushes that participants saw was 36.7, 33.6, and 32.7, respectively. They saw 35.1, 37.3, 38.7 bushes

for each block of the *Decreasing* condition, and 33.3, 33.3 and 34.5 bushes for each block of the *Neutral* condition. We used data from all the trials (bushes) apart from the first five practice trials and the last unfinished bush in each block (the last block was unfinished because the task ended when the point total was reached).

The average elapsed time spent in each block of the *Increasing* condition was 944.9 s, 860.8 s and 871.4 s. Elapsed times were 753.9 s, 712.5 s and 652.3 s in the *Decreasing* condition and 791.4 s, 758.8 s and 761.9 s in the *Neutral* condition. An ANOVA with Condition and Block as factors shows a main effect of Condition [F(2,36) = 9.316, P = 0.0005] because of the different reward patterns and different goals set for these conditions. There is also a main effect of Block [F(2,72) = 25.69, P < 0.0001], demonstrating that observers sped up over the blocks. The significant interaction of Condition and Block [F(4,72) = 4.006, P = 0.0055] reflects the

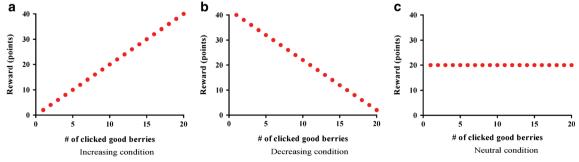


Fig. 3 The changing patterns of the reward in different conditions



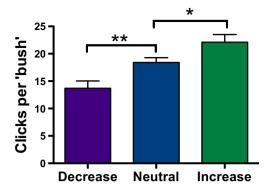
fact that the decrease in time per block is largest for the *Decreasing* condition and smallest for the *Neutral* condition.

#### Do observers react to different reward patterns?

Further analyses focused on the number of clicks in the different conditions. We removed clicks if the click time was greater than 10 s. The number of these outlier clicks was very low, with a total of 10 outliers out of 18,536 clicks in the *Increasing* condition, 14 outliers out of 13,217 clicks in the *Decreasing* condition and 1 outlier out of 15,935 clicks in the *Neutral* condition.

Figure 4 shows the average number of clicks (or number of "berries" picked) by observers for each berry bush in the three different conditions. The central question is whether observers react differently to different patterns of reward. A higher number of clicks shows that participants were searching longer in that condition. Unpaired *t*-test analyses revealed that observers picked more berries in the *Increasing* condition than in the Neutral condition  $[t(24) = 2.205, P = 0.0373, r^2 = 0.1685]^2$ and fewer berries in the *Decreasing* condition [t(24) = 2.976,P = 0.0066,  $r^2 = 0.2695$ ], although there were an equal number of good berries in all conditions. Not surprisingly, the difference between Increasing and Decreasing conditions was significant [t(24) = 4.267, P = 0.0003,  $r^2 = 0.4313$ ]. These results demonstrate that observers did react to different patterns of reward. When reward increased over time, observers searched longer. When reward decreased over time, participants quit searching more quickly.

As picking progresses in a bush, the chance that the most recent selection is a target berry goes down because observers tend to start with the reddest items (the proverbial "low hanging fruit") and progress to items less likely to be targets. We examined whether the 'positive predictive value (PPV)' of the final click in a bush differed between conditions. As shown in Fig. 5, unpaired t-test analyses revealed that the probability that the final click was a target is larger in the Neutral condition than in the *Increasing* condition [t(24) = 2.725, P =0.0118,  $r^2 = 0.2363$ ]. That probability is smaller than the PPV in the *Decreasing* condition [t(24) = 1.425, P = 0.1671, $r^2 = 0.0780$ ]. Not surprisingly, the difference between the Increasing and the Decreasing conditions is also significant  $[t(24) = 3.772, P = 0.0009, r^2 = 0.3721]$ . This result follows from the differences in the time spent in each bush. If you forage for a longer time (e.g., in the increasing condition), you will tend to pick more non-targets because all the obvious targets will have been picked.

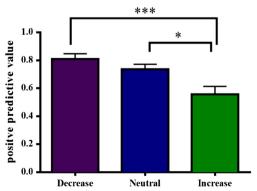


**Fig. 4** Average number of clicks observers made in each condition. Each error bar represents +/- 1 standard error of the mean (SEM)

As might be expected, observers' behaviors evolve as they learn from their experience with the different patterns of reward. This can be seen in Fig. 6, which plots average number of clicks in a patch as a function of block. Here, block serves as a coarse measure of experience. An ANOVA with Condition and Block as factors reveals a main effect of Condition [F(2,36) = 11.24, P = 0.0002], replicating the earlier analysis associated with Fig. 4. An interaction of Condition and Block [F(4,72) = 3.543, P = 0.0107] reflects the increasing size of the effect of Condition with experience, and indicates that humans acquire more knowledge about the reward pattern and change their foraging behaviors, accordingly, over the course of the experiment.

# Modeling foraging behavior under different reward patterns

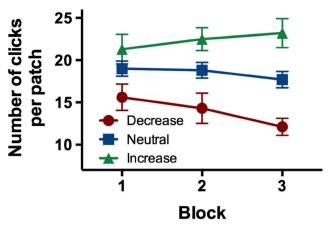
In previous work, we have found that the MVT provides a good description of average behavior on simple versions of this task, akin to the *Neutral* condition here (Wolfe, 2013; Zhang et al., 2015); MVT predicts that foragers should leave the current bush when the instantaneous rate of input drops below the average rate. In this section, we examine whether a



**Fig. 5** The probability that the final click in a patch will be a target. *Error bars* +/-1 SEM



<sup>&</sup>lt;sup>2</sup> For an estimate of effect size we are using  $r^2$ —the square of the effect size, r, not the correlation coefficient—calculated as  $t^2/(t^2+df)$ . See, for example, http://www.statisticslectures.com/topics/effectsizeindependentsamplest/#.  $r^2 = 0.01$  is considered a small effect, 0.09 a medium effect, and 0.25 a large effect.



**Fig. 6** Average number of clicks in a patch as a function of Condition and Block. *Error bars* +/- 1 SEM

model based on MVT is still a useful framework to understand human foraging behaviors under different patterns of changing reward.

The data reported above rejects any model that does not include a role for reward. For example, a model that proposes that observers keep picking until the PPV drops below some criterion value can be rejected because it would predict no difference between conditions.

A model that considers the reward pattern should do better. Thus, if we replace PPV with the expected value (EV) of the click (EV =  $PPV \times Reward$ ), then a threshold will be reached fastest in the Decreasing condition and slowest in the *Increasing* condition. The threshold is a free parameter, which can be constrained by adopting an MVT framework. To simulate the foraging behavior of this modified MVT model, we created simulated bushes containing 20 "good" and 20 "bad" berries. The signal strength of each berry ("redness" in the actual experiment) was drawn from one of two normal distributions; good and bad distributions, separated by d' = 2. Note that these simulated bushes mimic the actual experiment and match the experimental parameters. We assumed that different models would pick berries in order from the largest signal (reddest) to the smallest signal. While human observer would be imperfect in this ordering, the simulation models pick berries in a strictly decreasing order of signal strength. We simulated the *Increasing*, *Decreasing*, and *Neutral* conditions for 10,000 trials to get the predicted results of different models.

For the foraging models based on MVT, the time to leave the current patch is when the current yield from searching (e.g., targets found per unit time) drops below the average rate. To compute the current yield (instantaneous rate), we first calculate the EV for each click; how many points are awarded for the first click, second, and so forth? To obtain an instantaneous rate, the EV is divided by the average click time for each click where "click time" is the time since the preceding click. To compute an average rate, we calculated the total reward obtained divided by the total

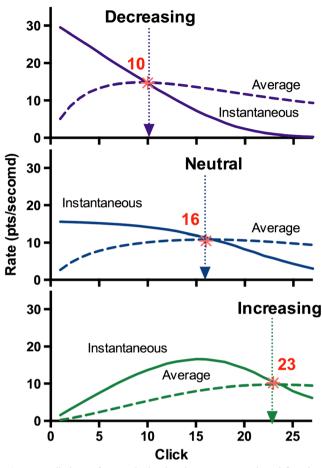
time spent in the experiment (including the "travel time" between bushes). The EV of each click is calculated as the average reward of each click based on 10,000 trials (e.g., the first click in the *Increasing* condition would have a EV of 2, assuming that observers always got the first click correct and collected 2 points).

To calculate the instantaneous rate, the EV is divided by the click time of each click. We employ the average time between clicks made by human foragers. Paired t-test analyses revealed that the time between clicks sped up over blocks during the task. The average click time between clicks was 1.28 s, 1.23 s and 1.22 s separately in three blocks of the Increasing condition [comparing the first block and the last block, t(12) = 2.387, P = 0.034, two-tailed]. The average click time between clicks was 1.3 s, 1.26 s and 1.23 s separately in three blocks of the *Decreasing* condition [comparing the first block and the last block, t(12)=3.902, p=0.002, twotailed]. The average click time between clicks was 1.27 s, 1.23 s and 1.22 s separately in three blocks of the Constant condition [comparing the first block and the last block, t(12) = 3.138, P = 0.008, two-tailed]. Paired t-test analyses revealed that there are no significant changes about the click time between clicks during a trial (bush). Based on these observations, we employ the same click time for different trials in the same block but different click times for different blocks (i.e., the instantaneous rate for the first block in the Increasing condition was computed as EV / 1.28 s, and the instantaneous rate for the second block in the Increasing condition was computed as EV / 1.23 s).

The simulated average rate was computed as the accumulated rewards in each simulation divided by the accumulated time (including a 6.2 s "travel time" between displays, derived from the experimental data). The actual travel time varies somewhat from trial-to-trial. Specifically, the actual "travel time" was greater than the preset "travel time" (3 s) because the actual delay included the time for the data recording and the generation of next display. We use the mean of the actual "travel time" to calculate the simulated average rate.

Figure 7 shows instantaneous and average rates as a function of click number for this model for all three conditions. The solid curves show instantaneous rates defined as the average point value of the current simulated click divided by the average time since the prior click. Taking the *Increasing* condition as an example, the rate starts low because the reward starts low. As the optimal model picks more berries in the bush, the instantaneous rate increases, reaching a peak at the 15th click. Note that, even though the reward keeps increasing, the instantaneous rate starts to decrease because the probability of picking a good berry decreases, and this probability decreases more than compensates for the increase in point value. The average rates are shown by dashed lines. In the *Increasing* condition, if observers always only pick a few berries, the average rate is low because the rewards for early





**Fig. 7** Predictions of a marginal value theorem (MVT)-based foraging model with knowledge about reward pattern in the three conditions of the experiment. The *red asterisk* marks the point at which the modeled instantaneous rate drops below the average rate and MVT predicts that the observer should quit. See text for details

berries are low, and the travel time is a larger and unproductive part of the total time. The average rises and declines somewhat as the chance of picking a good berry declines. The instantaneous rate drops below the average rate in the *Increasing* condition after 23 clicks. Accordingly, MVT predicts that observers should leave after those 23 clicks in the *Increasing* condition. As Fig. 7 shows, observers should leave after 16 clicks in the *Neutral* condition and 10 clicks in the *Decreasing* condition (or after 24, 17, and 11 clicks in the *Increasing*, *Neutral* and *Decreasing* conditions, if one assumes that observers leave after the first click whose yield falls below the average yield).

By comparing the prediction results of the modified MVT model with the actual data shown in Fig. 8, we can see that the model works quite well by the third block. For the *Increasing* condition, there is a marginally significant difference between the actual number of clicks and the predicted number of clicks in the first session [t(12) = 1.618, P = 0.0658,  $r^2 = 0.1791$ ] but no significant difference between them in the next two sessions [the second session: t(12) = 1.074, P = 0.1519,  $r^2 = 0.1519$ ,  $r^2 = 0.$ 

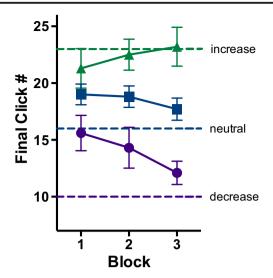


Fig. 8 Number of berries picked in each bush of three different blocks in each of the three conditions of the experiment. *Dashed lines* Model predictions of the point at which average and instantaneous rates are the same. *Error bars* SEM

0.0878; the third session: t(12) = 0.3635, P = 0.3613,  $r^2 = 0.3613$ 0.0109]. In the Decreasing condition, there is a significant difference between the actual number of clicks and the predicted number of clicks in the first two sessions [the first session: t(12) = 2.896, P = 0.0067,  $r^2 = 0.4114$ ; the second session: t(12) = 1.883, P = 0.0421,  $r^2 = 0.2281$ ] and no significant difference in the third session [t(12) = 1.067, P =0.1534,  $r^2 = 0.0867$ ]. Finally, in the *Neutral* condition, there is a significant difference between the actual and predicted number of clicks in the first two sessions [the first session: t(12) = 2.296, P = 0.0203,  $r^2 = 0.3052$ ; the second session: t(12) = 2.110, P = 0.0283,  $r^2 = 0.2706$ ] but no significant difference in the last session  $[t(12) = 0.7247, P = 0.2413, r^2]$ = 0.0419]. This analysis does not prove that the modified MVT model is the only correct model. However, it does show that human observers' foraging behaviors evolve with experience. Thus, a learning factor should be part of any successful foraging model that seeks to account for the effects of reward.

To illustrate this point, we added a learning parameter to the modified MVT model and calculated the learning rate of each block. We use a learning rate parameter, L, to denote how completely the observers have learned about the reward pattern. This value can vary from 0 to 1. The expected reward of the k-th target in the Increase condition becomes 2 + 2(k-1) \* L and the expected reward of the k-th target in the Increase condition is 40 - 2(k-1) \* L, shown in Fig. 9. For example, if L = 1, the expected reward would be 2,4,...40; if L = 0, the expected reward would be 2,2,...,2 in the Increase condition.

Based on the expected reward of each target, the current yield and the average rate for the learning MVT model can be computed to get the predicted click number when the observer will leave the current patch. By comparing the actual



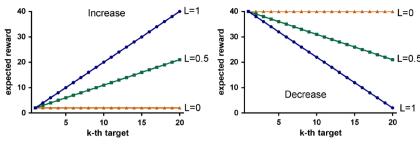


Fig. 9 Expected reward with partial knowledge about reward pattern. If L = 1, observers have full knowledge of the reward pattern. If L = 0, observers have no knowledge about the reward pattern

click number of each observer with the predicted click number, we can estimate the learning rate, L, that minimizes the error for each observer. The learning rates of different observers in the *Increase* and the *Decrease* condition can be seen in Fig. 10. In both the *Increase* and the *Decrease* condition, the learning rates appear to grow over different blocks of the experiment, suggesting that observers may gain and apply knowledge about the reward pattern. An ANOVA with Condition and Block as factors shows a main effect of Block [F(2,48) =5.872, P = 0.0052, indicating that observers were learning during the course of the experiment. There is no significant interaction of Condition and Block [F(2,48) = 0.07918, P =0.9240], reflecting the fact that there is no difference in how the learning rates emerge across the two conditions. We further analyze the data in the increase and decrease conditions separately: paired *t*-test on the difference between the learning rates of the first block and the final block in the Increase condition  $[t(12) = 1.837, P = 0.0911, r^2 = 0.2195]$ ; difference between the learning rates of the first block and the final block in the Decrease condition [t(12) = 2.259, P = 0.0433,  $r^2 = 0.2984$ ].

We compare the predicted click number of the learning MVT model with the actual click number, as shown in Fig. 11. A three-way ANOVA with Condition (increase, decrease), Block (1, 2, 3) and Model (predicted, actual) as factors shows a main effect of Condition [F(1,144) = 212.9, P < 0.0001], reflecting the fact that foraging model and observers would pick more berries in the *Increase* Condition. There is no significant effect of Model [F(1,144) = 0.0091, P = 0.9243], demonstrating that the learning MVT model does well in its prediction of human foragers' behaviors. Furthermore, unpaired *t*-test revealed that there is no significant difference

between the predicted click number and the actual click number in first block [t(24) = 1.147, P = 0.2626,  $r^2 = 0.0520$ ], second block [t(24) = 0.7051, P = 0.4875,  $r^2 = 0.0203$ ] or third block [t(24) = 0.2035, P = 0.8404,  $r^2 = 0.0017$ ] of the *Increasing* condition. Unpaired *t*-test also revealed that there is no significant difference between the predicted click number and the actual click number in first block [t(24) = 1.102, P = 0.2812,  $r^2 = 0482$ ], second block [t(24) = 0.7684, P = 0.4497,  $r^2 = 0.0240$ ] or third block [t(24) = 0.1318, P = 0.8962,  $r^2 = 0.0007$ ] of the *Decreasing* condition.

The preceding analyses suggests that the learning MVT model is doing what we wanted it to do, but this is a case of seeking support for a null hypothesis. Accordingly, we compared the predicted click number of the learning MVT model with the actual click number using Grouped Bayesian t-tests (Rouder, Speckman, Sun, Morey, & Iverson, 2009) to see if there is a strong preference for the null hypothesis in this case. Grouped Bayesian t-tests showed the favor of the null hypothesis in first block (Scaled JZS Bayes Factor = 1.702341), second block (Scaled JZS Bayes Factor = 2.293162) and third block (Scaled JZS Bayes Factor = 2.714465) of the *Increasing* condition. Similarly, Grouped Bayesian t-tests showed the favor of the null hypothesis in first block (Scaled JZS Bayes Factor = 1.765628), second block (Scaled JZS Bayes Factor = 2.215904) and third block (Scaled JZS Bayes Factor = 2.738877) of the Decreasing condition. Within a Bayesian framework, our results can be said to provide marginal evidence for the null over the alternative hypothesis.

The main goal of this section was to see whether MVT still provides a good start for a description of average behavior on simple versions of this task, as it did the earlier foraging works

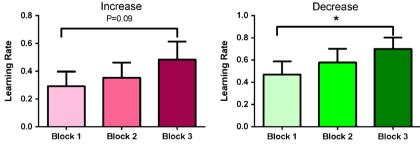


Fig. 10 Mean learning rates in the three blocks of the Increase and Decrease conditions. Error bars SEM



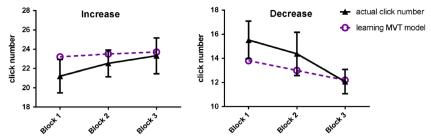


Fig. 11 Comparison between human observers and the learning MVT model. Purple dashed lines Predictions of the learning MVT model, black solid lines actual click number. Error bars SEM

(Wolfe, 2013; Zhang et al., 2015). Our data indicate that prediction models that do not include a role of reward, and a role of learning can be rejected since they would predict no difference between conditions and predict no difference between blocks. Our simulations indicate that an MVT model with a learning parameter and knowledge about the reward pattern gives a credible account of the data. We are not trying to rule out other models, only to argue that any successful model will need to include learning and knowledge of reward.

#### **Discussion**

A large body of work has examined the usefulness of the MVT and related concepts in the understanding of animal foraging (Stephens, Brown, & Ydenberg, 2007; Stephens & Krebs, 1986). More recently, this theoretical perspective has been applied to human behavior in visual search (Wolfe, 2013) and in human cognition more generally (Cain et al., 2012; Constantino & Daw, 2015; Fougnie et al., 2015; Hills, Jones, & Todd, 2012; Hutchinson et al., 2008; Wilke, Hutchinson, Todd, & Czienskowski, 2009). In the present work, we show that MVT continues to be a good starting point for an analysis of foraging behavior when we complicate the situation by adding different values to different foraged items.

The role of value in the control of visual attention has become a topic of interest in recent years (Anderson, Laurent, & Yantis, 2011, 2012; Anderson & Yantis, 2013; Della Libera & Chelazzi, 2009; Hickey, Chelazzi, & Theeuwes, 2010a, 2010b; Wang, Yu, & Zhou, 2013). The present work shows that value shapes foraging behavior in a manner predicted by MVT. The role of value appears to take some time to be learned in this experiment. Looking again at Fig. 10, the learning rate in the third block is significantly larger than that in the first block as the observers have learned to better extract the available value from each situation. These results suggest that a foraging model should consider both a role of reward and a role of learning in predicting human foraging behaviors. Simulation results of this learning variant of the MVT model demonstrate its power to predict human foraging behaviors under reward changing patterns. While the MVT model gives a good starting point for an analysis of foraging behaviors, non-MVT models which take the reward pattern and the learning factor into account, may also predict similar behaviors.

Beyond a fundamental interest in human foraging and decision-making, why should we care about these effects? People rarely search exhaustively when foraging for items of interest in the world. For example, when humans or animals harvest resources like berries in a bush, they are unlikely to pick all ripe berries before moving to a new bush (Wolfe, 2013), and this is what would be predicted by optimal foraging models. However, there are times when optimal foraging is not optimal. Sometimes, it is critically important to find all the targets; all the tumors, all the weapons, etc. In these situations, a tendency to stop short of exhaustive search remains even when the goal is to find everything (Ashman et al., 2000; Berbaum et al., 2000; Franken et al., 1994; Samuel et al., 1995). In medical settings, such errors have proven difficult to eradicate by means of interventions such as computer-aided diagnoses (CAD) or checklists (Berbaum et al., 2006, Berbaum et al., 2007). The present results show how a payoff strategy could be used to move an observer toward more complete search. If you counteract diminishing returns with increasing rewards, observers will stay with the current task for a longer period of time. At some level, this seems like common sense, but common sense is not an invariable guide to human behavior so it is useful to have data to back intuition.

An increasing pattern of reward is far from a cure for errors of omission in foraging search tasks. First, when stimuli are ambiguous, as they are here, increasing the number of clicks increases both hits and false alarm errors. In real world tasks like airport security or medical screening, false alarms incur a smaller cost than miss errors but they are much more common because targets are typically rare (Evans, Birdwell, & Wolfe, 2013; Wolfe, Brunelli, Rubinstein, & Horowitz, 2013). Increasing reward might be more applicable when targets, once found, are easily identified. If each successive typo in a manuscript was worth more than the last, proofreaders would probably find more of them. An experiment with stimuli of this sort would be an interesting sequel to the present work.

Second, a linearly increasing pattern of reward does increase the percentage of targets found but not to 100%. In the present experiment, observers found 57% of targets in



the *Decreasing* condition, 76% in the *Neutral* condition, and 82% in the *Increasing* condition. Pushing that 82% to 100% is difficult because so many of the berries are 'bad' berries once the bulk of the good berries have been picked. Again, it may be easier to move observers to 100% with unambiguous but still hard to find items.

Third, how different would the foraging behaviors be if a bonus could be awarded after some unknown and variable number of targets was found? Would the observers forage longer in the display and try to get more bonuses? An experiment with this kind of bonus would be another interesting sequel to this work.

In sum, it seems straightforward to tell your child to pick up all the Lego pieces in his room or to ask your radiologist to find every metastasis in a liver CT. However, once these tasks are understood as foraging tasks, it becomes clear that there are strong pressures that make it unlikely that the task will be performed perfectly. Understanding how humans forage is a start to helping them forage more successfully.

Acknowledgements This research was supported by grants to J.M.W. from ONR MURI (N000141010278), NIH-NEI (EY017001), Hewlett-Packard, Google, CELEST (NSF SBE-0354378) and by the National Natural Science Fund of China (Grant numbers 61233011, 61374006, 61473086); Major Program of National Natural Science Foundation of China (Grant number 11190015); Natural Science Foundation of Jiangsu (Grant number BK20170692, BK20131300); the Innovation Fund of Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education (Nanjing University of Science and Technology, Grant number JYB201601); the Innovation Fund of Key Laboratory of Measurement and Control of Complex Systems of Engineering(Southeast University, Grant number MCCSE2017B01); the Fundamental Research Funds for the Central Universities (2242016k30009).

#### References

- Anderson, B. A., Laurent, P. A. & Yantis, S. (2011), 'Value-driven attentional capture', Proceedings of the National Academy of Sciences of the United States of America 108(25), 10367–10371.
- Anderson, B. A., Laurent, P. A. & Yantis, S. (2012), 'Generalization of value-based attentional priority', Visual Cognition 20(6), 647–658.
- Anderson, B. A. & Yantis, S. (2013), 'Persistence of value-driven attentional capture.', *Journal of Experimental Psychology: Human Perception and Performance* 39(1), 6.
- Ashman, C., Yu, J. & Wolfman, D. (2000), 'Satisfaction of search in osteoradiology', American Journal of Roentgenology 177, 252–253.
- Berbaum, K., Caldwell, R., Schartz, K., Thompson, B. & Franken E. Jr. (2007), 'Does computer-aided diagnosis for lung tumors change satisfaction of search in chest radiography?', Academic Radiology 14, 1069–1076.
- Berbaum, K., Franken E. Jr., Caldwell, R. & Schartz, K. (2006), 'Can a checklist reduce SOS errors in chest radiography?', Academic Radiology 13, 296–304.
- Berbaum, K., Franken, E., Dorfman, D., Caldwell, R. & Krupinski, E. (2000), 'Role of faulty decision making in the satisfaction of search effect in chest radiography', Academic Radiology 7, 1098–1106.

- Berbaum, K., Franken, E., Dorfman, D., Rooholamini, S., Kathol, M., Barloon, T., Behlke, F., Sato, Y., Lu, C. & el Khoury, G. (1990), 'Satisfaction of search in diagnostic radiology', *Investigative Radiology* 25, 133–140.
- Brainard, D. H. (1997), 'The psychophysics toolbox', *Spatial vision* 10, 433–436.
- Cain, M. S., Vul, E., Clark, K. & Mitroff, S. R. (2012). A Bayesian optimal foraging model of human visual search. *Psychological Science*, 1–8.
- Charnov, E. L. (1976), 'Optimal foraging, the marginal value theorem', Theoretical Population Biology 9(2), 129–136.
- Constantino, S. M. & Daw, N. D. (2015), 'Learning the opportunity cost of time in a patch-foraging task', Cognitive, Affective, & Behavioral Neuroscience 15(4), 837–853.
- Della Libera, C. & Chelazzi, L. (2009), 'Learning to attend and to ignore is a matter of gains and losses', *Psychological Science* 20(6), 778–784.
- Desimone, R. & Duncan, J. (1995), 'Neural mechanisms of selective visual attention', *Annual Review of Neuroscience* 18, 193–222.
- Evans, K. K., Birdwell, R. L. & Wolfe, J. M. (2013), 'If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening', *PLoS One* 8(5), e64366.
- Fleck, M., Samei, E. & Mitroff, S. (2010), 'Generalized "satisfaction of search": adverse influences on dual-target search accuracy', *Journal* of Experimental Psychology: Applied 16, 60–71.
- Fougnie, D., Cormiea, S. M., Zhang, J., Alvarez, G. A. & Wolfe, J. M. (2015), 'Winter is coming: how humans forage in a temporally structured environment', *Journal of Vision* 15, 1–17.
- Franken, E., Berbaum, K., Lu, C., Kannam, S., Dorfman, D., Warnock, N., Simonson, T. & Pelsang R.E. (1994), 'Satisfaction of search in the detection of plain-film abnormalities in abdominal contrast studies', *Investigative Radiology* 29, 403–409.
- Hickey, C., Chelazzi, L. & Theeuwes, J. (2010a), 'Reward changes salience in human vision via the anterior cingulate', *Journal of Neuroscience* 30(33), 11096–11103.
- Hickey, C., Chelazzi, L. & Theeuwes, J. (2010b), 'Reward guides vision when it's your thing: trait reward-seeking in reward-mediated visual priming', *PLoS One* 5(11), e14087.
- Hills, T. T., Jones, M. N. & Todd, P. M. (2012), 'Optimal foraging in semantic memory.', *Psychological Review* 119(2), 431–440.
- Hutchinson, J. M. C., Wilke, A. & Todd, P. M. (2008), 'Patch leaving in humans: can a generalist adapt its rules to dispersal of items across patches?', *Animal Behaviour* 75, 1331–1349.
- Koch, C. & Ullman, S. (1985), 'Shifts in selective visual attention: towards the underlying neural circuitry', *Human Neurobiology* 4, 219–227.
- McNamara, J. (1985). An optimal sequential policy for controlling a Markov renewal process. *Journal of Applied Probability*. 324–335.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D. & Iverson, G. (2009), 'Bayesian t tests for accepting and rejecting the null hypothesis', Psychonomic Bulletin & Review 16(2), 225–237.
- Samuel, S., Kundel, H., Nodine, C. & Toto, L. (1995), 'Mechanisms of satisfaction of search: eye position recordings in the reading of chest radiographs', *Radiology* 194, 895–902.
- Stephens, D. W., Brown, J. S. & Ydenberg, R. C. (2007). Foraging: behavior and ecology, University of Chicago Press.
- Stephens, D. W. & Krebs, J. R. (1986). Foraging theory, Princeton University Press.
- Treisman, A. M. & Gelade, G. (1980), 'A feature-integration theory of attention', Cognitive Psychology 12, 97–136.
- Tsotsos, J., Culhane, S., Wai, W., Lai, Y., Davis, N. & Nuflo, F. (1995), 'Modeling visual attention via selective tuning', Artificial Intelligence 78, 507–545.
- Wang, L., Yu, H. & Zhou, X. (2013), Interaction between value and perceptual salience in value-driven attentional capture', *Journal of Vision* 13(3).



- Wilke, A., Hutchinson, J., Todd, P. & Czienskowski, U. (2009), 'Fishing for the right words: decision rules for human foraging behavior in internal search tasks.', *Cognitive Science* 33, 497–529.
- Wolfe, J. (1994), 'Guided search 2.0: a revised model of visual search', Psychonomic Bulletin & Review 1(2), 202–238.
- Wolfe, J. (2012), Approaches to visual search: feature integration theory and guided search. In: Nobre, A.C. & Kastner, S. (Eds) Oxford Handbook of Attention, New York: Oxford University Press.
- Wolfe, J. M. (2013), 'When is it time to move to the next raspberry bush? Foraging rules in human visual search', *Journal of Vision* 13, 1–17.
- Wolfe, J. M., Brunelli, D. N., Rubinstein, J. & Horowitz, T. S. (2013), 'Prevalence effects in newly trained airport checkpoint screeners: trained observers miss rare targets, too', *Journal of vision* 13(3), 33–33.
- Zhang, J., Gong, X., Fougnie, D. & Wolfe, J. M. (2015), 'Using the past to anticipate the future in human foraging behavior', *Vision Research* 111, 66–74.

