

# Probability Matching on a Simple Simulated Foraging Task: The Effects of Reward Persistence and Accumulation on Choice Behavior

Zack Ellerby<sup>1</sup> and Richard J. Tunney<sup>2</sup>

<sup>1</sup> School of Computer Science and School of Psychology, University of Nottingham, Nottinghamshire, England

<sup>2</sup> School of Psychology, Aston University, Birmingham, England

## ABSTRACT

Over a series of decisions between two or more probabilistically rewarded options, humans have a tendency to diversify their choices, even when this will lead to diminished overall reward. In the extreme case of probability matching, this tendency is expressed through allocation of choices in proportion to their likelihood of reward. Research suggests that this behaviour is an instinctive response, driven by heuristics, and that it may be overruled through the application of sufficient deliberation and self-control. However, if this is the case, then how and why did this response become established? The present study explores the hypothesis that diversification of choices, and potentially probability matching, represents an overextension of a historically normative foraging strategy. This is done through examining choice behaviour on a simple simulated foraging task, designed to model the natural process of accumulation of unharvested resources over time. Behaviour was then directly compared with that observed on a standard fixed probability task (cf. [Ellerby & Tunney, 2017](#)). Results indicated a convergence of choice patterns on the simulated foraging task, between participants who acted intuitively and those who took a more strategic approach. These findings are also compared with those of another similarly motivated study ([Schulze, van Ravenzwaaij, & Newell, 2017](#)).

## KEYWORDS

probability matching  
foraging  
apophenia  
heuristics  
judgement under uncertainty

## INTRODUCTION

Consider a fair die with four red and two green sides. If it were to be rolled 100 times, the maximum expected number of correct predictions would be 66.7, which can be achieved by exclusively predicting that it will land on red. However, given the task of maximising correct predictions, many respondents will instead tend to predict red on closer to two-thirds of rolls, although this will lead to an average of only 55.6 correct predictions. This is an example of probability matching, a behavioural tendency towards the suboptimal diversification of choic-

es, which is often characterised by selecting options in approximate proportion to their outcome contingencies (cf. [Vulkan, 2000](#); [Shanks, Tunney, & McCarthy, 2002](#)).

A recent study examined multiple putative factors in determining maximization versus diversification of choices ([Ellerby & Tunney,](#)

Corresponding author: Zack Ellerby, School of Computer Science, University of Nottingham, Nottinghamshire, England.

e-mail: [zack.ellerby@nottingham.ac.uk](mailto:zack.ellerby@nottingham.ac.uk)

2017). Behaviour was assessed on a standard probability learning task, in which reward probabilities were fixed, independent and identically distributed (IID). In this case, investigation focused on the contributions of two key factors, identified a priori. First, heuristics, in that matching might represent an overwhelmingly intuitive response. Second, apophenia, in that matching might be driven by the misperception of sequential dependencies in the reward sequence. These were not initially presumed to be mutually exclusive. In particular, it was considered that taking a more intuitive versus strategic approach might influence levels of apophenia, which could provide a vehicle for this effect. Results indicated that taking an intuitive versus strategic approach to the task is the predominant factor behind matching versus maximising behaviour. This is consistent with a heuristic-based account of probability matching, in which matching is the more readily available response (Koehler & James, 2010; Kogler & Kuehberger, 2007). In this case, probability matching is dominant when acting intuitively, but may be overruled upon the recognition and application of the maximising strategy through sufficient deliberation and subsequent cognitive control. By contrast, although self-reported apophenia was both prevalent and found to be associated with taking an intuitive approach to the task, results indicated no link between apophenia and choices. No evidence was found in support of apophenia-based accounts of probability matching (cf. Gaissmaier & Schooler, 2008; Unturbe & Corominas, 2007; Wolford, Newman, Miller, & Wig, 2004).

Therefore, if probability matching is truly a heuristic response, but this effect is not mediated by induced apophenia, then this raises the question of how and why this seemingly suboptimal behaviour came to be established. The present study was designed to explore one other possible answer to this question, which stems from fundamental differences between the artificial decision tasks upon which matching has been generally observed and their more natural counterparts.

Repeated choice between two options, each having fixed probabilities of reward that are independent of their level of exploitation, is a situation that is rare in nature. It therefore seems unlikely that sufficient selection pressure will have been applied in order to behaviourally establish the optimal solution to this specific problem. To better approximate problems that were commonly faced over a sustained period of evolutionary ancestry, such as the optimal allocation of foraging time, certain additional factors must be taken into account. Examples include competition, generally with conspecifics, and the potential for unharvested resources to accumulate and deplete over time. The latter may often occur in a manner that is contingent upon levels of exploitation, which can therefore provide a negative feedback mechanism. Crucially, accounting for these additional factors will radically alter what constitutes an optimal response pattern.

First, in a social environment, once a certain threshold of rivals is reached in competition for a more lucrative food source, taking advantage of a lesser but unchallenged resource becomes the better option. Moreover, the optimal spread of competitors is in direct proportion to the value of each resource. While natural selection does not primarily occur on the group level, the same principle is equally applicable at the individual level once it becomes prevalent in the population.

In fact, diversification of choices will confer a progressively greater competitive advantage in an environment in which competitors act as overmatchers or, in the most extreme case, as maximizers. Therefore, in terms of competitive foraging behaviour, there is potentially both a pressure for probability matching to become the dominant tendency and a means for it to remain so, as an evolutionarily stable strategy (cf. Gallistel 1990; Gigerenzer, 1996; 2000; Skyrms, 1996; 1997).

In behavioural ecology, the concept of the ideal free distribution, or IFD, (Fretwell & Lucas 1970, Fretwell 1972) has become unequivocally well-established. Foraging behaviour in accordance with the IFD has been consistently documented across a wide range of species (cf. Ellerby & Tunney, 2017; Weber, 1998). The phenomenon of near optimal group-level matching behaviour, when under the influence of competition, has also been experimentally demonstrated in humans. The majority of studies in this area used a variety of modified simple zero-sum probabilistic choice tasks (Kraft & Baum, 2001; Kraft, Baum, & Burge, 2002; Madden, Peden, & Yamaguchi, 2002; Sokolowski, Tonneau, & Baque, 1999), while Goldstone and Ashpole (2004) applied a more complex and realistic simulated foraging environment. A more recent study has also highlighted individual-level matching behaviour as being socially adaptive under competition, finding that choices were sensitive to the strategy employed by simulated competitors (Schulze, van Ravenzwaaij, & Newell, 2015).

It is also noteworthy that, although behaviour has been broadly consistent with the IFD across all of these studies, another common finding has been some degree of undermatching. This has been evident in the behaviour of both humans and other animals. Intriguingly, one study reported substantial undermatching when rewards were probabilistically distributed, whereas dividing points evenly between participants that had selected an option resulted in a shift to near IFD matching (Kraft et al., 2002).

There are pertinent environmental factors to take into account even when considering an individual in isolation. One such consideration is that many natural resources can become depleted, or even completely exhausted by overexploitation. In this case, it is no longer optimal to exclusively take advantage of a single resource, even if it begins at a very high-value, but instead to spread foraging behaviours across various options. The optimal allocation depends upon the rates of resource replenishment, but is, again, broadly proportional to each option's rate of reward.

To demonstrate this point, consider the simplified example of two areas of forest, with one area being more fertile and just over twice as likely to produce an edible item (e.g., fruit, nuts, or mushroom) over any given time period. In this instance an average of seven versus three items are generated in each area per day. A forager is able to harvest food from only one of these two areas each day and aims to allocate their foraging time between the two areas in a manner that maximizes average long-term reward. The first day, it is of course best to search within the more productive area. The next day is the same, with the first area still likely to contain more food despite the expected number of food items within the less fertile area having compounded over the two days to six. On the third day, however, the sum of food expected to

**TABLE 1.**  
Illustration of an Optimal Foraging Sequence, in a Simplified Two Area Model with Resource Generation Rates of Seven Versus Three and Satiety Limit of Ten Items Per Day

Trial number	Expected items: Area 1	Expected items: Area 2	Choice (assumes optimal)
1	3	7	Area 2
2	6	7	Area 2
3	9	7	Area 1
4	3	14	Area 2
5	6	11	Area 2
6	9	8	Area 1
7	3	15	Area 2
8	6	12	Area 2
9	9	9	Area 1
10	3	16	Area 2
11	6	13	Area 2
12	9	10	Area 2
13	12	7	Area 1
14	5	14	Area 2
15	8	11	Area 2
16	11	8	Area 1
17	4	15	Area 2
18	7	12	Area 2
19	10	9	Area 1
20 (repeat 10)	3	16	Area 2
21 (repeat 11)	6	13	Area 2

*Note.* This sequence is distinct from, though broadly analogous to, the probabilistic reward generation schedule applied in the simple simulated foraging task.

have grown in the unharvested region has reached nine. This surpasses the expected daily generation amount of the more productive area. It is therefore optimal in this instance to forage within the less productive area. The next day, it becomes even more advantageous to search in the more productive area than ever before, as the expected items of food within this area will be fourteen, compared with only three within the less fertile region. Following this point, however, the same pattern emerges of the less fertile area being the better option to forage within once every three days. Moreover, it is easy to slightly increase the complexity of this simple model so as to simulate the satiety of the forager - through introducing a maximum limit of ten items of food that can be taken from each area per day. Once incorporating this factor, the outcome for optimal foraging between the two areas repeats at an exact 7:3 ratio, closely matching the rates of resource replenishment (see Table 1).

Millions of generations of evolutionary ancestry have passed in which performance in natural foraging type situations would have been a prominent selection pressure. It seems likely that the tendency to diversify choices could have become established, through natural selection over this extended timescale, as what may be considered a default behaviour. By contrast, there has been a relatively negligible amount of time so far in which to adjust, at the genetic level, to the more rigid fixed probability schedules that have become more prevalent in

modern day situations, such as those underlying common gambling tasks (e.g., dice, slot-machines, or online equivalents) and which form the basis of the standard two-alternative forced-choice (2AFC) probability matching task. This provides a potentially compelling natural account of how probability matching might have become established, and remain to this day, as a prepotent intuitive response.

One might expect that the behaviour of nonhuman animals could valuably inform this account. Although group-level matching of foraging time to rewards available in different areas has been well-established in accordance with the IFD, examination of animal choice behaviour in social isolation, under closely controlled experimental conditions, has found somewhat equivocal results. Different behaviours have been found to predominate across the spectrum of matching to maximizing, depending upon species and specific task conditions (Behrend & Bitterman, 1961; Bullock & Bitterman, 1962; Graf, Bullock & Bitterman, 1964; Longo, 1964; Meyer, 1960; Wilson, Oscar & Bitterman, 1964). Although diversification of choices has often been observed, probability maximizing can be learned on concurrent variable-ratio schedules, in accordance with the generalized matching law (cf. Herrnstein, 1961; Herrnstein & Loveland, 1975; Davison & McCarthy, 1988; Poling, Edwards, Weeden & Foster, 2011), which nevertheless predicts probability matching on variable-interval schedules.

Although it is important to consider that some animal species have been found to approach probability maximizing under certain task conditions, we also propose that this behaviour does not necessarily preclude matching from being the initial instinctive response. It is possible that humans and other animals hold in common the tendency to instinctively match choices to reward contingencies, whether or not this may be overruled in certain cases by other processes. These include reinforcement learning and also, at least in humans, deliberative self-control. Moreover, although we expect that top-down executive control is generally applied to overrule the instinctive response according to reasoned understanding of the benefits of probability maximizing, it is also possible that such top-down control might sometimes be exerted with the opposite effect, instead incorrectly reinforcing the matching response as a result of mistaken heuristic- or apophenia- driven beliefs.

## Overview of the Experiment

While a body of existing literature has addressed whether human choice behaviour falls in line with the IFD under the influence of competition, at the time that the present study was conducted (2016), we were aware of none that had reported on equivalent behavioural comparisons for a purely individual situation, in which the adaptive benefit of choice diversification derives instead from the prospect of resource accumulation. Sugrue, Corrado and Newsome (2004) did include such an assessment of choice behaviour in Rhesus monkeys. In their study, choice proportions were found to match the ratio of rewards on a dynamic foraging task, in which choices were made between two options with independent and stochastic rates of reward assignment, but where assigned rewards would persist over time until the option was chosen. Since the time of conducting the present experiment, another study has been published with very similar theoretical motivation (Schulze,

van Ravenzwaaij & Newell, 2017). Here, the authors acknowledge that reward-persistence in an unchosen option will make diversification an effective choice strategy, then proceed to investigate the effects of persistent, or lingering, rewards on participants' actual patterns of choice behaviour.

The present study was originally designed to test whether the principles observed by Sugrue et al. (2004) extend to human behaviour. To do so, choice behaviour was examined on a simple simulated foraging task (SSFT). This task modifies the standard probability matching preparation to incorporate the stochastic accumulation and choice-driven depletion of resources over time into the underlying reward schedule. Task instructions and choice stimuli were also changed to superficially denote a foraging situation, although this still represents a simple 2AFC task. Further, choices were not examined exclusively on the SSFT. Behaviour on this task was also compared and contrasted with behaviour on the standard probability matching preparation, as reported in Ellerby and Tunney (2017). These two tasks were also compared for the effects of manipulating reward probability information and of self-reported intuition versus strategy use.

It will be informative to compare observed behaviour on this SSFT with the results obtained by Schulze et al. (2017), who compared human choices in a reward-hold condition, in which an unharvested reward would persist until collected, with a no hold condition. Although this manipulation is comparable with the accumulation of unharvested resources that is modelled in the SSFT, there are key differences between each of these conditions and their counterparts in the present study. In the reward-hold condition, only a single reward could persist at each option at any given time, irrespective of how long this option remained unexploited. By contrast, on the SSFT, resource accrual is permitted to continue indefinitely, in line with each option's respective resource generation probability. This means that there can be more than one reward held in the bank at an option that has not been chosen for multiple trials, although only a single reward can be taken on any given trial. In addition, in the no hold condition, Schulze et al. manipulated the fixed reward contingencies in order to match the discriminability between the two response options to that experienced by participants in the reward hold condition. As the present study draws its comparison against fixed-probability choices obtained from a pre-existing data-set, this procedure was omitted. Instead, an additional analysis was conducted to assess the relationship between the level of discriminability in outcome contingencies of the two options, as experienced by each respondent, and their subsequent choices.

It was hypothesised that responses on the SSFT would show a substantial degree of choice diversification, tending towards probability matching and potentially some degree of undermatching relative to the actual ratio of resource generation rates. It was also predicted that choices on the foraging task would show a relative absence of effects of either stated versus learned reward probabilities, or of intuition versus strategy use. When considering the SSFT together with the standard fixed probability task, the former of these should manifest as a significant interaction effect between task and probability condition.

## METHOD

### Participants

Sixty participants completed the SSFT. These were recruited in an opportunity sample, drawn primarily from the student population at the University of Nottingham. Of these, 17 were male and 43 female. Ages ranged from 17 to 44 ( $M = 24.03$ ,  $SD = 4.74$ ). This sample size was designed to match that of a fixed-probability task condition reported in a preceding study (cf. Ellerby & Tunney, 2017). In the standard task group (data taken from the aforementioned study), there were a further 60 participants, recruited in the same manner in a previous experimental session. In this group, 16 participants were male and 44 female, with an age range of 18 to 33 ( $M = 22.62$ ,  $SD = 3.09$ ).

Participants were paid an inconvenience allowance with a value contingent upon the choices they made during the task. Both optimal exploitation of resources on the SSFT and probability maximizing on the standard task would accumulate average total winnings of £5.88.

### Experimental Design

This experiment took a  $2 \times 42$  mixed model design, with probability condition (stated vs. learned) as a between-subjects factor and block (1 to 42) as a within-subject factor. In the stated probability condition, participants were informed that the two patches of land would generate resources probabilistically, at the different rates of 0.49 versus 0.21 and that resources would be allowed to accumulate over time, but that only a single reward could be taken on any one trial (see Appendix B). In the learned probability condition, participants were not provided with any information about reward accumulation rates. Which of the two choices had the higher rate of resource accumulation was counterbalanced between participants, in order to account for any bias in favour of a particular colour. All participants were randomly assigned to one of the two experimental conditions.

These data were also contrasted with equivalent data obtained from participants that completed a standard fixed probability task, in a  $2 \times 2 \times 42$  design, with task as an additional between-subjects factor. This comparison group comprised participants who completed the standard task condition of the preceding study (Ellerby & Tunney, 2017).

### Stimuli

A plain white background was maintained throughout the task. A black central fixation cross was used between trials. Stimuli consisted of images of two identical patches of woodland, with a blue and a yellow signpost to distinguish between them (see Figure 1). These were positioned equidistant from the centre and the left or right edge of the screen, with the side of presentation varying randomly between trials. The experimental procedure was designed and coded within Psychopy (Peirce, 2007).

### Task Procedure

For comparison purposes, the procedure was based closely upon that used in the preceding study (Ellerby & Tunney, 2017). The task was an

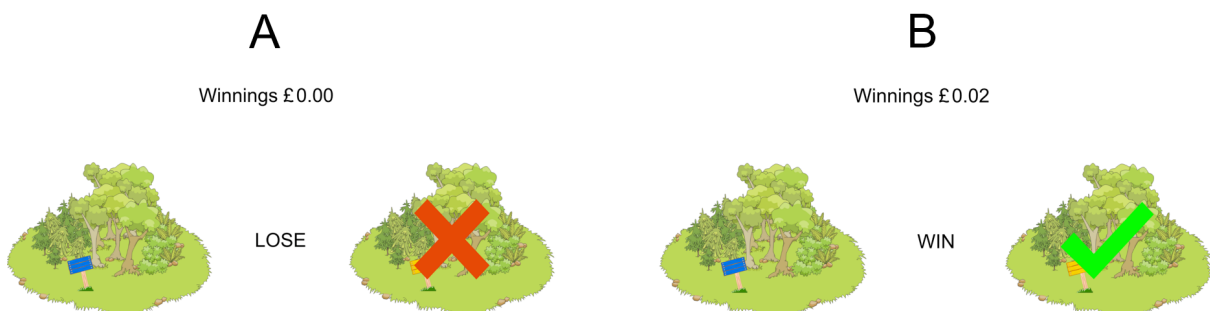
**FIGURE 1.**

Example choice stimuli for the simple simulated foraging task.

iterated 2AFC in which participants made repeated decisions between the two patches of land over 420 trials with the aim of maximizing total pay-offs over the entire task. There was no initial training period. Task instructions provided to participants are shown in Appendices A and B. Each trial began with a black central fixation cross upon a white background. The duration of the fixation varied randomly between one and two seconds. This was followed by the appearance of the two patches of land, to the left and right of the computer screen. The right-left position of the two patches of land varied randomly between trials and was not predictive of reward. Participants were explicitly told in the instructions to make their choices on the basis of identity (colour of the signpost) rather than the position of the patch of land (see Appendix A). Upon presentation of the stimuli, participants were free to make their predictions by pressing the *q* or *p* keys on a standard QWERTY keyboard to indicate the stimulus to the left or right side of the screen, respectively. Shortly after each prediction (500 ms) a green tick or red cross was overlaid on the image of the patch of land that had been chosen (see Figure 2). This graphic was accompanied by either the word *WIN* or *LOSE* shown to the centre of the screen. Participants received a £0.02 reward for each correct prediction.

## Simple Simulated Foraging Task Reward Schedule

The probabilities of a new resource appearing in each patch were fixed throughout the task, at 0.49 and 0.21 for the high and low accumulation rate options, respectively. However, uncollected rewards were allowed to both accumulate and persist over time, indefinitely over the entire task. This meant that actual reward contingencies varied over time in a manner that depended upon participants' previous choices.

**FIGURE 2.**

Example feedback stimuli for the simple simulated foraging task. Panel A shows an unsuccessful forage attempt. Panel B shows a successful forage attempt.

An illustration of the effects of this process of reward accumulation on the actual reward contingencies of each option, depending upon the number of trials since last observed to be exhausted, is shown in Table 2. Notably, although multiple rewards could accrue over time, only a single reward could be obtained on any one trial, leaving any remaining resources available for future exploitation. This stipulation was designed to mimic the effects of satiety in a natural environment. It also guaranteed two important principles. First, that the reward obtained on any given trial was always of equal magnitude. Second, that a substantial degree of diversification was necessary to fully exploit resources over the task. In practice, this precludes the potential strategy of allowing resources to extensively accrue within one unselected option, and then collecting all unharvested rewards through single sporadic selections of this underexploited area.

Nevertheless, a range of possible choice diversification ratios allowed an approximately optimal exploitation of resources on the SSFT. When considering the selection of the option with the highest likelihood of reward on each individual trial, choices repeat at a 2:1 ratio of high to low resource generation rate options. However, exploration ratios of either 3:1 or 1:1 could also keep expected resources in each option from exceeding one, and therefore prevent a wasteful "runaway" accumulation of unharvested resources in either option. We also conducted calculations to estimate the proportions of high and low generation rate choices necessary to bring back runaway accumulated resources to normal levels, given extreme initial choice conditions. First, in the case of 100 initial choices of the high generation rate option, we calculated that an average of 69.28 subsequent choices would be necessary to bring available resources in each option back to baseline, with a ratio of .79 high generation rate choices over the entire sequence (inclusive of the initial 100 trials). Following 100 initial choices of the low generation rate choices, we calculated that an average of 162.21 subsequent choices would be required, with a ratio of .49 high generation rate choices over the entire sequence.

The accumulation rates of both patches were determined with the primary aims of remaining proportional to the reward probabilities used in the fixed probability tasks reported in Ellerby and Tunney (2017), while also summing to 0.7. This means that an optimal strategy could expect to obtain a reward on an average of 70% of trials, maintaining maximum pay-offs on the SSFT as equivalent to those on the

**TABLE 2.**

Probability of Reward Presence and Overall Expected Reward at Each of the Two Options, Depending Upon Resource Generation Rate and Number of Trials Since Last Known to be Exhausted

Trials since exhaustion	Probability of one or more reward		Total expected reward	
	GR = .21	GR = .49	GR = .21	GR = .49
1	.210	.490	0.21	0.49
2	.376	.740	0.42	0.98
3	.507	.867	0.63	1.47
4	.610	.932	0.84	1.96
5	.692	.965	1.05	2.45
6	.757	.982	1.26	2.94
7	.808	.991	1.47	3.43
8	.848	.995	1.68	3.92
9	.880	.998	1.89	4.41
10	.905	.999	2.10	4.90

Note. GR = resource generation rate.

standard fixed-probability task. We considered this to be an important prerequisite for the purposes of directly comparing behaviour across the two tasks, as previous studies have found that increased financial incentives lead to a reduction in choice diversification (Brackbill, Starr, & Kappy, 1962; Shanks et al., 2002; Siegel & Goldstein, 1959). Participants in the stated probability condition were fully informed of this reward schedule before beginning the task, while those in the learned probability condition began completely naïve to it. Neither group received any external indicator of the amount of reward currently held in each option. Rather, they were required to estimate this in their own capacity, using the knowledge they had learned about the underlying reward schedules.

## Questionnaire

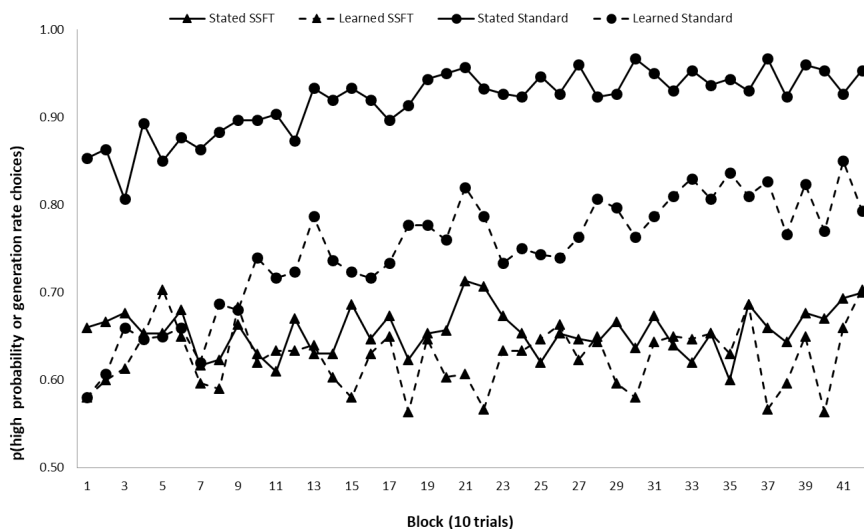
At the end of the experiment, participants completed a questionnaire, based on that used in Ellerby and Tunney (2017). Participants were asked to estimate the proportion of trials in which a new resource had

appeared in each colour patch (0% to 100%), the extent to which they relied upon intuition or strategy in making their predictions (1 to 5), and the extent to which they believed there to have been a predictable pattern in the reward sequence (1 to 5).

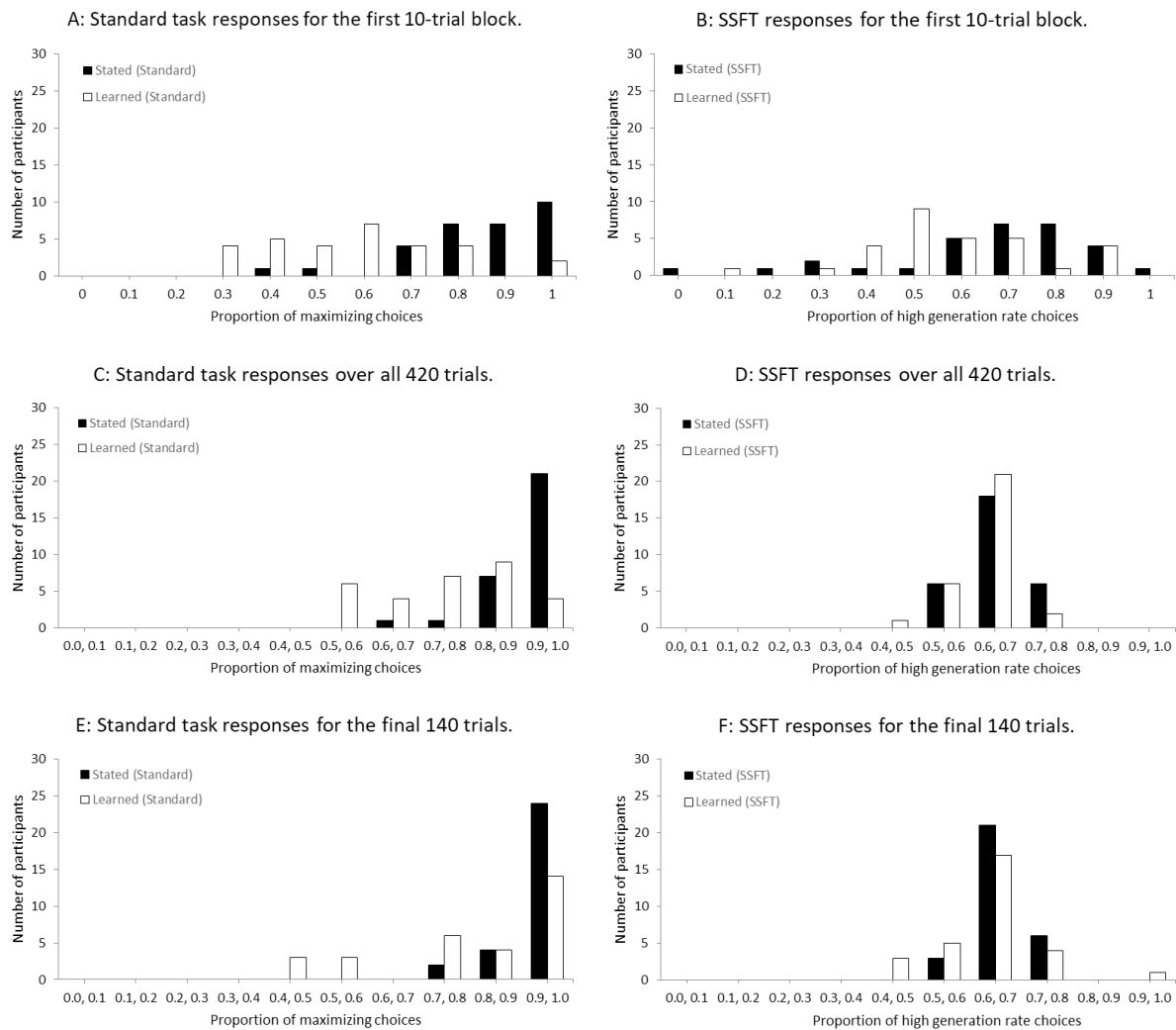
## RESULTS

The proportions of selection of the high resource generation rate option for each 10-trial block are shown in Figure 3, for each task and probability condition. Distributions of individual proportions of high generation rate choices are shown in Figure 4, Panels B, D, and F. The data for the SSFT were entered into a  $2 \times 42$  mixed model analysis of variance (ANOVA) with probability condition (stated vs. learned) as a between-subjects factor and block (1 to 42) as a within-subject factor. This found no significant main effect of block,  $F(16.64, 965.38) = 1.09$ ,  $MS_{\epsilon} = .08$ ,  $p = .36$ ,  $\eta_p^2 = .02$ ,<sup>1</sup> and no significant linear effect of block,  $F(1, 58) < 1.0$ , indicating that participants did not significantly increase their proportions of selections of the maximizing alternative as the experiment progressed. There was also no main effect of probability condition,  $F(1, 58) = 3.09$ ,  $MS_{\epsilon} = .17$ ,  $p = .08$ ,  $\eta_p^2 = .05$ , suggesting that participants who had been explicitly informed of the probabilistic accumulation rates did not allocate significantly more or fewer responses to the patch with the higher accumulation rate than those who had to learn these from experience alone. There was neither a significant interaction,  $F(16.64, 965.38) < 1.0$ ,<sup>1</sup> nor a linear interaction,  $F(1, 58) < 1.0$ , between block and probability condition.

Averaging steady state proportions of high accumulation rate choices on the SSFT over the final third of the task revealed values of .66 for the stated probabilities condition and .63 for learned probabilities condition with 95% CIs of .02 and .04, respectively (see Table 3). These indicate that there was no significant difference between the two groups, but average behaviour within each group significantly undermatched the actual ratio of reward generation probabilities (.7). Note,

**FIGURE 3.**

Proportions of selection of the high probability and high resource generation rate options. Shown for each block, probability, and task condition.



**FIGURE 4.** Histograms showing the number of participants who made different proportions of maximizing and high generation rate choices. Panels A and B show responses for only the first 10-trial block, Panels C and D - over the entire task, and Panels E and F - for the final third of the task. Panels A, C, and E show the standard condition and Panels B, D, and F - the SSFT condition.

**TABLE 3.** Steady State Proportions of Maximizing and High Resource Generation Rate Choices for the Standard Task and the Simple Simulated Foraging Task, Respectively.

	Task			
	Standard		SSFT	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
<b>Stated</b>	.944	.021	.659	.011
<b>Learned</b>	.805	.032	.631	.019

Note. SSFT = simple simulated foraging task.

however, that choices did not significantly undermatch the ratio of a locally optimal choice diversification strategy (2:1).

### Subjective Estimates of Resource Generation Rates

Subjective estimates of the relative resource generation rates of the two options were calculated by determining the proportion of total

estimated generation rate accounted for by the high accumulation rate option ( $grHest / [grHest + grLest]$ ). Group averages are shown in Table 4. Average reward probability estimates from the standard condition of the preceding study are also shown for comparison purposes. Of participants who completed the SSFT, an independent samples t-test revealed no significant difference between the relative estimated accumulation rates of those within the learned and stated probability conditions,  $t(58) = .32, p = .75$ . These estimates also did not significantly differ within any condition from the actual ratio of resource generation rates.

We also assessed the association between perceived ratio of SSFT resource generation rates and overall proportions of high generation rate choices. This revealed no significant overall effect,  $r = .16, p = .22, n = 60$ , neither within the stated probability,  $r = -.05, p = .79, n = 30$ , nor learned probability conditions,  $r = .29, p = .13, n = 30$ .

**TABLE 4.**  
Subjective Estimates of Outcome Probabilities and Relative  
Resource Generation Rates, by Probability Condition and Task.

	Task			
	Standard		SSFT	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
<b>Stated</b>	.685	.010	.680	.012
<b>Learned</b>	.727	.020	.674	.017

Note. SSFT = simple simulated foraging task.

## Heuristics and Apophenia

Self-reported use of strategy and belief that the outcome sequences contained patterns are shown in Tables 5 and 6. On the SSFT, these also did not differ between probability conditions,  $t(58) = .90$ ,  $p = .37$ ,  $t(58) = -.92$ ,  $p = .36$ .

## Experienced Discriminability

When reward contingencies are fixed, so is outcome discriminability – this being the difference in experienced proportion of rewarded trials for each option. However, on the SSFT, experienced outcome discriminability depends upon choices, and therefore varies between participants. Here, we conducted an additional analysis to assess the potential effect of experienced outcome discriminability on choices.

In order to avoid direct dependence of discriminability on choices, we correlated the level of outcome discriminability experienced over the initial 280 trials of the SSFT against proportions of high generation rate choices over the subsequent 140 trials. This revealed a significant overall negative association,  $r = -.42$ ,  $p < .001$ ,  $n = 60$ . This effect was found to be stronger in the learned probability,  $r = -.66$ ,  $p < .001$ ,  $n = 30$ , than the stated probability condition, where it was nonsignificant,  $r = -.25$ ,  $p = .19$ ,  $n = 30$ .

We also examined the effect of experienced outcome discriminability over the entire task on participants' estimated ratio of reward generation rates. This identified no relationship overall,  $r = .00$ ,  $p = .98$ ,  $n = 60$ , neither within the stated probability,  $r = .11$ ,  $p = .58$ ,  $n = 30$  nor learned probability conditions,  $r = -.06$ ,  $p = .74$ ,  $n = 30$ .

## Comparisons With a Standard Probability Matching Task

We also compared the data from the SSFT directly with corresponding behavioural data collected on a standard fixed-probability task. These data were obtained in the preceding study (Ellerby & Tunney, 2017). There are two principal differences between the SSFT and the standard task. The first of these is superficial. The SSFT was explicitly described as a foraging task, with choices made between two patches of land, whereas the standard task was framed as a probability task, with predictions made between which of two light bulbs would switch on. The second, more substantive difference is between the underlying reward schedules. On the SSFT, unharvested rewards remained and accumulated over time, while on the standard task they did not—reward contingencies were instead both fixed and IID. Otherwise, the two tasks were equivalent. Both were iterated two-alternative forced-choices, with the aim of maximizing total reward over a total of 420

trials. All stimulus durations, intertrial and interstimulus intervals also remained the same.

Selection proportions of the higher probability option on the standard fixed probability task are shown in Figure 3, alongside the selection proportions of the higher resource generation rate option on the SSFT. Distributions of individual proportions of maximizing choices are shown in Figure 4, Panels A, C, and E. These data were entered into  $2 \times 2 \times 42$  mixed model ANOVA with both probability condition (stated vs. learned) and task (standard vs. SSFT) as between-subjects factors and with block (1 to 42) as a within-subject factor. This revealed both a significant main effect of block,  $F(20.54, 2382.53) = 3.81$ ,  $MS_e = .05$ ,  $p < .001$ ,  $\eta_p^2 = .03$ ,<sup>1</sup> and a significant linear effect of block,  $F(1, 116) = 30.16$ ,  $MS_e = .09$ ,  $p < .001$ ,  $\eta_p^2 = .21$ , indicating that participants overall learned to allocate a significantly higher proportion of choices to the higher probability options (of either reward or resource accumulation) as the experiment progressed. There were also significant main effects of both probability condition,  $F(1, 116) = 35.66$ ,  $MS_e = .35$ ,  $p < .001$ ,  $\eta_p^2 = .24$ , and task,  $F(1, 116) = 131.22$ ,  $MS_e = .35$ ,  $p < .001$ ,  $\eta_p^2 = .53$ . These indicate that, overall, participants who had received explicit information regarding the probability rates chose the high probability option more often than those who had to learn this information for themselves, and that participants in the standard task group allocated significantly more choices to the maximizing option than participants in the SSFT group allocated to the high accumulation rate option.

This analysis also revealed theoretically informative interaction effects. First, there was a significant interaction between block and task,  $F(20.54, 2382.53) = 3.36$ ,  $MS_e = .05$ ,  $p < .001$ ,  $\eta_p^2 = .03$ ,<sup>1</sup> as well as a significant linear interaction,  $F(1, 116) = 20.96$ ,  $MS_e = .09$ ,  $p < .001$ ,  $\eta_p^2 = .15$ , representing significantly greater learning on the standard task than the SSFT. Second, there was also a significant interaction between probability condition and task,  $F(1, 116) = 18.04$ ,  $MS_e = .35$ ,  $p < .001$ ,  $\eta_p^2 = .14$ , with the effect of stated reward probabilities on participants' choices being significantly greater on the standard than the foraging task. When viewing these interaction effects in conjunction with the findings of the preceding analysis, conducted upon behaviour solely within the SSFT group, it is evident that the main effects of both block and probability condition were driven by differences within the standard condition. There was neither a significant interaction between block and probability condition,  $F(20.54, 2382.53) = 1.21$ ,  $MS_e = .05$ ,  $p = .23$ ,  $\eta_p^2 = .01$ ,<sup>1</sup> nor a linear interaction,  $F(1, 116) = 2.80$ ,  $MS_e = .09$ ,  $p = .10$ ,  $\eta_p^2 = .02$ . There was also neither a three-way interaction,  $F(20.54, 2382.53) < 1.0$ ,<sup>1</sup> nor a reliable three-way linear interaction,  $F(1, 116) = 2.77$ ,  $MS_e = .09$ ,  $p = .10$ ,  $\eta_p^2 = .02$ .

## Subjective Estimates of Reward Probabilities

Subjective estimates of outcome probabilities and relative resource accumulation rates are shown in Table 4. A  $2 \times 2$  ANOVA was conducted with probability condition (stated vs. learned) and task (standard vs. SSFT) as between-subjects factors. There was neither a significant effect of probability condition,  $F(1, 116) = 1.34$ ,  $MS_e = .01$ ,  $p = .25$ ,



$\eta_p^2 = .01$ , nor task,  $F(1, 116) = 3.69$ ,  $MS_e = .01$ ,  $p = .06$ ,  $\eta_p^2 = .03$ , nor a reliable interaction between the two,  $F(1, 116) = 2.54$ ,  $MS_e = .01$ ,  $p = .11$ ,  $\eta_p^2 = .02$ .

A significant positive association between estimated reward contingencies and maximizing choices was found within the learned probability condition of the standard task,  $r = .52$ ,  $p < .01$ ,  $n = 30$ . However, no significant effect was evident either within the stated probability condition,  $r = -.17$ ,  $p = .37$ ,  $n = 30$ , or overall,  $r = .14$ ,  $p = .30$ ,  $n = 60$ .

## Heuristics

Self-reported use of strategy vs intuition (see Table 5) was entered into a  $2 \times 2$  ANOVA with probability condition and task as between-subjects factors. A main effect of probability condition,  $F(1, 116) = 5.19$ ,  $MS_e = 1.00$ ,  $p = .03$ ,  $\eta_p^2 = .04$ , showed that participants in the stated probability condition reported greater strategy use than those in the learned probability condition. There was no effect of task,  $F(1, 116) < 1.0$ , and no significant interaction between the two,  $F(1, 116) = 1.01$ ,  $MS_e = 1.00$ ,  $p = .32$ ,  $\eta_p^2 = .01$ .

Bartlett's test indicated no evidence for either heterogeneity of variance or departure from normality in reported strategy use between the two tasks,  $\chi^2(1) = .04$ ,  $p = .84$ .

## Apophenia

Self-reported beliefs that the outcome sequences contained a pattern (see Table 6) were entered into another  $2 \times 2$  ANOVA. This revealed a significant main effect of probability condition,  $F(1, 116) = 4.20$ ,  $MS_e = 1.45$ ,  $p = .04$ ,  $\eta_p^2 = .04$ , indicating that participants in the learned probability condition reported significantly higher levels of apophenia than those in the stated probability condition. A main effect of task,  $F(1, 116) = 18.74$ ,  $MS_e = 1.45$ ,  $p < .001$ ,  $\eta_p^2 = .14$ , was also found, indicating that participants that had completed the SSFT reported significantly higher levels of belief that there was a pattern in the sequence than those who had completed the standard task. There was no interaction between the two factors,  $F(1, 116) < 1.0$ .

**TABLE 5.**  
Subjective Reports of Strategy Use by Condition and Task.

	Task			
	Standard		SSFT	
	M	SE	M	SE
<b>Stated</b>	3.800	0.176	3.467	0.205
<b>Learned</b>	3.200	0.188	3.233	0.157

Note. Scores ranged from 1 = pure intuition to 5 = pure strategy. SSFT = simple simulated foraging task.

**TABLE 6.**  
Subjective Reports of Belief That Outcome Sequences Contained a Pattern, by Condition and Task

	Task			
	Standard		SSFT	
	M	SE	M	SE
<b>Stated</b>	2.000	0.209	3.133	0.227
<b>Learned</b>	2.633	0.256	3.400	0.177

Note. Scores ranged from 1 = strongly disagree to 5 = strongly agree. SSFT = simple simulated foraging task.

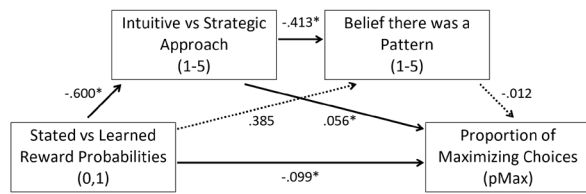
## Mediation Analysis

Independent mediation analyses were conducted within both the standard and SSFT conditions, using the PROCESS macro for SPSS (Hayes, 2012, 2013). The primary purpose of these analyses was to compare and contrast the relative effects of the probability condition manipulations between the two tasks, and, in particular, self-reported intuition versus strategy use and apophenia on choice behaviour. For this analysis, proportions of maximizing choices were calculated for the final third of the tasks. This was done with the aim of accounting for learning effects within the standard condition and within the SSFT group for consistency, despite no evidence of significant learning on this task. Mediation Model 6 was used in order to maintain consistency with the preceding study (Ellerby & Tunney, 2017). This model included a single independent variable and two mediator variables, of which the former (intuition vs. strategy) was allowed to influence the latter (pattern belief). This model reflects the conception of intuition versus strategy as a more general approach which may be taken from the beginning of the task, whereas apophenia depends upon experience, and so might be influenced by whether a more or less strategic approach was taken. Effects were calculated for each of 10,000 bootstrapped samples. Model results are presented in Figures 5 and 6 (Variable coding is as follows: Stated = 0 Learned = 1, Intuition-Strategy = 1-5, Pattern belief = 1-5, Proportion of maximizing choices = 0-1).

For participants who completed the standard task, the direct effect of probability condition on strategy use was found to be significant,  $t(58) = -2.33$ ,  $p = .02$ , with this predictor accounting for 9% of the sample variance ( $R^2 = .09$ ). The direct effect of probability condition on apophenia was found to be nonsignificant,  $t(57) = 1.17$ ,  $p = .25$ , while the direct effect of strategy use on apophenia was significant,  $t(57) = -2.57$ ,  $p = .01$ , with these two predictors accounting for 16% of the sample variance ( $R^2 = .16$ ). The direct effects of probability condition,  $t(56) = -3.05$ ,  $p = .004$ , and strategy use,  $t(56) = 3.37$ ,  $p = .001$ , on proportion of maximizing choices were each found to be significant, whereas the direct the effect of apophenia was not,  $t(56) = -.90$ ,  $p = .37$ . These three predictors accounted for 39% of the sample variance ( $R^2 = .39$ ).

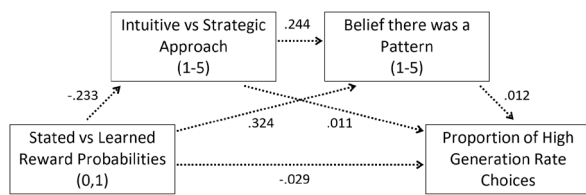
The unstandardized indirect effects of probability condition on the proportion of maximizing choices were as follows: Through intuition versus strategy use,  $(-.600)(.056) = -.034$ , 95% CIs  $[-.074, -.008]$ , significant at  $\alpha = .05$ . Through apophenia,  $(.385)(-.012) = -.005$ , 95% CIs  $[-.034, .005]$ , nonsignificant at  $\alpha = .05$ . Through first strategy use then apophenia,  $(-.600)(-.413)(-.012) = -.003$ , 95% CIs  $[-.019, .003]$ , nonsignificant at  $\alpha = .05$ .

For participants who completed the SSFT, all effects were found to be nonsignificant. The direct effect of probability condition on strategy use,  $t(58) = -.23$ ,  $p = .37$ , accounted for 1% of the sample variance ( $R^2 = .01$ ). The direct effects of probability condition,  $t(57) = 1.13$ ,  $p = .26$ , and strategy use,  $t(57) = 1.70$ ,  $p = .10$ , on apophenia together accounted for 6% of the sample variance ( $R^2 = .06$ ). The direct effects of probability condition,  $t(56) = -1.30$ ,  $p = .20$ , strategy use,  $t(56) = .98$ ,  $p = .33$ , and apophenia,  $t(56) = 1.15$ ,  $p = .26$ , on proportions of high accumulation rate choices together accounted for 8% of the sample variance ( $R^2 = .08$ ).



**FIGURE 5.**

Mediation model for the standard task condition. Unstandardized regression coefficients are shown with an asterisk if significant. Solid and dotted lines indicate significant and nonsignificant effects, respectively.  $\alpha = .05$ .



**FIGURE 6.**

Mediation model for the simple simulated foraging task condition. Unstandardized regression coefficients are shown with an asterisk if significant. Solid and dotted lines indicate significant and nonsignificant effects, respectively.  $\alpha = .05$ .

The unstandardized indirect effects of probability condition on the proportion of high accumulation rate choices were as follows: Through intuition versus strategy use,  $(-.233)(.011) = -.003$ , 95% CIs  $[-.025, .025]$ . Through apophenia,  $(.324)(.012) = .004$ , 95% CIs  $[-.002, .023]$ . Through first strategy use then apophenia,  $(-.233)(.244)(.012) = -.001$ , 95% CIs  $[-.009, .0003]$ . All indirect effects were also nonsignificant at  $\alpha = .05$ .

### Win-Stay Lose-Shift

The proportion of choices that were consistent with a win-stay lose-shift (WSLS) strategy was also assessed for participants both within the SSFT and the standard task conditions (see Table 7).

Overall proportions of WSLS-consistent choices were entered into a  $2 \times 2$  ANOVA, with probability condition (stated vs. learned) and task (standard vs. SSFT) as between-subjects factors. This found neither a main effect of probability condition,  $F(1, 116) < 1.0$ , nor of task,  $F(1, 116) = 1.42$ ,  $MS_e = .01$ ,  $p = .24$ ,  $\eta_p^2 = .01$ . However, a significant two-way interaction was found,  $F(1, 116) = 6.20$ ,  $MS_e = .01$ ,  $p = .01$ ,  $\eta_p^2 = .05$ . This reflects the finding that, on the SSFT, those in the learned resource generation rate condition made more WSLS-consistent

**TABLE 7.**

Proportions of Win-Stay Lose-Shift-Consistent Choices, Shown for Each Task and Probability Condition

	Task			
	Standard		SSFT	
	M	SE	M	SE
Stated	.694	.011	.627	.022
Learned	.661	.017	.685	.023

choices than those in the stated condition, whereas on the standard task, the opposite was true.

There is an important caveat when interpreting these comparisons. Expected baseline proportions of WSLS-consistent choices will differ between tasks, whether or not WSLS was actively pursued as a strategy. Similarly, the expected proportion of WSLS-consistent choices within the standard task will systematically increase in line with the proportion of maximizing choices. Here, for example, maximizing will entail a proportion of .7 WSLS-consistent choices (all being win-stay) and WSLS-independent matching—only .58. On a fixed-probability task, it is therefore possible to control for between-subjects variance in expected WSLS choices derived from each participants' proportion of maximizing choices (cf. Gaissmaier & Schooler, 2008; Ellerby & Tunney, 2017). However, calculating equivalent values according to overall choice proportions is not feasible for the SSFT due to dependencies between trials. That is, equal proportions of high generation rate choices may lead to differing proportions of WSLS-consistent choices, depending upon the sequence in which they were made. In light of this, we chose to compare uncorrected proportions of WSLS-consistent choices from both tasks.

The interaction effect observed here is likely partially explained by the aforementioned mechanism; lower proportions of WSLS-consistent choices on the learned probability condition of the standard task corresponded with significantly lower proportions of maximizing choices observed in the same condition. However, the inverse effect found between the SSFT conditions cannot be similarly explained. This might instead reflect a greater number of participants in the learned probability SSFT condition adopting a genuine WSLS-driven approach to their choices.

Further insight into choice behaviour could be afforded by investigating the proportions of win-shift and lose-shift choices independently. Here, we report how often participants were found to switch to the alternative option following a win or a loss on the preceding trial (see Tables 9 and 10). Expected baseline amounts of win- and lose-shift choices were also calculated for different choice strategies from each of the two tasks. These are shown in Table 8 for comparison purposes.

Win-shift choices as a proportion of all choices following a win, were entered into a  $2 \times 2$  ANOVA, with probability condition (stated vs. learned) and task (standard vs. SSFT) as between-subjects factors. This revealed no main effect of probability condition,  $F(1, 116) < 1.0$ . However, a significant main effect of task was found,  $F(1, 116) = 47.39$ ,  $MS_e = .02$ ,  $p < .001$ ,  $\eta_p^2 = .29$ , representing significantly higher proportions of win-shift choices on the SSFT as compared to the standard task. A significant two-way interaction was also found,  $F(1, 116) = 19.42$ ,  $MS_e = .02$ ,  $p < .001$ ,  $\eta_p^2 = .14$ , reflecting the finding that participants on the standard task were more likely to shift following a win in the learned probability condition, whereas participants on the SSFT were more likely to do so in the stated probability condition.

Lose-shift choices as a proportion of all choices following a loss, were also entered into a  $2 \times 2$  ANOVA, with probability condition (stated vs. learned) and task (standard vs. SSFT) as between-subjects factors. This revealed significant main effects of both probability con-

dition,  $F(1, 116) = 13.70$ ,  $MS_e = .03$ ,  $p < .001$ ,  $\eta_p^2 = .11$ , and of task,  $F(1, 116) = 93.08$ ,  $MS_e = .03$ ,  $p < .001$ ,  $\eta_p^2 = .45$ . These represent significantly higher proportions of lose-shift choices both when probabilities were learned instead of stated and on the SSFT relative to the standard task. A significant two-way interaction was also found,  $F(1, 116) = 15.63$ ,  $MS_e = .03$ ,  $p < .001$ ,  $\eta_p^2 = .12$ . This reflects the finding that participants on the standard task were substantially more likely to shift following a loss in the learned probability than in the stated probability condition, whereas participants on the SSFT were approximately equally likely to switch following a loss in each probability condition.

## DISCUSSION

Following millions of years of natural selection, it can be expected that the most highly adaptive rules should have emerged as instinctive, or default, behavioural responses. While heuristic accounts provide a superficial explanation of nonnormative choice diversification, they do not address the question of precisely how such a seemingly suboptimal tendency as probability matching could have become established and remain to this day as a dominant intuitive strategy. The current study was motivated by this question. One intriguing proposal is that choice diversification, including the case of probability matching, might reflect the expression of a generally optimal decision criterion for the purposes of time allocation in foraging situations, which has become established over an extended period of evolutionary ancestry. In humans, this rule is therefore expressed predominantly when the actor behaves in accordance with their intuitions. Under this interpretation, humans do, in fact, have what has been generally the most adaptive response pattern readily and naturally available. This leads to the dominance of this choice strategy when alternatives lack availability, as discussed by Koehler and James (2010). Therefore, the expression of probability matching behaviour when confronted with the standard 2AFC

**TABLE 8.**  
Expected Proportion of Shift Choices for Different Response Patterns, Given an Immediately Preceding Win or Loss

	Shift after win	Shift after loss
Probability maximizing (standard)	.000	.000
Probability matching (standard)	.362	.500
Win-stay lose-shift (either task)	.000	1.000
Optimal diversification (SSFT)	.620	.777

Note. Optimal diversification represents the selection of high and low generation rate options in a repeating 2:1 ratio. SSFT = simple simulated foraging task.

**TABLE 9.**  
Observed Proportion of Choice Shifts Following a Win, for Each Task and Condition

	Task			
	Standard		SSFT	
	M	SE	M	SE
Stated	.071	.013	.357	.030
Learned	.207	.028	.270	.026

**TABLE 10.**  
Observed Proportion of Choice Shifts Following a Loss, for Each Task and Condition

	Task			
	Standard		SSFT	
	M	SE	M	SE
Stated	.199	.028	.595	.032
Learned	.422	.028	.587	.027

task in which reward contingencies are fixed and IID, might represent an overextension of what has been an otherwise normative behaviour.

This study investigated human behaviour on a simple simulated foraging task (SSFT), which was designed to incorporate the factor of resource accumulation over time into the standard 2AFC probability matching preparation. Behaviour on this task was then compared with that on the standard version, obtained in a preceding study (Ellerby & Tunney, 2017). Differences were also examined in how a number of other factors relate to choice behaviour and each other across the two tasks. When viewing behaviour on the SSFT alone, no significant learning effect was found, coupled with no effect of explicitly stating the reward accumulation rates. For participants who completed the SSFT, stating reward contingencies also had no effect on either subjective estimates of accumulation rates, the use of strategy over intuition, or levels of apophenia. A tendency towards group-level undermatching was evident in relation to the ratio of resource generation rates, though not in relation to a locally optimal 2:1 ratio of choice allocations across the task.

Direct comparison of these results with those on a standard, fixed-probability task revealed a stark contrast between effects. Simple main effects indicated substantially lower proportions of high resource generation rate choices on the SSFT than proportions of maximizing choices on the standard fixed probabilities task. Moreover, significant interactions were found involving task and both learning and the effects of stated reward probabilities. These findings represent robust effects of each of these factors within the standard condition, paired with their absence on the SSFT. Furthermore, a mediation analysis revealed that the significant effect of intuition versus strategy use on choice behaviour, which was present on the standard task, disappeared on the SSFT. This was true despite no evidence being found for either a significant difference in overall levels of strategy use or for heterogeneity of variance in levels of strategy use between the two tasks. In line with these findings, the indirect effect of stating reward probabilities on choices through facilitation of strategy use was also absent on the SSFT.

This evidence indicates that while on the standard task, participants behave differently depending upon whether they adopt a predominantly intuitive or strategic approach, on the SSFT, the degree to which a more deliberative strategy is applied no longer has any substantive influence on choice behaviour. This finding is consistent with heuristic-based accounts which propose that probability matching is prepotent when participants act intuitively. As choice diversification, including probability matching, can represent an optimal choice strategy on tasks in which unharvested rewards can persist and accumulate, participants acting either intuitively or strategically can converge upon equivalent response proportions when completing the SSFT. By contrast, on the standard task in which prob-

abilities are fixed, taking an intuitive versus strategic approach will lead to a systematic divergence in choice behaviour.

Incidentally, levels of apophenia were found to be significantly higher on the SSFT than on the standard task as well as in the learned than stated probability conditions. This makes sense, as on the SSFT, rewards were contingent in some manner upon previous choices, and therefore not entirely IID. Nonetheless, as on the standard task, levels of apophenia were found to have no substantive effect on choice behaviour.

The findings of another recent publication (Schulze et al., 2017) conducted with a similar premise to the current study can provide a useful comparative measure. When comparing reward-hold with no hold conditions, Schulze et al. found neither a significant main effect of reward retention on average choice proportions nor any reliable group-level interaction effects. Nevertheless, further probing of distributions of individual participant level choices suggested that substantial underlying differences were, in fact, present between the two reward schedules. In the reward-hold condition, the modal peak response was near optimal diversification of choices. By contrast, individual choice proportions demonstrated a relatively platykurtic distribution within the no hold condition. Here, while the modal response was maximizing, a tail of participants extended substantially beyond probability matching, with a small subset of participants even reaching chance levels of responding. When considering the final block of the task, potentially effective diversification was 6.94 times more prevalent in the reward-hold than the no hold condition. By contrast, maximizing was 9.63 times more prevalent in the no hold condition. Therefore, the authors argue that their results provide clear evidence for qualitatively different response patterns between the two conditions, despite the lack of significant group-level differences.

The results of the present study are broadly consistent with these findings. However, they also indicated a clear group-level difference between choices on the SSFT and the standard task. There were two key differences between the reward schedules of these two studies, which merit clear reiteration. Together, these might explain much, if not all, of the apparent discrepancy in observed group-level effects. First, in the Schulze et al. (2017) study's reward-hold condition, a single reward could persist over time in the unchosen option, but multiple rewards were not permitted to accumulate. On the SSFT, the rewards not only remained in an underexploited option but were also allowed to continually accrue until they were collected. This entails that while the SSFT and the reward-hold condition are analogous in principle, with some degree of diversification becoming advantageous on each task, the specific levels of diversification that are optimal are not the same. In fact, while probability matching is an optimal response-pattern on the SSFT and diversification is necessary to maximise rewards over the task, it is also possible to do so almost equivalently through a relatively broad range of choice diversification patterns. For example, 1:1, 2:1, and 3:1 ratios of option selection can each approximate optimal outcomes over the task, even though they do not all optimise likelihood of reward on each individual trial.

Second, Schulze et al. (2017) controlled for perceived discriminability over the two conditions. This entailed substantially closer mean reward contingencies in their no hold condition (.76 vs .63-.65), which also fluctuated during the task, as compared with the fixed and substantially more

easily discriminable contingencies (.70 vs .30) used in the standard condition of the present study, against which behaviour on the SSFT was compared. We conducted analyses to assess the extent to which differences in experienced outcome discriminability influenced choices on the SSFT. These revealed that higher experienced discriminability over the first two-thirds of the task was associated with more, rather than less, diversification of choices over the final third. We consider the most likely explanation for this possibly counter-intuitive finding to be that diversification of choices on the SSFT directly entails higher outcome discriminability, and between-subjects differences in the tendency to diversify may have remained relatively stable across the task. Nonetheless, this finding fails to support the hypothesis that experienced discriminability is an important driving factor behind diversification on the SSFT, although it is still likely to be so on a fixed-probability task.

Relatedly, we also examined subjective estimates of the ratio of resource generation rates between the two options, both on the SSFT and the standard task. Here, the results showed that perceived SSFT reward generation rates were entirely unrelated to experienced outcome discriminability within both conditions. This indicates that participants must have derived their perceptions of reward generation in each option from more than the experienced outcome discriminability alone. We propose that this reflects a general understanding of the principle of reward accumulation in the unchosen option, even within the learned probability condition. In addition, although we found no effect of subjective estimates of SSFT resource generation rates on choices, equivalent analyses for the standard task revealed a significant relationship within the learned probability condition. This is not surprising, as this is the only task condition in which participants rely entirely upon their own experience to discern outcome contingencies, and these contingencies can be reliably determined from experienced outcomes alone.

Further analyses were conducted to explore overall proportions of WSLS-congruent choices over the two tasks, paired with a more specific examination of how often participants tended to switch their response immediately following rewarded or unrewarded trials. Here, participants were found to be substantially more likely to shift following a loss than a win, as would be predicted by any reinforcement learning-based model. Nevertheless, recorded proportions of win-shift choices were clearly nonzero in all task conditions, despite a maximizing strategy on the standard task entailing no switching between options at all. However, significantly greater proportions of win-shift choices were observed on the SSFT than the standard task. This difference was particularly exaggerated between the two stated probability conditions, in which respondents were five times as likely to switch choices following a win on the SSFT than they were on the standard task. This finding demonstrates that respondents on the SSFT diversified in a manner that was not driven by WSLS, at least not exclusively so. It instead suggests a substantive degree of anticipatory switching, which was likely driven by varying degrees of cognizance of resource accumulation in the unchosen option. This interpretation is supported by the finding that on the SSFT win-shift choices were significantly more common in the stated probability condition, in which knowledge of the process by which resources accumulate was made explicitly available from the outset.

Significant interaction terms were evident between task and condition. This was true for overall WSLs-consistent choices and, inversely so, for both proportions of win- and lose-shift choices when considered independently. The former two of these terms are congruent and likely reflect opposing effects of stated reward contingencies between the two tasks. As discussed in the preceding paragraph, providing respondents with prior knowledge of the SSFT reward schedule encouraged the use of win-shift choices, reflecting periodic exploration of the more rarely exploited option in anticipation of resources having accumulated therein. By contrast, on the standard task stating reward probabilities led to a higher incidence of win-stay choices through engendering probability maximizing. The pattern of lose-shift choice proportions runs counter to that observed in total WSLs-consistent choices, which serves to illustrate that these overall effects were driven primarily by differences in choices following rewarded rather than unrewarded trials.

On the whole, the results of present study are consistent with the theory that choice diversification, and potentially probability matching, represent an overextension of an instinctive response which has historically represented an evolutionarily adaptive foraging strategy. However, it would be inappropriate to claim that the current findings alone constitute conclusive evidence in support of this account. There are a range of potential explanations for why choice diversification might be an intuitive response on the standard probability matching task, which we acknowledge that the findings of the present study do not rule out.

First, the presence of competitors provides a valid alternative mechanism for establishing matching as an optimal response. Competition for limited resources has likely provided an equally long-standing and consistently important selection pressure as the potential for resources to accumulate and deplete over time. Several previous studies, in relation to both the IFD and its possible application to probability matching, have focused on competition rather than resource accumulation as the primary basis for this (cf. Kraft & Baum, 2001; Kraft et al., 2002; Madden et al., 2002; Sokolowski et al., 1999; Schulze et al., 2015; Weber, 1998). Behavioural tendencies that were originally established in response to this factor could drive instinctive diversification of choices in noncompetitive environments.

Alternatively, rather than probability matching being manifest as an inherently instinctive response, the tendency to diversify choices might be an overlearned behaviour, developing from common environmental principles that are encountered in everyday life and which can establish heuristic responses. If, for instance, outcomes are not assumed to be fixed, a characteristic that is relatively uncommon in real-life situations, then the periodic exploration of alternatives is an important behaviour for the purposes of maintaining an accurate and up to date representation of the reward value of different options. This is crucial in permitting the behavioural flexibility to more optimally track potential shifts in reward contingencies over time. A study conducted by Green, Benson, Kersten, and Schrater (2010) offers an overt empirical illustration of this point. This study revealed that diversification naturally results from model-based learning when there is uncertainty regarding the fixedness of reward schedules. Here, an optimal Bayesian algorithm converged upon probability matching when it was initialized with an incorrect world-model, based upon the ecologically plausible assumptions that reward probabilities

were neither IID. nor coupled to those of their alternatives. In addition, the representativeness heuristic offers another example of how probability matching could result from the mistaken application of seemingly common sense logic. Irrespective of whether or not someone is aware that reward contingencies are fixed and IID, they are likely to assume that the best response pattern will generally approximate salient characteristics of the actual outcome sequence.

To summarise, in this study, we developed a simple simulated foraging task. This was done by making superficial modifications to a standard 2AFC probability matching task, along with modelling the persistence and accumulation of unharvested resources within the task reward schedule. Analyses contrasted observed behaviour between this and a standard fixed-probability task. Discrepancies were identified in both choice behaviour and the effects of learning, provision of reward probability information, and self-reported strategy use. These combined to indicate that the effect of deliberation over intuition in engendering maximizing behaviour is particular to the standard fixed probability task. When the potential for uncollected resources to accumulate over time was incorporated into the reward schedule, we no longer observed any effect of overall levels of strategy use on choice behaviour. Rather, a convergence of choice patterns was found across participants, with those acting intuitively now behaving statistically equivalently to those that take a more deliberative approach. Previous research has supported heuristic-based accounts of probability matching. These propose that diversification is a readily-available strategy, which can be expressed in situations where it will not result in maximum pay-offs whenever insufficient deliberation is applied. Although they do not rule out alternative explanations, the current findings are compatible with the further hypothesis that the ready-availability of the matching response might have become established due to having represented a historically adaptive foraging strategy. Probability matching might therefore represent an instinctive overextension of an otherwise normative foraging behaviour. However, further research is required in order to more conclusively determine the validity of this theory.

## FOOTNOTES

<sup>1</sup> Degrees of freedom were adjusted using the Greenhouse-Geisser correction for violations of sphericity.

## REFERENCES

- Behrend, E. R., & Bitterman, M. E. (1961). Probability-matching in the fish. *The American Journal of Psychology*, *74*, 542–551. doi: 10.2307/1419664
- Bullock, D. H., & Bitterman, M. E. (1962). Probability-matching in the pigeon. *The American Journal of Psychology*, *75*, 634–639. doi: 10.2307/1420288
- Brackbill, Y., Starr, R. H., & Kappy, M. S. (1962). Magnitude of reward and probability-learning. *Journal of Experimental Psychology*, *63*, 32–35. doi:10.1037/h0045051
- Davison, M., & McCarthy, D. (1988). *The matching law: A research review*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ellerby, Z. W., & Tunney, R. J. (2017). The effects of heuristics and apophenia on probabilistic choice. *Advances in Cognitive*

- Psychology*, 13, 280–295. doi: 10.5709/acp-0228-9
- Fretwell, S. D. (1972). *Populations in a seasonal environment*. Princeton, NJ: Princeton University Press.
- Fretwell, D. S., Lucas, H. L. (1970). On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica* 19, 16–32. doi: 10.1007/BF01601955
- Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, 109, 416–422. doi: 10.1016/j.cognition.2008.09.007
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Graf, V., Bullock, D. H., & Bitterman, M. E. (1964). Further experiments on probability-matching in the pigeon 1. *Journal of the Experimental Analysis of Behavior*, 7, 151–157. doi: 10.1901/jeab.1964.7-151
- Green, C. S., Benson, C., Kersten, D., & Schrater, P. (2010). Alterations in choice behavior by manipulations of world model. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 16401–16406. doi: 10.1073/pnas.1001709107
- Gigerenzer, G. (1996). Rationality: Why social context matters. In P. B. Baltes & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspectives on the social foundation of cognition* (pp. 319–346). Cambridge, UK: Cambridge University Press.
- Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. New York City, NY: Oxford University Press.
- Goldstone, R. L., & Ashpole, B. C. (2004). Human foraging behavior in a virtual environment. *Psychonomic Bulletin & Review*, 11, 508–514. doi: 10.3758/bf03196603
- Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. Retrieved from <http://www.afhayes.com/public/process2012.pdf>.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267–272. doi: 10.1901/jeab.1961.4-267
- Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior*, 24, 107–116. doi: 10.1901/jeab.1975.24-107
- Koehler, D. J., & James, G. (2010). Probability matching and strategy availability. *Memory & Cognition*, 38, 667–676. doi: 10.3758/mc.38.6.667
- Kogler, C., & Kuehberger, A. (2007). Dual process theories: A key for understanding the diversification bias? *Journal of Risk and Uncertainty*, 34, 145–154. doi: 10.1007/s11166-007-9008-7
- Kraft, J. R., & Baum, W. M. (2001). Group choice: The ideal free distribution of human social behavior. *Journal of the Experimental Analysis of Behavior*, 76, 21–42. doi: 10.1901/jeab.2001.76-21
- Kraft, J. R., Baum, W. M., & Burge, M. J. (2002). Group choice and individual choices: modeling human social behavior with the ideal free distribution. *Behavioural Processes*, 57, 227–240. doi: 10.1016/s0376-6357(02)00016-5
- Longo, N. (1964). Probability-learning and habit-reversal in the cockroach. *The American Journal of Psychology*, 77, 29–41. doi: 10.2307/1419269
- Madden, G. J., Peden, B. F., & Yamaguchi, T. (2002). Human group choice: Discrete-trial and free-operant tests of the ideal free distribution. *Journal of the Experimental Analysis of Behavior*, 78, 1–15. doi: 10.1901/jeab.2002.78-1
- Meyer, D. R. (1960). The effects of differential probabilities of reinforcement on discrimination learning by monkeys. *Journal of Comparative and Physiological Psychology*, 53, 173. doi: 10.1037/h0045852
- Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017
- Poling, A., Edwards, T. L., Weeden, M., & Foster, T. M. (2011). The matching law. *Psychological Record*, 61, 313–322. doi: 10.1007/BF03395762
- Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2015). Of matchers and maximizers: How competition shapes choice under risk and uncertainty. *Cognitive Psychology*, 78, 78–98. doi: 10.1016/j.cogpsych.2015.03.002
- Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2017). Hold it! The influence of lingering rewards on choice diversification and persistence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 1752–1767. doi: 10.1037/xlm0000407
- Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, 15, 233–250. doi: 10.1002/bdm.413
- Siegel, S., & Goldstein, D. A. (1959). Decision-making behavior in a 2-choice uncertain outcome situation. *Journal of Experimental Psychology*, 57, 37–42. doi: 10.1037/h0045959
- Skyrms, B. (1996) *The evolution of the social contract*. Cambridge, England: Cambridge University Press.
- Skyrms, B. (1997). Game theory, rationality and evolution. In M. L. Dalla Chiara, K. Doets, D. Mundici, & J. van Benthem (eds.), *Structures and norms in science* (pp. 73–85), Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Sokolowski, M. B. C., Tonneau, F., & Baque, E. F. I. (1999). The ideal free distribution in humans: An experimental test. *Psychonomic Bulletin & Review*, 6, 157–161. doi: 10.3758/bf03210824
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004, June). Matching behavior and the representation of value in the parietal cortex. *Science*, 304, 1782–1787. doi: 10.1126/science.1094765
- Unturbe, J., & Corominas, J. (2007). Probability matching involves rule-generating ability: A neuropsychological mechanism

- dealing with probabilities. *Neuropsychology*, *21*, 621–630. doi: 10.1037/0894-4105.21.5.621
- Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*, 101–118.
- Weber, T. P. (1998). News from the realm of the ideal free distribution. *Trends in Ecology & Evolution*, *13*, 89–90. doi: 10.1016/s0169-5347(97)01280-9
- West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition*, *31*, 243–251. doi: 10.3758/bf03194383
- Wilson, W. A., Oscar, M., & Bitterman, M. E. (1964). Visual probability-learning in the monkey. *Psychonomic Science*, *1*, 71–72. doi: 10.3758/BF03342795
- Wolford, G., Newman, S. E., Miller, M. B., & Wig, G. S. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology*, *58*, 221–228. doi: 10.1037/h0087446

RECEIVED 07.07.2018 | ACCEPTED 22.04.2019

## APPENDIX A

### TASK INSTRUCTIONS. SLIDE 1.

Welcome to the experiment. You will be presented with a series of choices. On each trial, a yellow and a blue signposted patch of land will be shown to either side of the screen. Your task is to choose to forage in one or the other on each trial. For each successful forage, 2p will be added to your total winnings. You can choose between these two foraging areas by pressing the q key for whichever option is shown to the left of the screen, or the p key for whichever option is shown to the right of the screen. Remember, it is the colour of the signpost in each area that is important. The side to which each colour patch is presented will vary randomly between trials and does not have any effect on likelihood of reward. The aim is to win as much money as possible throughout the task, by foraging successfully on as many trials as possible. Once you have read and understood the task instructions, press the space bar to continue. If you have any questions, please ask the experimenter.

## APPENDIX B

### TASK INSTRUCTIONS. ADDITIONAL REWARD GENERATION RATE INFORMATION (ONLY SHOWN IN STATED PROBABILITIES GROUP).

The two foraging areas will accumulate resources at different rates. On each trial, the probability of a new resource appearing in the blue signposted patch will be 0.49, while the probability of a new resource appearing in the yellow patch will be 0.21. This means that a new reward can be expected to become available in the blue patch, on average, in 49% of trials, and for the yellow patch, on average, in 21% of trials. Any resources that appear in the patch that you do not select will remain available, and accumulate over time, until they are taken. This means that one patch of land may contain more than one available reward. However, only a single reward may be taken on any one trial. Once you have read and understood this information, press the space bar to continue. If you have any questions, please ask the experimenter.

## APPENDIX C

### WORDING OF QUESTIONNAIRE ITEMS (WITH THE BLUE PATCH AS THE HIGH GENERATION RATE OPTION).

On every trial each patch of land had a certain probability of gaining an extra resource. Please estimate on what proportion of trials (%) you think that a new reward appeared in the blue and yellow signposted patches (over the entire task).

Blue Patch: %  
Yellow Patch: %

Overall, to what extent do you feel that you used your intuition to make your choices on the task, as opposed to any explicitly held plan or strategy?

(1 = pure intuition, 5 = pure strategy)

1 2 3 4 5

Please answer whether you agree or disagree with the following statements (1 = strongly disagree, 2 = slightly disagree, 3 = unsure, 4 = slightly agree, 5 = strongly agree)

There was a pattern in the reward sequence.

1 2 3 4 5