

Testing the limits of optimality: the effect of base rates in the Monty Hall dilemma

Walter T. Herbranson · Shanglun Wang

© Psychonomic Society, Inc. 2013

Abstract The Monty Hall dilemma is a probability puzzle in which a player tries to guess which of three doors conceals a desirable prize. After an initial selection, one of the nonchosen doors is opened, revealing that it is not a winner, and the player is given the choice of staying with the initial selection or switching to the other remaining door. Pigeons and humans were tested on two variants of the Monty Hall dilemma, in which one of the three doors had either a higher or a lower chance of containing the prize than did the other two options. The optimal strategy in both cases was to initially choose the lowest-probability door available and then switch away from it. Whereas pigeons learned to approximate the optimal strategy, humans failed to do so on both accounts: They did not show a preference for low-probability options, and they did not consistently switch. An analysis of performance over the course of training indicated that pigeons learned to perform a sequence of responses on each trial, and that sequence was one that yielded the highest possible rate of reinforcement. Humans, in contrast, continued to vary their responses throughout the experiment, possibly in search of a more complex strategy that would exceed the maximum possible win rate.

Keywords Choice · Monty Hall dilemma · Optimality · Pigeons · Probability learning · Sequence learning

A longstanding assumption in psychology and economics is that people make decisions that maximize gains and minimize losses (von Neumann & Morgenstern, 1944). More recently, however, behavioral science has yielded results that put in question such sweeping assumptions of optimality (Kahneman & Tversky,

1979). Optimal foraging theory (Stevens & Krebs, 1986) makes similar assumptions about the behavior of nonhuman animals, and again, recent research has revealed that actual behavior sometimes falls well short of optimal (Zentall, 2011). It is uncertain whether the situations in which humans fail to behave optimally will be the same as those that produce suboptimal behavior in other animals.

One fascinating example of suboptimal choice in humans is the “Monty Hall dilemma” (MHD). In the MHD, an individual is presented with a choice from among three doors, only one of which conceals a desirable prize. The two remaining doors have behind them unappealing gag prizes. After a door has been selected, the host, Monty Hall, opens one of the nonwinning doors and then offers the contestant a choice between staying with the initial selection and switching to the other remaining door. Most people choose to stay with their initial choice, a suboptimal decision because switching yields a greater chance of winning (Granberg & Brown, 1995).

One explanation for why switching is better is as follows: Given that a contestant makes the initial selection randomly, that selection will correspond to the prize location one-third of the time. In those cases, the prize is behind the contestant’s selected door, and so it follows that staying must win and switching must lose. In contrast, a contestant’s initial choice will *not* correspond to the prize location the remaining two-thirds of the time. In those cases, switching must win and staying must lose. Switching must win because the participant did not choose the prize, and the opened door did not reveal the prize. Consequently, the prize can only be located behind the other remaining door.

Whereas the game show scenario ordinarily described in the MHD involves a single trial, Granberg and Brown (1995) had participants complete 50 trials with feedback after each, in order to determine whether people could learn to switch on the basis of experience. Although their participants became more likely to switch, they quickly (within about 20 trials) reached a stable plateau that was still short of the optimal strategy of switching on every trial. The reasons for such

W. T. Herbranson · S. Wang
Whitman College, Walla Walla, WA, USA

W. T. Herbranson (✉)
Department of Psychology, Whitman College, 345 Boyer Avenue,
Walla Walla, WA 99362, USA
e-mail: herbrawt@whitman.edu

consistent failure to maximize the probability of winning remain unsettled, though several compelling possibilities have been suggested (see Gilovich, Medvec, & Chen, 1995; Tubau & Alonso, 2003).

It may be that the suboptimal tendency for humans to stay in the MHD does not extend to other animals. Herbranson and Schroeder (2010) showed that pigeons performed nearly optimally on a task based on the MHD. The three response keys (left, center, and right) in a standard operant chamber stood in for the three doors, and any winning trial was followed by access to mixed grain. At the beginning of a trial, all three keys were illuminated with white light, and one was designated as the prize location (but remained indistinguishable from the others). Following a single peck to any of the three keys, all keys were darkened for 1 s, and subsequently, two of the three keys were illuminated with green light. The nonilluminated (“opened”) key was selected randomly within the constraints of the MHD: It could not be the one that had been pecked, nor could it be the prize location. A second peck resulted either in access to mixed grain (if it was to the designated prize key) or the end of the trial with no reinforcement. Over the course of 30 days, all of their pigeons learned to switch on virtually every trial. Their pigeons also learned to stay on virtually every trial of a second experiment in which staying was the optimal response, indicating that their performance was based on experience with the task, rather than on an immediate understanding of the structure of the problem. Humans, meanwhile, completed parallel procedures and failed to adopt the best strategies, even with extensive training.

Though it had no influence on the outcome of the experiment, most of Herbranson and Schroeder’s (2010) pigeons eventually adopted a preferred key for their first response on most trials (specifically, 96 % of trials). Because the prize location was randomly determined on each trial, initial choices had no bearing on the chance of winning, so pigeons’ position biases were inconsequential. However, note that as a result of adopting a preferred starting key, pigeons usually received reinforcement immediately after pecking one of the other, nonpreferred keys in the second portion of a trial. They were much less likely to receive reinforcement immediately after pecking their initially preferred key (reinforcement could only be presented after the second peck of a trial, never before it). At first glance, this result is curious, in that at the beginning of a trial, when all three keys were illuminated, pigeons did not choose the key that was most strongly associated with reinforcement. This implies that pigeons learned something aside from a preference for a specific response key location. One possibility draws from previous research on response sequence learning (Herbranson & Stanton, 2011; Nissen & Bullemer, 1987): Pigeons might have learned to execute a specific sequence of two key pecks on each trial. This is a viable possibility, because on average, sequences that involved more than one key (i.e., switching sequences) would be reinforced more frequently than sequences that involved

only a single key (i.e., stay sequences). This possibility was the subject of our first experiment.

Experiment 1

As we mentioned previously, the first response in the standard MHD has no effect on the probability of winning. However, it can be made to be important if the prize locations are not equiprobable. In such a situation, the optimal strategy is to first choose the lowest-probability location and then switch away from it, regardless of which locations remain and which location is no longer available. Doing so results in a higher probability of winning. The precise probabilities associated with staying and switching can be computed in two ways, depending on the assumptions made about the decision process (see Granberg, 1999). If both decisions (the initial choice and whether to stay or switch) are made at the beginning of a trial, before the host opens one of the unselected doors (i.e., if one commits to switching or staying, regardless of which door is opened by the host), the probabilities of winning for switching and staying can be calculated before the host acts. Following Granberg, we refer to these as *unconditional probabilities*. The unconditional probability of winning for staying is equal to the probability (P) that the prize was initially placed behind that same, initially chosen door. The unconditional probability of winning for switching must be the remaining probability ($1 - P$), which is equal to the sum of the probabilities for the other two locations. Thus, in terms of the unconditional probability of winning, one can think of switching as trading the initially chosen door for *both* of the nonchosen doors. Even though one of those two does not contain the prize and will be opened, the remaining door will contain the prize, with probability $1 - P$.

If the second decision (whether or not to switch) is made *after* the host has revealed one of the nonwinning doors, one can recalculate the probabilities, taking into account new information revealed by the host’s decision of which door to open. Because the host’s decision is constrained in part by the prize location, it can reveal additional information about which door conceals the prize. This additional information determines the *conditional probability* of winning, given the host’s choice of which door to open, and the precise conditional probability can be computed using Bayes’s theorem. Equation 1 shows how to calculate the probability that the prize is in the initially chosen location (I), given that the host opened a particular door (o).

$$P(I|o) = \frac{P(o|I)P(I)}{P(o)}. \quad (1)$$

This is the *conditional probability* of winning by staying, and the probability of winning by switching is its complement,

$1 - P(I | o)$. If the prize is equally likely to be allocated to any of the three keys, then this conditional approach yields the same probabilities as the previously described, unconditional approach. However, if there is a bias, such that the three locations are not equally likely to contain the prize, then the conditional and unconditional probabilities will differ. In particular, if the host has the option of opening either of two doors with different probabilities of containing the prize, and opens the one with the higher probability, then the conditional probability of winning for switching decreases relative to the unconditional probability (note that this only applies if the initial choice was to the prize location). Conversely, if the lower-probability location is opened by the host, the conditional probability increases relative to the unconditional probability.

With this in mind, we developed two alternative procedures, summarized in Table 1, along with the standard MHD for comparison. Condition H features one key with a higher likelihood of containing the prize, whereas Condition L features one key with a lower likelihood of containing the prize. In both cases, the highest probability of winning is achieved by choosing the option associated with the lowest probability (or, in the case of Condition H, either of the two options associated with the lower probability) and then switching. Although the two approaches produce slightly different probabilities, they both dictate the same optimal strategy. In fact, the two are tightly linked, in that use of Bayes's theorem to compute the conditional probabilities requires knowledge of the unconditional probabilities.

Note that in both conditions, the multiple possible response sequences yield different reinforcement probabilities. Also note that switching is superior to staying, regardless of which key was initially selected, and regardless of whether one uses the conditional or unconditional probabilities. The optimal strategy in the two new versions (Condition H and Condition L) is to initially select a low-probability location and then to switch away from it, and by doing so one can achieve a higher likelihood of winning than in the classic MHD.

Whereas pigeons have been previously shown to perform optimally by switching in a standard MHD, the conditions outlined here are more complex, in that switching is by itself no longer sufficient for optimal performance. A pigeon could switch on every trial and still not maximize the payout if it started by pecking the wrong key. Similarly, a pigeon could consistently choose any one of the three initially available choices and not maximize payout if it failed to switch. In order to gain the greatest number of reinforcers, a pigeon must choose properly during both phases of a trial. In other words, it must execute a specific *sequence* of responses across the two choices that constitute a trial. Thus, there were two dependent variables, corresponding to the two elements of that sequence: Choosing the low-probability key during the first phase of a trial, and switching during the second phase of a trial.

Table 1 Variants of the Monty Hall dilemma task and associated probabilities of winning by switching

First Choice	Opened Door	Resulting Choice	Probability of Winning by Switching	
			Unconditional	Conditional
Condition H (.40, .30, .30)				
L	C	LR	.60	.60
L	R	LC	.60	.60
C	L	CR	.70	.67
C	R	CL	.70	.72
R	L	RC	.70	.67
R	C	RL	.70	.72
Condition L (.20, .40, .40)				
L	C	LR	.80	.80
L	R	LC	.80	.80
C	L	CR	.60	.67
C	R	CL	.60	.50
R	L	RC	.60	.67
R	C	RL	.60	.50
Classic MHD (.33, .33, .33)				
L	C	LR	.67	.67
L	R	LC	.67	.67
C	L	CR	.67	.67
C	R	CL	.67	.67
R	L	RC	.67	.67
R	C	RL	.67	.67

L, C, and R refer to the left, center, and right locations, respectively. Each probability presented in the table is the probability of winning by switching away from the initially selected key. Probabilities of winning by staying can be calculated by subtracting each from 1.0

Method

Animals Six white Carneaux pigeons (*Columba livia*) were obtained from Double-T Farm (Glenwood, IA). Each was maintained at approximately 80 %–85 % of its free-feeding weight (Poling, Nickel, & Alling, 1990) in a temperature-controlled colony room with a 14-h light/10-h dark cycle. Each pigeon was housed in an individual home cage with free access to water and grit. Some birds had previous experience in the same chambers on a serial response task (Herbranson & Stanton, 2011). All experimental sessions took place at approximately the same time, five days per week.

Apparatus Four BRS/LVE operant chambers were used. The front wall of each chamber contained three response keys and a feeder through which birds could gain access to mixed grain. Each chamber was interfaced to a personal computer that controlled all experimental events, recorded the data, and calculated daily statistics.

Procedure All birds were pretrained in sessions consisting successively of habituation, magazine training, and autoshaping (Brown & Jenkins, 1968) until consistent responding was achieved on each of the three response keys. The stimuli during autoshaping were white and green key lights.

A daily experimental session consisted of a series of individual trials (ten trials on the initial day, gradually increased to 100 trials over the first five days of the experiment). Each trial consisted of an initial choice response from among three options, a second choice response from among two options, food delivery if the second response was correct, and an intertrial interval (ITI). A house light was located directly above the center response key and remained on throughout each experimental session.

At the beginning of a trial, the computer selected one of the three keys as the prize location, according to the distribution that defined each condition. Once the prize location had been determined, all three keys were illuminated with white light. After a single peck at any key, all three keys were darkened for 1 s. During the 1-s delay, the computer pseudorandomly selected one of the three keys to deactivate for the remainder of the trial, with the constraint that the deactivated key could be neither the one previously selected as the prize location for that trial nor the key that the bird had already pecked on that trial. For the remainder of a trial, the deactivated key remained darkened, and any pecks on it were ignored. The two remaining keys (the key that the bird had just pecked and one other key) were then illuminated with green light. If a bird pecked the key that corresponded to the prize location, it was given approximately 2.5 s access to mixed grain (times varied from bird to bird, in order to maintain individual running weights) and, following a 5-s ITI, moved along to the next trial. If a bird's second response did not correspond to the prize location, no reinforcement was provided, and the procedure simply continued on to the ITI and the next trial.

For each trial, two results of interest emerged: the initial choice (from among three white keys) and whether the bird pecked the same key on its second choice (from among two green keys). If it did, the bird was said to have *stayed*. If the bird pecked a different key with its second choice than it did with its first, it was said to have *switched*.

Conditions Pigeons were arbitrarily assigned to one of two conditions. In Condition H, the chances of the three keys containing the prize were .4, .3, and .3. The odd (.4) key was in a different location for each of the three birds. In Condition L, the chances of the three keys containing the prize were .2, .4, and .4. Again, the odd key (.2) was in a different location for each of the three birds. Birds were run until they individually reached a stable performance criterion, defined as five consecutive days over which the daily percentage of initial choices on each key and the percentage of switch responses each varied by less than 10 %. In Condition H,

the three birds reached the criterion after 26, 50, and 54 days. In Condition L, the three birds reached the criterion after 17, 38, and 80 days.

Results

Condition H The left group of bars in Fig. 1 display performance in terms of the two dependent variables: initial key selection and switching. Over the final 5 days, pigeons in Condition H chose one of the two keys that had a lower chance of containing the prize on 98.33 % of trials, 95 % CI¹ [94.11, 102.56]. The confidence interval encompasses 100, meaning that pigeons' performance was indistinguishable from the optimal strategy of choosing a low-probability key 100.00 % of the time ($d = 0.98$,² $p = .232$). It was, however, reliably different from a random strategy, whereby one of the two low-probability keys would be chosen 66.67 % of the time, as that value falls outside of the confidence interval ($d = 18.62$, $p < .001$). Also over the final 5 days, pigeons switched on 94.27 % of trials, 95 % CI [89.29, 99.24]. This rate of switching was slightly lower than the optimal strategy of switching 100 % of the time ($d = 2.86$, $p = .038$), and considerably higher than a random strategy of switching 50 % of the time ($d = 22.10$, $p < .001$).

Over the final 5 days, the three birds in Condition H were reinforced on 65.28 % of trials, 95 % CI [63.61, 66.94]. This is slightly lower than the theoretical maximum of 70.00 %³ ($d = 1.17$, $p < .001$). That theoretical maximum assumes that a bird responds optimally during both phases of every trial: that it chooses a low-probability location and then switches. Two other models of pigeons' responding are also of interest: One corresponds to a bird that chooses optimally during the first stage of a trial but not the second, and the other corresponds to a bird that chooses optimally during the second stage of a trial but not the first. In the first case, a bird that optimally selected one of the two low-probability starting keys on every trial, but randomly chose to switch or stay would be reinforced on 50.00 % of all trials. Pigeons easily outperformed this value ($d = 6.57$, $p < .001$). In the second case, a theoretical bird that responded randomly on its first choice, but always switched, would be correct on 66.67 % of all trials. The earned reinforcement was very similar to this value ($d = 0.60$, $p = .091$), though note that the previous analysis

¹ All confidence intervals throughout the article are 95 % confidence intervals.

² Throughout the article, d is effect size, defined as the difference between the sample mean and the relevant comparison value, expressed in standard deviations.

³ Theoretical comparison values are based on the unconditional probabilities from Table 1, because they produce a single value that is dependent solely on the birds' (and not the host's) behavior. One can still test the viability of other theoretical values by noting whether or not they fall within the specified 95 % confidence intervals.

indicated that these pigeons did not make their initial choices randomly. Thus, performance was most consistent with a strategy that considered both stages of a trial.

Condition L The right group of bars in Fig. 1 shows performance in terms of the two dependent variables: initial key selection and switching. Over the final 5 days, pigeons in Condition L chose the key that had the lowest chance of containing the prize on 95.73 % of the trials, 95 % CI [90.15, 101.32]. This performance was not reliably different from the optimal strategy of choosing the low-probability key 100 % of the time ($d = 1.90, p = .081$). It was reliably different from a random strategy, whereby the low key would be chosen 33.33 % of the time ($d = 27.76, p < .001$). Also over the final 5 days, pigeons switched on 91.87 % of trials, 95 % CI [82.63, 101.10]. This rate of switching was also not different from the optimal strategy of switching 100 % of the time ($d = 2.19, p = .063$), but was different from a random strategy of switching 50 % of the time ($d = 11.26, p = .003$).

Over the final 5 days of the experiment, the three birds in Condition L were reinforced on 73.33 % of trials, 95 % CI [70.54, 76.13]. This was slightly lower than the theoretical maximum of 80 % for a bird that chose optimally during both stages of a trial ($d = 1.71, p < .001$). As was done for the data from Condition H, we also compared the numbers of earned reinforcers to the expectations for a bird that chose optimally during one stage of a trial but not both. A bird that selected the best starting key, but randomly chose to switch or stay would be reinforced on 50.00 % of all trials. Pigeons earned significantly more reinforcers than such a strategy would earn ($d = 5.98, p < .001$). A theoretical bird that responded randomly on its first choice but always switched would be correct on 66.67 % of all trials. Birds also gained significantly more reinforcers than such a strategy would earn ($d = 1.62, p < .001$).

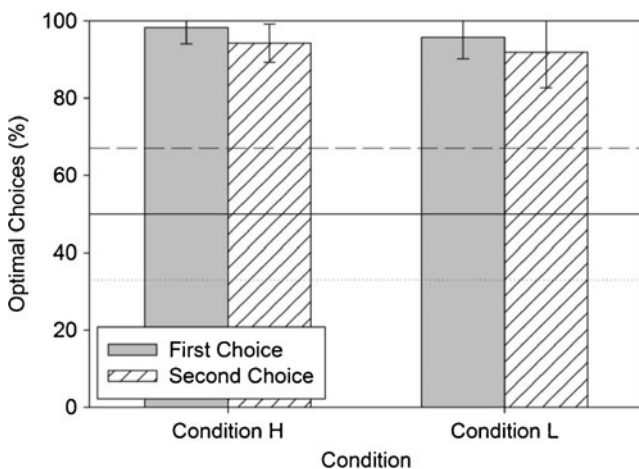


Fig. 1 Percentages of optimal choices in Condition H (left cluster) and Condition L (right cluster) of Experiment 1. The optimal first choice (gray bars) is to peck a low-probability key. The optimal second choice (striped bars) is to switch and peck a different key. Reference lines are at 33 %, 50 %, and 67 %. Error bars depict 95 % confidence intervals

Analysis of two-response sequences

The two choices on a trial together constituted a short response sequence, and six different sequences could be produced on any given trial: left–left, center–center, right–right, left–switch, center–switch, and right–switch. Notice that this list of possible sequences combines each pair of possible switch sequences (e.g., left–center and left–right) into a single category (left–switch). This compression acknowledges the fact that only one of the two possible switching choices can be available on a given trial, depending on which key was deactivated between responses. Not all of the nine key-specific paths could be available on any given trial, but any of the six sequences characterized above could be produced on any trial. This scheme thus acknowledges that the task involved an initial choice from among three options and a subsequent choice from among two, yielding $2 * 3 = 6$ possibilities.

Figure 2 shows how pigeons' tendencies to produce each of the six possible sequences changed over the course of the experiment. In order to combine data across individuals, location identities were shifted so that the odd-probability key would always appear as the “left” key. Because each bird completed a different number of sessions before reaching the stability criterion, days of training were divided into four equal

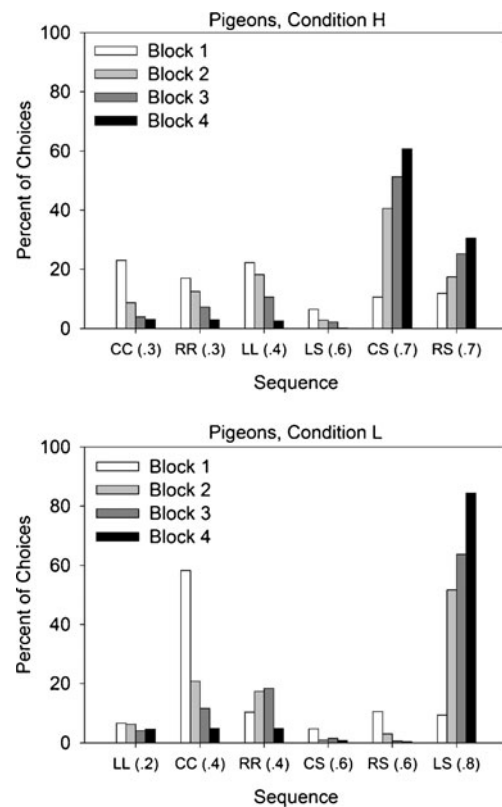


Fig. 2 Percentages of trials on which each of six possible response sequences was produced in Conditions H (top panel) and L (bottom panel) of Experiment 1. The sequences in each panel are ordered from left to right according to the probability that each will result in reinforcement

blocks based on the total number of trials that each bird completed during the entire experiment.

Inspection of Fig. 2 reveals some informative patterns. In particular, it is quite apparent that pigeons began the experiment by producing all six response sequences (white bars), but ended it with a strong preference for the sequence (or sequences) that had the highest probability of reinforcement (black bars). Because many of the bars displayed in Fig. 2 are based on small numbers of trials, we did not perform tests for statistical reliability. Instead, we computed these numbers specifically because they can be combined into a more robust analysis of overall sequence variability, presented below.

Figure 3 summarizes changes in four different measures of performance over the course of the experiment: initial key selections (top left panel), switches (top right panel), wins (bottom left panel), and U (bottom right panel). U (Miller & Frick, 1949) is a measure of sequence uncertainty, defined as

$$U = -\sum_{i=1}^6 \frac{RF_i \times \log_2(RF_i)}{\log_2(6)}. \quad (2)$$

RF is the relative frequency of occurrence for each of the six possible sequences, which corresponds to the sequence data displayed in Fig. 2. Possible U values range from 0 to 1, and higher values indicate similar frequencies of occurrence for each sequence. The highest possible U value (1.0) would be obtained from an individual that produced each of the six possible sequences exactly the same number of times, and thus was completely unpredictable. In contrast, lower U values indicate that some sequences were more likely than others. The lowest possible U value (0.0) would correspond to an individual that produced only a single response sequence, and did so on each and every trial.

We ran separate 4 (block: 1, 2, 3, 4) \times 2 (condition: H, L) analyses of variance (ANOVAs) on each of the four measures of performance shown in Fig. 3. In each case, we found a significant main effect of blocks. Over the course of the experiment, pigeons became more likely to choose the best starting key [top left panel; $F(3, 12) = 8.37, p = .003$] and to switch [top right panel; $F(3, 12) = 15.34, p < .001$]. Apparently these two elements of the optimal strategy were learned simultaneously, since neither began to rise before the other. As birds learned to do both of these things, their proportions of wins increased [bottom left panel; $F(3, 12) = 10.45, p < .001$] and their response sequences became more stereotyped [bottom right panel; $F(3, 12) = 6.60, p = .007$]. No significant main effects or interactions involved condition (all p s $> .165$).

Discussion

Pigeons approximated the optimal solution in both conditions. This means that they learned to do two things. First, they

learned to begin each trial by selecting the key with the lowest probability of being the prize location. Second, they learned to switch during the second stage of a trial. Only by doing both of these things could pigeons earn close to the maximum possible number of reinforcers.

These results are similar to those of Herbranson and Schroeder (2010), in that pigeons approximated the optimal strategy. In the present experiment, however, the optimal strategy was more complex, in that birds not only needed to switch, but also needed to account for the uneven distribution of prize locations by selecting the lowest-probability key during the first stage of each trial. Correspondingly, one might expect that it would take longer to learn the optimal strategy. This was indeed the case. In the present experiment, pigeons reached the criterion for stable performance after an average of 43 and 45 days of training in Conditions H and L, respectively (and one bird in Condition L required all of 80 days to reach the criterion). Herbranson and Schroeder's (2010) pigeons were all responding nearly optimally after less than 30 days of training. In the present experiment, only one of the three birds in each condition had reached the stability criterion by the 30th day of training.

These results contain an interesting quirk, in that the optimal strategy learned by pigeons required that they specifically avoid the key that had the strongest association with food reward. If pigeons simply responded by selecting the available key that had been consistently paired with access to grain, their trial-initiating responses would have drifted toward that key, and their overall probability of reinforcement would have dropped, even if they continued to switch. Thus, pigeons must have been learning something aside from which individual keys were strongly associated with food.

What, then, did pigeons learn? Analysis of the response sequences suggests that pigeons learned to emit a specific sequence of two responses on each trial. Note, though, that the concept of a sequence here involves more flexibility than it might carry in some other contexts. In the present task, the sequence learned consisted of an initial selection of a specific key, followed by a decision to switch to one of two alternate keys. The specific alternative available was not always the same, and could not be predicted with a high degree of accuracy. Consequently, the motor pattern required for optimal responding by necessity varied from trial to trial. Thus, one should be careful not to equate sequence learning with a rigid motor pattern. Several sequence-learning tasks have shown learning despite similar kinds of variability within a sequence (Herbranson & Stanton, 2011; Howard & Howard, 1997).

Experiment 2

The pigeons in Experiment 1 performed optimally by doing two things: They learned to choose the lowest-probability key (or keys) at the beginning of a trial and then to switch. Several

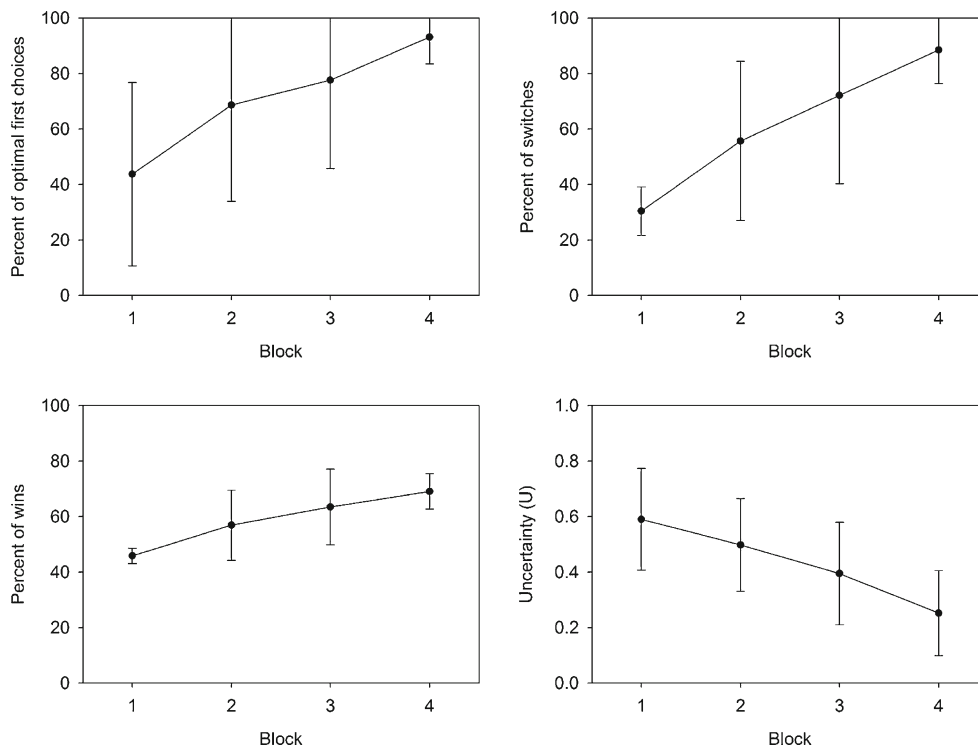


Fig. 3 Four measures of performance from the pigeons in Experiment 1. Top left: Percentages of optimal first choices. Top right: Percentages of switches. Bottom left: Percentages of wins. Bottom right: Uncertainty

(*U*). All values combine data from Conditions H and L. Error bars depict 95 % confidence intervals

research programs have established that humans ordinarily fail to perform optimally on a standard MHD. However, optimality in most of these experiments is based solely on whether a participant stays or switches, and the initial choice has no bearing on the likelihood of success. Whereas switching is apparently difficult for humans to learn, it is possible that the other aspect of optimal performance (the initial choice) may not be, or that a different context like the one in Experiment 1 might facilitate switching.

Granberg (1999) conducted an experiment involving a variant on the MHD with different base rates. As in Experiment 1, the optimal solution was to first choose the location with the lowest probability of containing the prize and then to switch. Although they achieved some learning over 60 trials, Granberg's participants failed to adopt the optimal solution: They did not consistently switch, and they did not consistently start by choosing the low-probability door. However, it is important to note that Granberg's methods differed from those of Experiment 1 in two crucial ways. First, there were four doors to choose from, rather than the customary three. Second, participants were explicitly informed that the different doors had specific probabilities of containing the prize. Therefore, we cannot be certain how human participants would respond to a procedure that more closely paralleled Experiment 1.

Experiment 2 replicated the logic of Experiment 1 with human participants. As in Experiment 1, two conditions were presented: one that featured a single location with a higher

likelihood of containing the prize (Condition H), and one that featured a single location with a lower likelihood of containing the prize (Condition L). Optimal performance in either case involved selecting a location with the lowest likelihood of being the prize location, and then switching. The two variables of interest were the same as in Experiment 1: initial choice and switching.

Method

Participants A group of 12 participants were recruited from undergraduate psychology courses. Some of the participants received partial course credit for their participation.

Procedure Each participant completed 100 trials of a procedure that mirrored Experiment 1. Rather than being projected onto pecking keys, the stimuli were presented on a 17-in. flat-panel computer monitor. Participants responded by guiding a cursor with a mouse and clicking on the stimuli.

At the beginning of each trial, the computer selected one of three locations as the prize location. Three white squares, 3×3 cm, were then presented, evenly spaced in a horizontal row. Each had an identifying letter in the center: "L," "C," and "R" for the left, center, and right locations, respectively. A single mouse click on one of those three squares cleared the screen and triggered a 1-s delay. Responses to other areas of the screen had no consequences. During the delay, the computer

selected one location to deactivate, using the standard constraints of the MHD. After the delay, two green squares with identifying letters were presented in the same locations, and a single mouse click on either square produced visual feedback in the center of the screen: the word “win” if the location corresponded to the prize location, and the word “lose” if it did not. The feedback remained visible for 2 s, after which the screen was cleared, and a 3-s ITI preceded the next trial.

During the experimental session, care was taken to avoid phrases associated with the Monty Hall dilemma, such as “door,” “prize,” “switch,” or “goat.” Participants were told that they were in an experiment on choice, and that they should try to win as many trials as possible.

Conditions The participants were arbitrarily assigned to one of two conditions, mirroring those in Experiment 1. In Condition H, the chances of the three locations containing the prize on each trial were .4, .3, and .3. Two of the six participants in Condition H had the odd probability (.4) in each of the left, center, and right locations. In Condition L, the chances of the three doors containing the prize were .2, .4, and .4. Again, two of the six participants had the odd probability (.2) in each spatial location.

Results

Condition H The left bars in Fig. 4 show performance over the final 50 trials in terms of the two dependent variables: initial choice and switching. Participants in Condition H chose one of the two locations with a lower chance of containing the prize on 67.67 % of trials, 95 % CI [53.34, 81.99]. This performance was reliably different from the optimal strategy

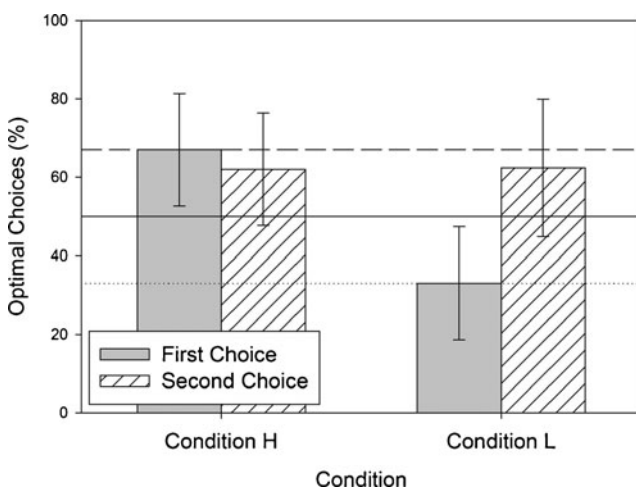


Fig. 4 Percentages of optimal choices in Condition H (left cluster) and Condition L (right cluster) of Experiment 2. The optimal first choice (gray bars) is to choose a low-probability location. The optimal second choice (striped bars) is to switch and choose a different location. Reference lines are at 33 %, 50 %, and 67 %. Error bars depict 95 % confidence intervals

of choosing a low-probability key 100 % of the time ($d = 2.37$, $p = .002$). It was not reliably different from a random strategy, whereby one of the two low-probability keys would be chosen 66.67 % of the time ($d = 0.07$, $p = .865$). Also over the final 50 trials, participants switched on 62.00 % of trials, 95 % CI [47.70, 76.30]. This rate of switching was different from the optimal strategy of switching 100 % of the time ($d = 2.79$, $p < .001$), but not different from a random strategy of switching 50 % of the time ($d = 0.88$, $p = .083$).

Participants in Condition H won 55.33 % of the final 50 trials, 95 % CI [48.48, 62.19]. This is reliably lower than the theoretical maximum of 70.00 % ($d = 2.25$, $p = .003$). That theoretical value corresponds to a participant that chose optimally during both stages of a trial. As with the pigeon data from Experiment 1, we could also compare human participants’ win rates with expectations for a participant that responded optimally during one stage of a trial, but not both. A participant that began each trial by optimally selecting one of the two low-probability starting keys, but then randomly chose to switch or stay would be reinforced on 50.00 % of all trials. The actual number of wins obtained was not significantly different from this figure ($d = 0.82$, $p = .102$), even though the previous analysis showed that participants chose one of the two low-probability keys almost precisely two-thirds of the time, corresponding to a random selection process. In contrast, a theoretical participant that made his initial choice randomly but always switched would be correct on 66.67 % of all trials. Participants won a significantly lower percentage of trials ($d = 1.74$, $p = .008$), and this is consistent with the previous analysis, indicating that they did not learn to switch consistently.

Although the previous analysis indicates that participants as a group did not learn to respond optimally, it was possible that some individual participants might have learned to do so. The leftmost panels in Fig. 5 show the performance of individuals from Condition H in blocks of 25 trials (the top left panel displays optimal first choices, and the bottom left panel displays optimal second choices).

As can be seen in Fig. 5, none of the individual participants approached the level of performance achieved as a group by the pigeons in Experiment 1. Nevertheless, note that considerable within- and between-subjects variability occurred across the experiment, and that some individuals achieved relatively high percentages of optimal choices during the final two blocks. In order to see whether any individual participants were significantly better than chance over the final 50 trials, we ran a series of binomial tests on our two dependent variables, summarized in Table 2. As can be seen in both Table 2 and Fig. 5, Participant H4 learned to perform better than chance on both phases of a trial (though still not 100 % of the time), and participant H3 switched significantly more than chance would dictate. Other participants showed very little change across the experiment, and in at least one case (H1), a participant seemed to abandon the strategy of consistently switching.

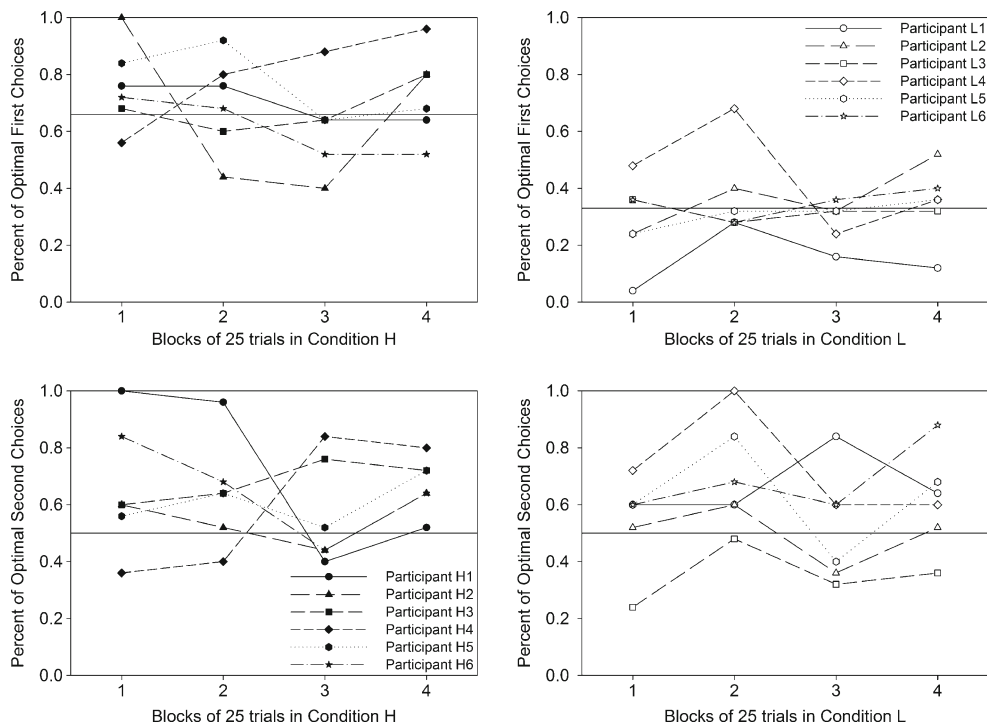


Fig. 5 Initial choices (top panels) and switching (bottom panels) in Condition H (left panels) and Condition L (right panels) of Experiment 2. Data points represent the percentages of optimal choices for individual

participants in blocks of 25 trials. The reference lines in each panel correspond to chance performance

Condition L The right bars in Fig. 4 show performance over the final 50 trials in terms of the two dependent variables: initial choices and switching. Participants in Condition L chose the location with a lower chance of containing the prize on 33.00 % of trials, 95 % CI [18.60, 47.40]. This was reliably different from the optimal strategy of choosing the low-probability key 100 % of the time ($d = 4.88, p = .001$). It was not reliably different from a random strategy, whereby the low-probability key would be chosen 33.33 % of the time ($d = 0.02, p = .955$). Also over the final 50 trials, participants switched on 62.33 % of the trials, 95 % CI [44.80, 79.86]. This rate of switching was different from the optimal strategy of switching 100 % of the time ($d = 2.25, p = .003$), but not different from a random strategy of switching 50 % of the time ($d = 0.74, p = .130$).

Over the final 50 trials, participants in Condition L won 58.67 % of trials, 95 % CI [46.03, 71.31]. This is significantly lower than the theoretical maximum of 80.00 % ($d = 1.77, p = .007$). Again, that theoretical maximum corresponds to a hypothetical participant that made optimal choices on both stages of a trial, and again we could also compare participants' win rates with the expected win rates for hypothetical participants that chose optimally during one stage but not both. A participant that optimally began each trial by selecting the low-probability location but randomly chose to switch or stay would be reinforced on 50.00 % of all trials. The number of wins actually obtained was not significantly different from this value

Table 2 Optimal choices from individual participants over the final 50 trials of Experiment 2

Participant	First Choice		Second Choice	
	Optimal Choices	Binomial Probability	Optimal Choices	Binomial Probability
Condition H				
H1	32	.713	23	.664
H2	30	.874	27	.336
H3	36	.261	37	.001*
H4	46	.001*	41	.001*
H5	33	.605	31	.060
H6	26	.989	27	.336
Condition L				
L1	7	.999	37	.001*
L2	21	.126	22	.760
L3	16	.631	17	.984
L4	15	.729	30	.101
L5	17	.513	27	.336
L6	19	.287	37	.001*

Binomial probabilities are for N or more optimal choices out of 50 trials. The test proportions are .67 for the first choice in Condition H, .33 for the first choice in Condition L, and .5 for the second choice in both conditions. * $p < .001$

($d = 0.72$, $p = .138$), even though the previous analyses indicated that participants selected the low-probability key about one-third of the time, corresponding to a random selection process. In contrast, a participant that responded randomly on his first choice but always switched would be correct on 66.67 % of all trials. The average number of wins was also not significantly different from this value ($d = 0.66$, $p = .165$), although again the previous analyses indicated that participants did not switch consistently.

As was done with the data from Condition H, we also considered the possibility that although our human participants did not respond optimally as a group, some individuals might have. The performances of individuals, in blocks of 25 trials, are shown in the rightmost panels of Fig. 5. As was the case for Condition H, considerable variability occurred across the experiment, and (at least with regard to switching) some participants achieved high percentages of optimal responses over the final blocks. This observation from Fig. 5 is confirmed by the binomial tests summarized in Table 2. Although no participant selected the best starting key at a rate that was significantly greater than chance, two participants (L1 and L6) did apparently learn to switch at levels greater than chance (though, again, they did not approach the optimal value of 100 %).

Analysis of two-response sequences

Figure 6 is the equivalent of Fig. 2 from Experiment 1, and it shows the percentages of trials on which human participants produced each of the six possible response sequences across four equal blocks of training. Because each participant completed exactly 200 trials, each of the four blocks of trials shown in the figure corresponded to 25 trials.

An inspection of Fig. 6 shows a very different pattern than was seen in Fig. 2. Human participants began by producing all six sequences during the first block of trials (white bars), but unlike pigeons, did not settle on any consistent preference by the final block of trials (black bars). They may, however, have developed some subtle biases toward higher-probability sequences. The tallest bars in each condition are those that correspond to the highest probabilities of reinforcement (the rightmost sequences in Fig. 6), even though the differences were very subtle. Similarly, the shortest bars in each condition are those that correspond to the lowest probabilities of reinforcement (leftmost sequences). Nevertheless, even during the final block, differences between sequences were very small indeed.

Figure 7 parallels Fig. 3 from Experiment 1, and it shows the same four measures of performance: starting key selection, switching, wins, and U .

We ran separate 4 (block: 1, 2, 3, 4) \times 2 (condition: H, L) ANOVAs on each of the four measures of performance

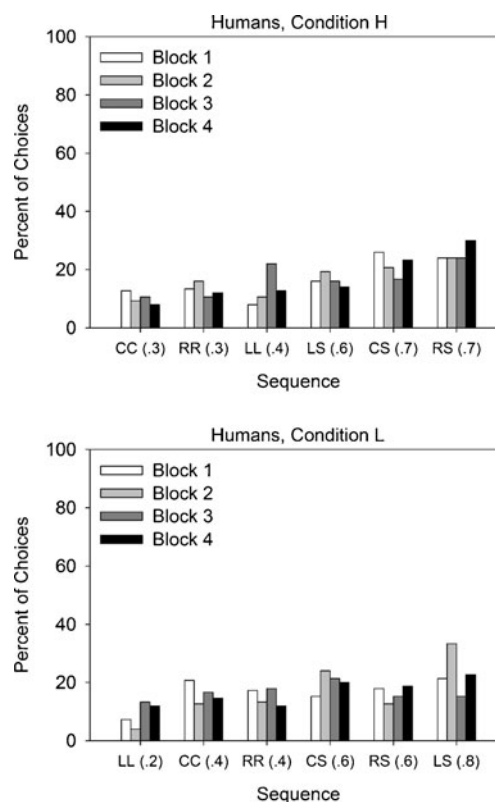


Fig. 6 Percentages of trials on which each of six possible sequences was produced in Conditions H (top panel) and L (bottom panel) of Experiment 2. The sequences in each panel are ordered from left to right according to the probability that each will result in reinforcement

displayed in Fig. 3. Humans, in contrast to pigeons, did not increase their likelihoods of choosing the optimal starting key [top left panel; $F(3, 30) = 1.65$, $p = .199$] or of switching [top right panel; $F(3, 30) = 1.33$, $p = .283$]. However, the probability of switching during Block 1 was considerably higher than that of the pigeons during Block 1 of Experiment 1 (dotted reference line). Part of the mystique of the MHD is that most humans initially approach it with a suboptimal tendency to stay. Our participants actually switched on 60.3 % of the first 25 trials. Though this figure was not significantly greater than 50 %, it is worth noting that it favors switch responses over stay responses. If participants did indeed have an initial preference to stay (nine out of the 12 participants stayed on the very first trial), it was overcome very quickly. Fast adjustment toward switching has been demonstrated before in investigations of the MHD: Granberg and Brown's (1995) participants shifted their strategy toward switching over the first 20 trials, and thereafter performance remained stable. Because humans did not increase their tendency to respond optimally during either phase of a trial, their proportions of wins did not increase over the course of the experiment [bottom left panel; $F(3, 30) = 0.02$, $p = .997$]. Finally, U values were uniformly high throughout the experiment [bottom right panel; $F(3, 33) = 0.63$, $p = .602$]. That uniformity could possibly be the consequence

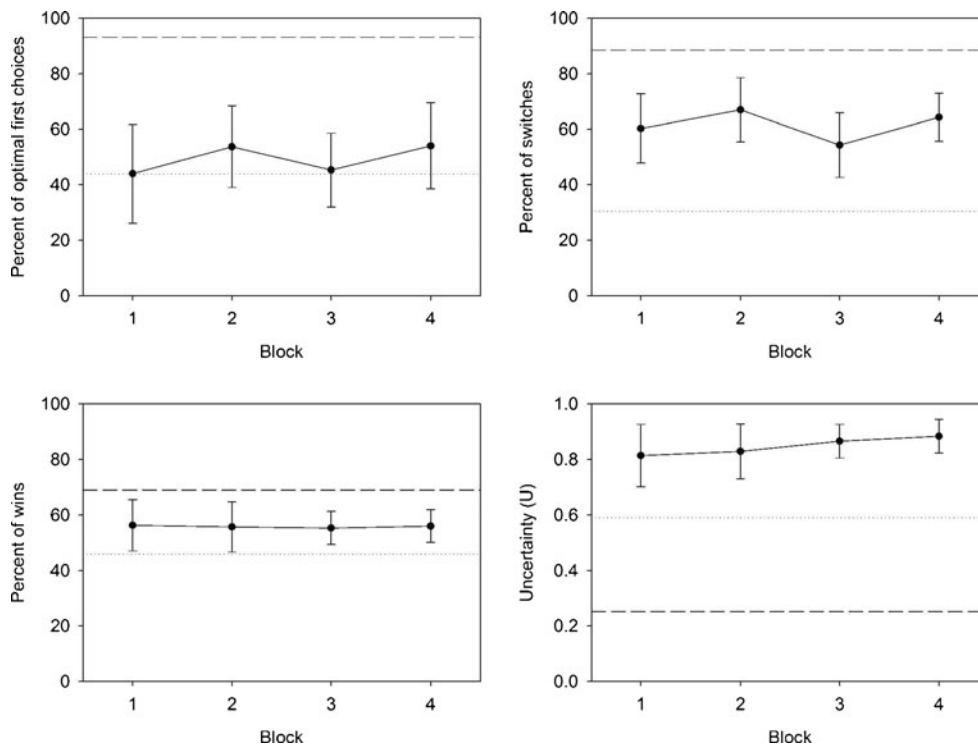


Fig. 7 Four measures of performance from humans in Experiment 2. Top left: Percentages of optimal first choices. Top right: Percentages of switches. Bottom left: Percentages of wins. Bottom right: Uncertainty (U). All values

combine data from Conditions H and L. Error bars depict 95 % confidence intervals. The reference lines correspond to pigeons' mean performance during the first (dotted) and fourth (dashed) blocks of Experiment 1

of a ceiling effect, as each subsequent block had a slightly higher U value, but little room to increase before reaching the maximum possible value. Ceiling effect or not, such a high degree of sequence variability is consistent with the possibility that humans persistently searched for a strategy that would be correct on every trial, one proposed explanation for suboptimal performance on other difficult probability-learning tasks (Fantino & Esfandiari, 2002). The only statistically significant effect from these ANOVAs was a main effect of condition on initial choice [$F(1,10) = 23.62, p < .001$], though this likely only reflects the different values corresponding to chance performance in the two conditions (.67 for Condition H vs. .33 for Condition L). None of the other main effects or interactions involving condition were significant (all other $ps > .390$).

Discussion

As a group, the human participants in Experiment 2 failed to optimize during both phases of a trial: They did not begin each trial by choosing the lowest-probability location, and they did not learn to switch consistently. Their performance is thus different from the nearly optimal performance of pigeons in Experiment 1 but is consistent with previous results from human participants (Granberg, 1999). Furthermore, some participants in each condition adopted at least part of the optimal solution, indicating that some degree of learning was possible

and that contingencies in effect had some influence over behavior.

Even though the performance of our human participants was not optimal, the responses revealed some informative patterns. First, participants' initial choices were evenly distributed across the three locations. When a single low-probability option was available among the three locations (Condition L), it was selected almost precisely one-third of the time. When two low-probability options were among the three locations (Condition H), one of them was selected approximately two-thirds of the time. This is exactly what would be expected if participants were making their initial choices in a random fashion, or at least allocating equal numbers of responses to each of the three available options. Even when data were analyzed on an individual level, only a single participant out of 12 eventually showed a significant preference for the optimal first choice. Apparently, participants did not learn much, if anything, about the different base rates over the course of the experiment.

The other element of the optimal strategy was to switch on every trial, and doing so would result in a greater probability of winning, regardless of which location was initially selected. The mean rate of switching for our participants, however, was not reliably greater than 50 % in either condition, meaning that we cannot conclude that they learned this aspect of optimal strategy, either. Note, however, that previous MHD experiments (Granberg & Brown, 1995; Herbranson & Schroeder,

2010) have shown switching probabilities greater than chance, with means very similar to those in the present experiment (about 60 %). Thus, the lack of overall significance here may have been due to our small number of participants: We recruited 12, whereas Granberg and Brown, for example, had 114. Given that some of our participants (two in each condition) showed individual switching levels that were greater than chance, we think it is a reasonable possibility that a larger sample might yield similar statistical conclusions.

Granberg's (1999) similar MHD experiment with unequal base rates yielded comparable results, in that participants in general did not consistently select the low-probability location, did not consistently switch, and did not maximize their expected wins. However, one important difference is that a small number of Granberg's participants did eventually stumble across and adopt the optimal strategy. He characterized the process as one of sudden realization (a "Eureka" moment), and proposed that the unequal-base-rate MHD is a useful procedure for studying insight learning. It is interesting that none of our participants experienced this kind of insight (participant H4 might have been close), despite completing a larger number of trials (100 vs. 60). One possible reason is that the optimal strategy in Granberg's experiment yielded a 90 % win probability, which was higher than either of the optimal strategies in the present experiment. Furthermore, given that Granberg's procedure utilized four locations, rather than the three in the present experiments, the optimal strategy existed within a larger field of possible response sequences, some of which had even lower probabilities of winning than the sub-optimal sequences in the present experiment. Thus, the superiority of the optimal response sequence in the present experiment may not have been as easy for participants to notice and appreciate.

We might now consider the same question that we asked at the close of Experiment 1: What did humans learn? Unlike the pigeons, our human participants did not learn to execute a specific sequence of responses that maximized their win rate. They also did not learn either of the two elements of the optimal strategy: initial selection of the low-probability location or switching. They did, however, produce highly variable responses, as revealed by our analysis of sequence variability. We would point out that although highly variable responses did not produce a high rate of reinforcement during the 100 trials that constituted the experiment, response variability was precisely what might eventually lead to adoption of the optimal strategy. Pigeons in Experiment 1 were able to produce the best sequence during later blocks only because they had tried all possible sequences during the earliest blocks. Similarly, some of Granberg's (1999) participants gained insight into the optimal solution by trying a variety of strategies, eventually stumbling upon the best one and noticing its superiority. We suspect that such a realization could take considerable time, given our methods. However, some subtle

changes to parallel Granberg's methods might serve to facilitate such an insight. Among these are the implementation of a higher possible win rate, explicit mention of the best possible win rate (emphasizing that it is less than 100 %), and specification of the probabilities associated with each location. If such a realization were to happen, note that acquisition might still look quite different from our pigeons' gradual adoption of the optimal strategy seen in Fig. 3.

General discussion

In these experiments, pigeons adopted strategies that were very close to optimal, whereas humans did not. The procedures utilized were different from the standard MHD, in that responses during both phases of a trial had influence over the likelihood of winning. Thus, an interpretation of the results reported here must first consider two things: the tendency to switch and initial choice.

Tendency to switch

The importance of switching is the element that the present methods have in common with the standard MHD. In both cases, the probability of winning on any given trial is considerably higher if one switches, and so the optimal strategy is to switch on every trial. Thus, with respect to the second stage of a trial, the results reported here are consistent with those of previous studies. Pigeons learned to switch on virtually every trial, whereas humans did not.

The difference between the switching performances of pigeons and humans on these experiments may possibly illustrate the difference between two contrasting response strategies: maximizing and probability matching (Herrnstein, 1997). Maximizing results in the greatest possible number of wins over the long run, and is what pigeons achieved by consistently switching. Probability matching involves matching the distribution of responses to the probability that each response will win, and is consistent with the average performance of humans. In the MHD (including both the standard version and the variants tested here), switching wins on average two-thirds of the time, and staying wins on average one-third of the time. Probability matching therefore would involve switching on two-thirds of all trials and staying on the remaining one-third, and this is very close to the mean switching rates of human participants in both conditions of Experiment 2 and in Herbranson and Schroeder (2010). Although a matching strategy is inferior to maximizing, it is important to note that both strategies are sensitive to the probabilities in effect and indicate some kind of learning. Nevertheless, whereas our aggregated human results were consistent with matching, we are hesitant to conclude that our human participants were indeed matching. First, it is

important to point out that those group results included considerable variability and were also consistent with a random decision to stay or switch. Second, some individual participants (two out of six in each condition) learned to consistently switch and actually outperformed the figure that would correspond to matching.

Initial choice

The element of the methods employed here that differs from the standard formulation of the MHD is the implementation of unequal base rates across the three initially available options. A consequence of those unequal base rates is that the initial choice on a trial influences the subsequent probabilities that switching and staying will win. The optimal strategy is to begin each trial by choosing the location having the lowest probability of containing the prize and subsequently to switch. Again, pigeons learned to do this consistently, whereas humans did not and produced a distribution of initial choices that was very close to random.

Consistent with the results of Herbranson and Schroeder (2010), pigeons maximized, and in doing so, also adopted a preferred key for their initial choices. The methods employed here were different, however, in that a specific initial key preference was necessary for optimal performance. In the standard MHD, the first selection has no bearing on the probability of winning. Pigeons' initial key preference in the standard MHD, then, is simply a consequence of the fact that no advantage is to be had by varying their initial selections. Different birds in Herbranson and Schroeder's study adopted different key preferences, without any influence over the probability of gaining reinforcement. In the present experiments, however, birds had to adopt a specific key preference in order to maximize the probability of gaining reinforcement, and they did so. Each of the three birds adopted a preference for a different key (left, right, or center), depending on which key was associated with the lowest available probability.

Humans' initial selections did not vary as a function of win probability. Our human participants chose the low-probability key one-third of the time in Condition L, and one of the two low-probability keys two-thirds of the time in Condition H, a pattern consistent with an even distribution of selections across all three keys. Thus, they apparently did not appreciate the different base rates, or realize that initial key selection had any bearing on the probability of winning.

Conclusions

The variations on the MHD introduced here are informative for behavioral science in part because they address a puzzlingly suboptimal behavior pattern that is characteristic of humans, yet not seen in pigeons. Perhaps the simplest possible strategy to learn and implement in the MHD would be one that

involves always choosing the location that has the highest association with reinforcement. Doing so is both simple and intuitive, but is also suboptimal and would correspond to only 40 % reinforcement in both conditions (a loss of either 30 % or 40 % as compared with the optimal strategy, depending on the condition). Neither species adopted such a simple strategy. Humans were not drawn to the higher-probability locations, but rather distributed their initial responses evenly across all three. Pigeons learned specifically to avoid the higher-probability keys with their initial choices. Subsequently, the two species diverged even further, in that pigeons switched on virtually all trials, whereas humans did not (though some individual participants did show a significant tendency toward switching). When the two stages of a trial were combined, it meant that pigeons learned to execute a specific sequence of two responses, and in doing so earned close to the theoretical maximum payout, whereas humans won at a slower rate that was close to the expected payout for a random strategy.

It seems that with training, pigeons learned to execute specific, short sequences of responses. During the initial days of the experiment, pigeons' responses appeared to be random, since a large number of response sequences were produced in approximately equal proportions.⁴ By the end of the experiment, the sequences that resulted in the highest probabilities of reinforcement were produced almost exclusively. Humans also produced a wide variety of response sequences at the beginning of the experiment, but they did not eventually display a strong bias toward the optimal sequences, perhaps because they did not perceive the sometimes subtle differences in reinforcement probabilities associated with different response sequences. Note that humans received less training than the pigeons in Experiment 1, and this difference may or may not have accounted for the discrepant results. Although our human participants completed more trials than had been required to reach a stable level of performance in previous investigations of the MHD (Granberg & Brown, 1995; Herbranson & Schroeder, 2010), we cannot discount the possibility of an extremely slow, extremely variable learning process. This possibility is supported by the results from one of our participants (H4), who appeared to be very close to the optimal strategy by the end of the experiment (see Fig. 5), and by those of Granberg (1999), who also reported that some individual participants eventually gained insight into the optimal solution to a similar MHD-based task. It is apparent, then, that humans *can* learn the optimal solutions to various versions of the MHD. Nevertheless, the present data support previous work indicating that the MHD constitutes a particularly challenging probability-learning scenario; humans often master other probability-

⁴ Although the proportions of sequences varied from bird to bird, each bird produced all of the possible sequences during the initial days of training, and no single sequence strongly predominated in any individual bird's responses.

learning tasks (e.g., Edwards, 1961; Gardner, 1957) in far fewer trials than were completed in Experiment 2.

It has been proposed that humans fail at the MHD because of numerous decision traps (see Herbranson, 2012, for a summary), and these decision traps may also explain the difference between pigeon and human performance in the present experiment. In other words, it may be that pigeons perform optimally by treating the task as nothing more than a straightforward probability-learning task. Humans, on the other hand, are influenced by something other than the probabilities of reinforcement in effect. Plausible explanations for humans' failure on the standard MHD that have been offered previously would presumably be just as relevant here. Among them are an equiprobability bias (De Neys, 2007), the illusion of control (Granberg & Dorr, 1998), faulty causal reasoning (Burns & Wieth, 2004), and anticipation of regret (Gilovich et al., 1995).

Finally, the two-stage nature of the methods utilized here may hold potential for additional investigations of cognition in pigeons, humans, and perhaps other animals. It may, for example, constitute a simple method for studying planning, in that an animal must make a low-probability choice in order to make available later a higher-probability choice. It might also yield further insights into sequence learning, optimality, or probability matching. Thus, aside from being a delightful, if not aggravating puzzle, the MHD may constitute a useful tool for investigating various aspects of comparative cognition.

References

- Brown, P. L., & Jenkins, H. J. (1968). Autoshaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior*, *11*, 1–8.
- Burns, B. D., & Wieth, M. (2004). The collider principle in causal reasoning: Why the Monty Hall dilemma is so hard. *Journal of Experimental Psychology: General*, *133*, 434–449.
- De Neys, W. (2007). Developmental trends in decision making: The case of the Monty Hall Dilemma. In J. A. Elsworth (Ed.), *Psychology of decision making in education, behavior, and high risk situations* (pp. 271–281). Hauppauge, NY: Nova Science.
- Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology*, *62*, 385–394.
- Fantino, E., & Esfandiari, A. (2002). Probability matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology*, *56*, 58–63. doi:10.1037/h0087385
- Gardner, R. A. (1957). Probability-learning with two and three choices. *American Journal of Psychology*, *70*, 174–185.
- Gilovich, T., Medvec, V. H., & Chen, S. (1995). Commission, omission, and dissonance reduction: Coping with regret in the “Monty Hall” problem. *Personality and Social Psychology Bulletin*, *21*, 182–190.
- Granberg, D. (1999). A new version of the Monty Hall dilemma with unequal probabilities. *Behavioural Processes*, *48*, 25–34.
- Granberg, D., & Brown, T. A. (1995). The Monty Hall dilemma. *Personality and Social Psychology Bulletin*, *21*, 711–723.
- Granberg, D., & Dorr, N. (1998). Further exploration of two-stage decision making in the Monty Hall dilemma. *American Journal of Psychology*, *111*, 561–579.
- Herbranson, W. T. (2012). Pigeons, humans and the Monty Hall dilemma. *Current Directions in Psychological Science*, *21*, 297–301.
- Herbranson, W. T., & Schroeder, J. (2010). Are birds smarter than mathematicians? Pigeons (*Columba livia*) perform optimally on a version of the Monty Hall Dilemma. *Journal of Comparative Psychology*, *124*, 1–13.
- Herbranson, W. T., & Stanton, G. L. (2011). Flexible serial response learning by pigeons (*Columba livia*) and humans (*Homo sapiens*). *Journal of Comparative Psychology*, *125*, 328–340.
- Hermstein, R. J. (1997). *The matching law*. Cambridge, MA: Harvard University Press.
- Howard, J. H., Jr., & Howard, D. V. (1997). Age differences in implicit learning of higher-order dependencies in serial patterns. *Psychology and Aging*, *12*, 634–656.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
- Miller, G. A., & Frick, F. C. (1949). Statistical behavioristics and sequences of responses. *Psychological Review*, *56*, 311–324.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32. doi:10.1016/0010-0285(87)90002-8
- Poling, A., Nickel, M., & Alling, K. (1990). Free birds aren't fat: Weight gain in captured wild pigeons maintained under laboratory conditions. *Journal of the Experimental Analysis of Behavior*, *53*, 423–424.
- Stevens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Tubau, E., & Alonso, D. (2003). Overcoming illusory differences in a probabilistic counterintuitive problem: The role of explicit representations. *Memory & Cognition*, *31*, 596–607.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Zentall, T. R. (2011). Maladaptive “gambling” by pigeons. *Behavioural Processes*, *87*, 50–56.