Context-dependent utility overrides absolute memory as a determinant of choice

Lorena Pompilio[1] and Alex Kacelnik[2]

Zoology Department, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom

Edited by Charles R. Gallistel, Rutgers, The State University of New Jersey, Piscataway, NJ, and approved November 11, 2009 (received for review June 29, 2009)

A core problem of decision theories is that although decision-makers' preferences depend on learning, their choices could be driven either by learned representations of the physical properties of each alternative (for instance reward sizes) or of the benefit (utility and fitness) experienced from them. Physical properties are independent of the subject's state and context, but utility depends on both. We show that starlings' choices are better explained by memory for context-dependent utility than by representations of the alternatives' physical properties, even when the decisionmakers' state is controlled and they have accurate knowledge about the options' physical properties. Our results support the potential universality of utility-driven preference control.

state-dependent preferences | rationality | starling | optimal foraging | decision-making

Almost three centuries ago Daniel Bernoulli identified a major issue at the core of decision theory and anticipated the rationale for our present study. He wrote: "...the determination of the value of an item must not be based on its price, but rather on the utility it yields. The price of the item is dependent only on the thing itself and is equal for everyone; the utility, however, is dependent on the particular circumstances of the person making the estimate. Thus there is no doubt that a gain of one thousand ducats is more significant to a pauper than to a rich man though both gain the same amount." (1)

The significance of Bernoulli's insight cannot be overestimated. Although he was dealing with directly perceptible information, we believe that the idea is also crucial for problems involving learning and memory and, here, we develop this argument. Most decisions depend on acquired information, and individuals use learning to adjust preferences. Once learning is invoked, the significance of Bernoulli's dichotomy jumps to the fore. Suppose that a subject learns that foraging option A yields prey after a chasing time t, whereas option B yields the same prey type after $t + \Delta t$. Superficially, it would seem logical that given a choice the decisionmaker (DM) would prefer A. However, if the two food sources are encountered in different circumstances, the ranking of their absolute properties may differ from the ranking of the corresponding utilities. One possible cause for a mismatch in ranking is that the DM may learn about the different options when its needs are different. If the objectively inferior option B is typically encountered when the subject is in greater need, then B can be remembered as yielding greater utility than A. This phenomenon is called state-dependent valuation learning and has been shown in mammals (2), birds (3–5), fish (6) and insects (7). In humans, contradictions between expectations based on utility versus absolute gains are observed in various context-dependent phenomena that violate the principle of invariance, suggesting that the same concept may apply to human decisionmaking (8, 9).

This taxonomic and situational breadth of manifestation suggests that Bernoulli's functional dichotomy may cause evolutionary convergence in the relation between memory and choice. Additional support for this view is that the underlying mechanism seems to differ across taxa: In birds the manipulation of energetic state during learning affects preferences but not accuracy of representation of absolute properties (10); in insects, in contrast, the effect of past needs may alter the perception of each option (7).

Although convergent evolution suggests adaptive advantage, there is as yet no established theoretical reason why a mechanism that drives choice through remembered utility gains would be selected against one where preferences resulted from a direct comparison of the remembered metrics of the options. Despite Bernoulli's distinction, for any given DM in a fixed state, if the utility function is monotonic (as in all of the cases we consider) the ranking of objective metrics is identical to the ranking of utilities. Using absolute metrics would always yield the right choice, although using remembered utilities is vulnerable to the manipulation we use: An option with poorer metrics may be preferred because the DM remembers utilities accrued in different states for each option.

The concept would be strengthened if it worked when remembered utility is manipulated by a diversity of factors, rather than just need, and especially if the effect occurred without interfering with memory for absolute metrics. Here, we test the idea with an experimental design that pitches absolute against comparative processes of choice (11). We use European starlings (*Sturnus vulgaris*) as subjects and manipulate the utility of different options by altering their relative ranking through context manipulations, as in earlier experiments with pigeons (12–14). The starlings pecked at colored response keys (options) to obtain food rewards. Options differed in delay to reward, but to make general points we also refer to its reciprocal I, or 'Immediacy' ($I_i = d_i^{-1}$, where d_i is the delay to reward associated with option i). In Experiment 1, each bird was exposed to two phases: Training and Choice. In Experiment 2, we added two extra phases, so that each bird had to complete the Pre-Training, Training, Choice, and Post-Choice phases.

During the Training phase of both experiments (6 days), all birds were exposed to two different contexts. In each session under Context AB, the subjects were exposed to 10 trials with option A and 10 trials with option B, in random order and separated by an intertrial interval (ITI) of 50 s (see State equalization in *Methods*). Similarly, under Context CD, there were 10 trials with option C and 10 trials with option D, also in a random order and with 50 s of ITI. Options were programmed so that $I_A > I_B = I_C > I_D$ in Experiment 1 and $I_A > I_B > I_C > I_D$ in Experiment 2 (Fig. 1*a*). We denote each option by $X_{(z)}$, where X is the name of the option and z gives the option's delay to reward in seconds.

The Choice phase differed from the Training phase only in the inclusion of two choices before, during, and after each exper-

Author contributions: L.P. and A.K. designed research; L.P. performed research; A.K. contributed new reagents/analytic tools; L.P. and A.K. analyzed data; and L.P. and A.K. wrote the paper.

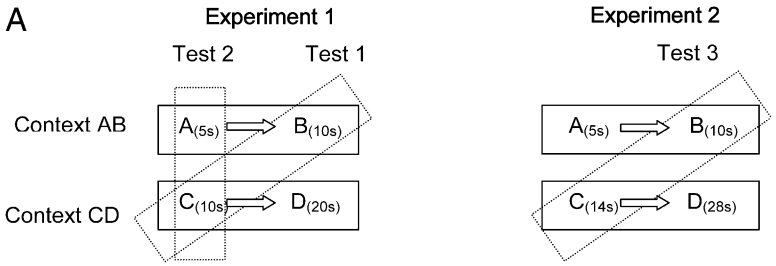
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Departamento de Ecología, Genética y Evolución. Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón II Ciudad Universitaria, C1428EHA Buenos Aires, Argentina.

[2]To whom correspondence should be addressed. E-mail: alex.kacelnik@zoo.ox.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0907250107/DCSupplemental.



B

Hypothesis	Predictions: Test 1 $B_{(10s)}$ vs. $C_{(10s)}$	Test 2 $A_{(5s)}$ vs. $C_{(10s)}$	Test 3 $B_{(10s)}$ vs. $C_{(14s)}$
a) Absolute memory	Indifference	A	B
b) Value by association	B	A	B
c) Remembered ranking	C	Indifference	C
d) Lexicographic combination:			
-Absolute memory priority	C	A	B
-Ranking priority	C	A	C
e) Non-lexicographic combination of absolute memory and ranking	C	A	Metrics favour B, Ranking favours C
RESULTS	**C**	**A**	**Indifference**

Fig. 1. (*A*) Description of Experiment 1 on the left and Experiment 2 on the right. The horizontal rectangles represent the options comprising each context. Context AB: options A and B; Context CD: options C and D. The numbers between parentheses represent seconds of delay to food. The arrows within each context indicate the ranking relationships, pointing toward the lower ranking option in each context. The dotted rectangles describe the choices presented to each subject during the choice phase. (*B*) Predictions from hypotheses a–e for Tests 1, 2, and 3. The bottom row shows the results obtained in each test.

imental session, so that the subjects continued to be exposed to sequential presentations of single options within specific contexts (Table S1). Choices could be among: (*i*) Options with equal immediacy but different ranking (Test 1 in Experiment 1, Choice of $B_{(10s)}$ vs. $C_{(10s)}$). (*ii*) Options with different immediacy but equal ranking (Test 2 in Experiment 1, choice of $A_{(5s)}$ vs. $C_{(10s)}$). (*iii*) Options differing in both immediacy and ranking with opposite sign (Test 3 in Experiment 2, Choice of $B_{(10s)}$ vs. $C_{(14s)}$).

In Experiment 2, the Pre-Training and Post-Choice phases were aimed at testing whether the birds discriminated between 10 s and 14 s of delay in the absence of context effects.

We consider the following hypothetical mechanisms of choice (see Fig. 1*b*):

(a) *Absolute memory.* If subjects' memory for the immediacy of each option drives their choices, and temporal memory is not affected by context, then in Experiment 1 they should be neutral between $B_{(10s)}$ and $C_{(10s)}$ in Test 1 and prefer $A_{(5s)}$ over $C_{(10s)}$ in Test 2. In Experiment 2 (Test 3) they should prefer $B_{(10s)}$ over $C_{(14s)}$.

(b) *Value by association.* If options acquire attractiveness by being in a richer environment, they should prefer in Experiment 1 $B_{(10s)}$ over $C_{(10s)}$ (Test 1) and $A_{(5s)}$ over $C_{(10s)}$ (Test 2). In Experiment 2 they should prefer $B_{(10s)}$ over $C_{(14s)}$ (Test 3).

(c) *Remembered ranking.* If options acquire positive or negative value depending exclusively on their ranking against alternatives in the same context, then $A_{(5s)}$ and $C_{(10s)}$ would become the most (and equally) attractive options. This predicts favoring $C_{(10s)}$ over $B_{(10s)}$ (Test 1) and indifference between $A_{(5s)}$ and $C_{(10s)}$ (Test 2) . $C_{(14s)}$ should be preferred over $B_{(10s)}$ in Experiment 2 (Test 3).

(d) *Lexicographic combination of absolute memory and ranking.* In lexicographic processes there is a priority dimension. If the alternatives differ in the priority dimension then choices are determined by this difference, but if they are tied then a secondary dimension exerts influence. Thus, if remembered immediacy had priority and remembered ranking were secondary, in Experiment 1 $C_{(10s)}$ would be preferred to $B_{(10s)}$ in Test 1 through the effect of the secondary dimension, and $A_{(5s)}$ would be preferred to $C_{(10s)}$ in Test 2 through the primary dimension. In Experiment 2 $B_{(10s)}$ would be preferred to $C_{(14s)}$ (Test 3) through the primary dimension. If the priority dimension were remembered ranking, then $C_{(10s)}$ or $C_{(14s)}$ would be preferred to $B_{(10s)}$ in both Experiment 1 (Test 1) and Experiment 2 (Test 3) trough the primary dimension, whereas $A_{(5s)}$ would be preferred to $C_{(10s)}$ in Experiment 1, Test 2, through the secondary dimension.

(e) *Nonlexicographic combination of absolute memory and ranking.* This hypothesis implies that instead of one of the dimensions having total priority, absolute memory and ranking can be traded against each other. Under this hypothesis an option is compared with the alternatives in all available dimensions (in this study, immediacy and ranking). An important difference with hypothesis d is that here magnitude, as well as sign of the differences in each dimension matter. This mechanism predicts preference for $C_{(10s)}$ over $B_{(10s)}$ in Test 1 because the only difference

(ranking) favors C, and for $A_{(5s)}$ over $C_{(10s)}$ in Test 2 (Experiment 1) because the only difference (immediacy) favors A. But the crucial comparison for this hypothesis comes from Experiment 2. Here in the Pre-Training and Post-Choice phases there is no context effect, and $B_{(10s)}$ should be preferred over $C_{(14s)}$ because of immediacy. However, in the Choice phase context should introduce the effect of remembered ranking in favor of C. Since, ranking and immediacy have opposite signs, this predicts a shift so that when facing $B_{(10s)}$ vs. $C_{(14s)}$ in Test 3 (Experiment 2) preference for $B_{(10s)}$ over $C_{(14s)}$ should be less than the preference shown in the pre and posttraining phases, perhaps even reaching indifference or preference for $C_{(14s)}$, depending on the relative strength of the ranking and absolute memory effects.

(f) *Dual memory vs. Distorted memory*. This distinction runs through the previous hypotheses. State and context effects can be mediated either by a dual system in which perception and memory of absolute metrics are unaltered and remembered utility is context- or state-dependent, or by modulating the perception or representation of the metrics of the options. For instance, if, in the comparison between B_{10} and C_{10}, the latter is preferred, this result could be mediated by a memory for the ranking of each option in its context, or by (incorrectly) remembering C as having greater immediacy than B, that is, by storing a distorted temporal representation. An additional complication in terms of mechanism is the discriminability of the options. For example, if the predictions of hypothesis e are confirmed, and context reduces or eliminates preference for the option with better absolute metrics this could be mediated by the animal's ability to discriminate between 10 s and 14 s being compromised by the presence of different contexts. Although the focus on this paper is not so much on the mechanism but the effect, these possibilities can be established empirically because the accuracy of memory for delays is exposed by the subjects' pecking behavior.

Results

Fig. 2 shows preferences in terms of choice proportions. In Test 1 all subjects showed a preference for the option with higher ranking [$C_{(10s)}$ dominated $B_{(10s)}$], whereas in Test 2 all preferred the option with greater immediacy [$A_{(5s)}$ dominated $C_{(10s)}$] (one-sample Student t test of magnitude of preference against indifference: $t_{(1.5)} = 16.57$, $P < 0.001$ and $t_{(1.5)} = 21.64$, $P < 0.001$, respectively, for each test).

In Experiment 2, Pre-Training and Post-Choice phases, all subjects strongly preferred the option with greater immediacy [$B_{(10s)}$ dominated $C_{(14s)}$], choosing the longer-delay option [$C_{(14s)}$] in only 12% ± 3.6 of the opportunities (mean ± SE). During the Choice phase (Test 3) preference for $C_{(14s)}$ increased to reach a mean (±SE) of 49.7% ± 2.8. The effect of phase (namely, of including contexts) was significant (repeated measures ANOVA $F_{2.5} = 47.11$, $P < 0.001$). This shift is caused both by differences between the Choice and Pre-Training phases and between Choice and Post-Choice phases (post hoc comparisons, $P < 0.001$). During the Choice phase preference did not depart significantly from indifference ($t_{(1.5)} = 0.11$, $P = 0.92$; one sample t test against indifference).

In the two experiments, pecking patterns reflected the absolute properties of the options. Qualitatively, the pecking pattern during options $B_{(10s)}$ and $C_{(10s)}$ in Experiment 1 was remarkably similar as shown in Fig. 3 and (standardized respect to its maximum) in Fig. S1. To test whether the preference for $C_{(10)}$ over $B_{(10)}$ in Test 1 was due to distorted timing, we compared the pecking rates and the location of pecking peaks between these two options. We found that the rate of pecking was significantly higher for option $C_{(10)}$ (as expected from any motivational hypothesis), but there was no interaction between time and option (repeated measures ANOVA $F_{2.5} = 11.62$, $P = 0.02$ and $F_{2.5} = 0.96$, $P = 0.52$, respectively, for each comparison). This result implies that there is no evidence of a timing effect. Similarly, the analysis of peak location did not detect any reliable difference between these options ($t_{(1.5)} = 1.29$, $P = 0.25$, see *SI Text*). In Experiment 2, options $C_{(14s)}$ and $B_{(10s)}$ differed in both their delays and in their ranking. Fig. 4 shows that the starlings discriminated between the delays, exhibiting displaced pecking rate curves as expected from the programmed delays (repeated measures ANOVA using each individual's ratio between mean pecking rate for option C over B as a dependent variable, and time-bin as a within-subjects fixed factor, $F_{1.5} = 4.58$, $P < 0.003$). These displaced curves contrast with the indifference shown in choices between these options (Test 3; Fig. 2), highlighting the separation between the effects of temporal memory and memory for ranking on choice.

Discussion

We summarize our results by noting the following: (*i*) when options differed in ranking but had equal immediacy (Test 1) preference was predicted by ranking; (*ii*) when options differed in immediacy but had equal ranking (Test 2) preference was predicted by immediacy; and (*iii*) when both ranking and immediacy differed with opposite signs (Test 3), they counteracted each other. Context manipulations did not noticeably affect temporal memory, indicating that preference effects are unlikely

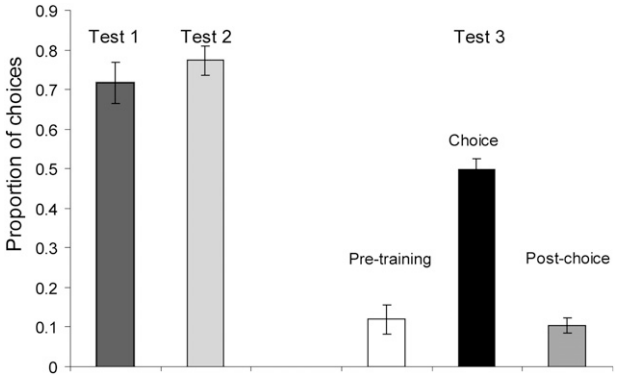


Fig. 2. Mean proportion of choices (±SE) for option $C_{(10s)}$ vs. $B_{(10s)}$ in Test 1 (Experiment 1), for option $A_{(5s)}$ vs. $C_{(10s)}$ in Test 2 (Experiment 1) and for option $C_{(14s)}$ vs. $B_{(10s)}$ in Pre-Training, Choice (Test 3), and Post-Choice phases (Experiment 2).

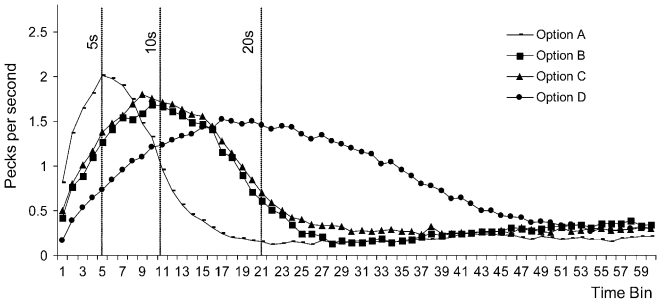


Fig. 3. Experiment 1: Mean pecks per second in probe trials for options $A_{(5s)}$, $B_{(10s)}$, $C_{(10s)}$, and $D_{(20s)}$. The vertical lines indicate the time bin where food was due for each option.

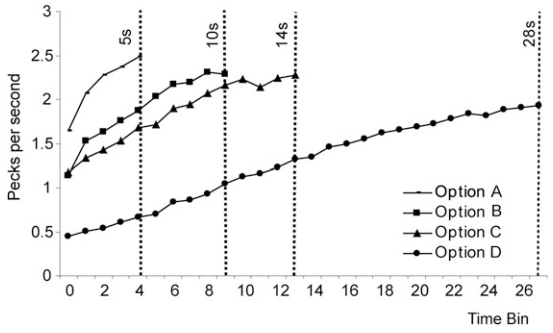


Fig. 4. Experiment 2: Mean pecks per second in single-option trials at the end of Training and during the Choice phase for options $A_{(5s)}$, $B_{(10s)}$, $C_{(14s)}$, and $D_{(28s)}$. The vertical lines indicate the time bin where food was delivered for each option.

to be mediated by distortions of temporal information processing. These findings support the hypothesis that proposes that choices are driven by a nonlexicographic combination of absolute memory and ranking. Further work is necessary to quantify the weights assigned by the subjects to each dimension. Our present analysis only confirms that when context factors (remembered ranking) compete with absolute memory, choice shifts, proving that absolute memory alone cannot account for the results. A similar choice mechanism has been proposed to explain choices in humans (15) and states that subjects use a comparative nonhierarchical evaluation mechanism whereby dimensions are taken into account simultaneously. The fact that context effects cause indifference between options with different time costs, or even preference for the worse alternative, has unchartered potential consequences in nature.

It is pertinent to ask why did these choice mechanisms evolve, given that the animals appear to store accurately the metrics of each option. The fact that absolute metrics for the delays involved are available to the animals and control their pecking pattern discards the possible benefit of avoiding the information-processing cost of remembering the properties of the options. A variant could be that it may be costly or time consuming to use temporal information at the time of choice rather than to encode it at the time of learning. We suspect that this alternative is equally invalid, because choice processes do not seem to cause time costs in this species (16). One possible advantage is that utility might simplify comparisons between alternatives that contribute to the same component of fitness (say energy balance) but differ in absolute properties such as prey size, handling time, or energetic density. Another is that utility could serve as an efficient common currency to compare options that differ in the fitness component to which they contribute, as when a decisionmaker has to choose between an option offering food and another offering refuge from cold.

Methods

Subjects. Subjects were 12 wild-caught European Starlings (*Sturnus vulgaris*) with previous experimental histories (English Nature license No. 20.020.068). A different group of birds participated in each experiment. The birds were kept outdoors before starting the experiment, when they were transferred to indoor individual cages (120 cm × 60 cm × 50 cm) that served as housing and testing chambers. Lights were on between 0600 and 1900 hours, and the room temperature ranged from 20 °C to 23 °C. Subjects were visually but not acoustically isolated.

Apparatus. Each cage had a panel with a central food hopper and two response keys that could be transilluminated with yellow, red, blue, or green. Computers running Arachnid language (Paul Fray Ltd.) served for control and data collection. Food rewards were in all cases 2 units of semicrushed and sieved Orlux pellets (≈0.02 g/unit) delivered at a rate of 1 unit/s by automatic pellet dispensers (Campden Instruments). Birds were permitted to feed ad libitum on turkey starter crumbs and supplementary mealworms for 5 h at the end of each day. This regime allows the starlings' body weights to remain stable at ≈90% of their free feeding value (17). During the rest of the day they only obtained food by working on the programmed schedules.

Experimental Protocol. ***Training phase.*** All subjects were exposed to a Training phase in which they experienced four sources of reward, arranging that two of them were encountered in one context and the other two in a different context. Context AB led to encounters with either option A or option B, whereas context CD led to encounters with either option C or D. To obtain the reward the birds had to peck at transilluminated colored response keys (a particular color was, for each bird, always associated with the same immediacy to reward). The quality of each option was given by its immediacy to reward I_i ($I_i = d_i^{-1}$, where d_i is the delay to reward associated with option i). As shown in Fig. 1, options were programmed so that $I_A > I_B = I_C > I_D$ in Experiment 1 and $I_A > I_B > I_C > I_D$ in Experiment 2. When necessary, we denote each option by $X_{(z)}$, where X is the appropriate letter and the suffix z gives the option's delay to reward in seconds. In both experiments the delay to food for option B was twice as long as that for option A, and the delay for option D was twice as long as that for C to keep constant the discriminability of the differences between the options from the same context. The colors associated with the options were counterbalanced across subjects and remained the same for each bird during the whole experiment.

The order and sides in which the options were presented was randomized. The order of presentation of contexts alternated across subjects and days. Half of the subjects started the first session with context AB and the other half with context CD. On one day, the subjects who began with context AB ended with context CD, and on the next day, they began with context CD and ended with context AB.

Three sessions per day were programmed, separated by 20 min of interval (see Table S1). Each session was defined by one presentation of both contexts AB and CD, separated by 5 min of interval. Each context was defined by the presentation of 10 single-option trials of each option comprising a context (for example, 10 presentations of A and 10 presentations of B in a random order). Notice that during the Training phase the subjects had no opportunity to make any choices at all, so that subjects had equal exposure between options and could not form preference habits. Single-option trials provided the birds information about the properties of each alternative and served to collect data on their pecking patterns that were used to reveal their knowledge of the temporal delays programmed for each option. Single-option trials started with one key blinking (0.7 s on and 0.3 s off). The first peck caused the light to turn steadily on and the programmed delay to start lapsing. The first peck after the programmed delay had elapsed triggered the delivery of food, followed by an intertrial interval (ITI) of 50 s during which all keys were off. If no peck was registered during the 5 s after the programmed delay had lapsed the bird lost the reward, and the same option was presented after the ITI. This procedure equalized the number of rewards experienced per option.

In Experiment 1 probe trials in which the reward was omitted were included. Probe trials were as single-option trials but, after the fist peck, the light stayed on for 60 s and no reward was given. Two probe trials of each option were programmed to occur randomly in each session. This procedure, referred to as "peak procedure" (18, 19), typically leads to the observation of an increase in average pecking rate toward the time when a reward is normally due, followed by a decline after this time has passed with no reward, thus providing information about the birds' knowledge of the temporal properties of the options (immediacy to reward). In rewarded trials only the rising shoulder of this function is observed, as the trial ends at the time of the reward. Because the peak procedure analysis from Experiment 1 showed no distortion in the estimation of the immediacy to reward between options with the same programmed immediacy but different past utility ($B_{(10s)}$ and $C_{(10s)}$), probe trials were not included in Experiment 2 to simplify the experimental design. Instead, in Experiment 2 the rising shoulders of the pecking patterns were analyzed to study whether subjects were able to distinguish the options' intrinsic properties, namely their typical delay to reward, regardless of their preferences in the choice phase.

In Experiment 2 choice trials between options B and C were included before Training (phase Pre-Training) and after the Choice phase (phase Post-Choice) to examine whether, in the absence of contextual influences, the birds discriminated between the delays of the options comprising Test 3 (10 s and 14 s of delay). Choice trials started with two keys simultaneously blinking. The first peck on either of the keys caused the chosen key to turn steadily on and the other to turn off. After that, the trial continued as in single-option trials.

Choice phase. The choice phase lasted for 12 days for Experiment 1 and 6 days for Experiment 2 and was similar to the Training phase with the exception

that choice trials with simultaneous presentations of two options were included. In Experiment 1 the birds had cross-context pairings of B vs. C (Test 1) and A vs. C (Test 2), whereas in Experiment 2 they only had choices of B vs. C (Test 3). In Experiment 1, in the first 6 days of the choice phase (from day 1–6), half of the subjects were confronted with choice trials between options $A_{(5s)}$ and $C_{(10s)}$. For the other half, the choice trials were between options $B_{(10s)}$ and $C_{(10s)}$. During the following 6 days of the choice phase (from day 7–12), those subjects that previously chose between $A_{(5s)}$ and $C_{(10s)}$ had choices between $B_{(10s)}$ and $C_{(10s)}$, whereas those that had previously chosen between $B_{(10s)}$ and $C_{(10s)}$ were presented with choices between $A_{(5s)}$ and $C_{(10s)}$.

As shown in Table S1, 18 choice trials per day were presented in pairs separated by an ITI of 50 s. Each pair of choice trials were presented (*i*) before the beginning of each session (the session began 50s after the last choice trial of the pair was completed), (*ii*) after the end of each session (the first choice trial of the pair was presented 50 s after the session was completed), (*iii*) within each session, during the transitions between the contexts (50 s after the context presented in first place finished). This distribution of choices aimed at decreasing any possible impact of current context on choice, as we were interested in effects of memory.

State equalization. Because each context comprised options differing in reward delay, subjects in context CD received the same number of rewards as in context AB, but through longer trials. In Experiment 1, the ITI was 50 s, and the delays to reward were 5 s and 10 s in A and B, respectively, so that completing two trials in context AB took $2 \times 50\text{ s} + 5\text{ s} + 10\text{ s} = 115$ s, whereas in context CD (where C and D were 10 s and 20 s, respectively) it took $2 \times 50\text{s} + 10\text{ s} + 20\text{ s} = 130$ s, or 13% longer. In Experiment 2, with the delays in A and B being as in Experiment 1 but those in C and D being 14 s and 28 s, respectively, the difference is 23%. To compensate for the rate differences and control for energetic state without having to deliver potentially confounding food rewards, we adopted the following procedure: We calculated the maximum delay to food that could be experienced in context CD (the context offering the higher average delay), dividing the session in blocks of five trials, and in both contexts delivered supplementary delays at the end of each block so as to reach this maximum. For example, in Experiment 1 the maximum summed delay potentially experienced in a block was 100 s (because it could be possible that, by chance, option $D_{(20s)}$ appeared five times in a row). Thus, if in context CD the sequence of options in a block was, for example, $C_{(10s)}$–$D_{(20s)}$–$C_{(10s)}$–$D_{(20s)}$–$C_{(10s)}$ (where the total sum of delays is 70 s), then the supplementary delay given after the block finished was 30 s (that is, 100 s − 70 s). The same was done in context AB, where the maximum of 100 s was compared with the sum of the delays experienced in each block, and then the compensatory interval was added to the ITI after the block finished.

ACKNOWLEDGMENTS. We thank Cynthia Schuck-Paim for helpful discussion and Alasdair Houston, John McNamara, Sara Shettleworth, Miguel Angel Rodríguez-Gironés, Stuart West, Armando Machado, Wolfram Schultz and two helpful referees for valuable feedback on the manuscript. Financial support was received from Clarendon Fund from Oxford University Press (to L.P.) and Biotechnology and Biological Sciences Research Council Grants 43/S13483 and G007144/1 (to A.K.).

1. Bernoulli D (1954) Exposition of a new theory on the measurement of Risk, 1738. Translation. *Econometrica* 22:23–36.
2. Capaldi E, Myers DE, Champbell DH, Sheffer JD (1983) Conditioned flavor preferences based on hunger level during original flavor exposure. *Anim Learn Behav* 11:107–115.
3. Kacelnik A, Marsh B (2002) Cost can increase preference in starlings. *Anim Behav* 63: 245–250.
4. Schuck-Paim C, Pompilio L, Kacelnik A (2004) State-dependent decisions cause apparent violations of rationality in animal choice. *PLoS Biol* 2:2305–2315.
5. Marsh B, Schuck-Paim C, Kacelnik A (2004) Energetic state during learning affects foraging choices in starlings. *Behav Ecol* 15:396–399.
6. Aw JM, Holbrook RI, Burt de Perera T, Kacelnik A (2009) State-dependent valuation learning in fish: Banded tetras prefer stimuli associated with greater past deprivation. *Behav Processes* 81:333–336.
7. Pompilio L, Kacelnik A, Behmer S (2006) State-dependent learned valuation drives choice in an invertebrate. *Science* 311:1613–1615.
8. Tversky A, Kahneman D (1986) Rational choice and the framing of decisions. *J Bus* 59: 251–278.
9. Slovic P (1995) The construction of preference. *Am Psychol* 50:364–371.
10. Pompilio L, Kacelnik A (2005) State-dependent learning and suboptimal choice: When starlings prefer long over short delays to food. *Anim Behav* 70:571–578.
11. Shafir EB, Osherson DN, Smith EE (1993) The advantage model: A comparative theory of evaluation and choice under risk. *Organ Behav Hum Decis Process* 55:325–378.
12. Belke TW (1992) Stimulus preference and the transitivity of preference. *Anim Learn Behav* 20:401–406.
13. Gibbon J (1995) Dynamics of time matching: Arousal makes better seem worse. *Psychon Bull Rev* 2:208–215.
14. Gallistel CR, Gibbon J (2000) Time, rate and conditioning. *Psychol Rev* 107:289–344.
15. Wedell DH (1991) Distinguishing among models of contextually induced preference reversals. *J Exp Psychol Learn Mem Cogn* 17:767–778.
16. Shapiro MS, Siller S, Kacelnik A (2008) Simultaneous and sequential choice as a function of reward delay and magnitude: Normative, descriptive and process-based models tested in the European starling (Sturnus vulgaris). *J Exp Psychol Anim Behav Process* 34:75–93.
17. Bateson M (1993) Currencies for decision making: The foraging starling as a model animal. PhD dissertation (Univ of Oxford, Oxford).
18. Catania AC (1970) Reinforcement schedules and psychophysical judgments. *Theory of reinforcement schedules*, ed Schoenfeld WN (Appleton-Century-Crofts, New York), pp 1–42.
19. Roberts S (1981) Isolation of an internal clock. *J Exp Psychol Anim Behav Process* 7: 242–268.