Starlings uphold principles of economic rationality for delay and probability of reward

Tiago Monteiro¹, Marco Vasconcelos³,² and Alex Kacelnik¹

¹Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK
²School of Psychology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

Rationality principles are the bedrock of normative theories of decision-making in biology and microeconomics, but whereas in microeconomics, consistent choice underlies the notion of utility; in biology, the assumption of consistent selective pressures justifies modelling decision mechanisms as if they were designed to maximize fitness. In either case, violations of consistency contradict expectations and attract theoretical interest. Reported violations of rationality in non-humans include intransitivity (i.e. circular preferences) and lack of independence of irrelevant alternatives (changes in relative preference between options when embedded in different choice sets), but the extent to which these observations truly represent breaches of rationality is debatable. We tested both principles with starlings (Sturnus vulgaris), training subjects either with five options differing in food delay (exp. 1) or with six options differing in reward probability (exp. 2), before letting them choose repeatedly one option out of several binary and trinary sets of options. The starlings conformed to economic rationality on both tests, showing strong stochastic transitivity and no violation of the independence principle. These results endorse the rational choice and optimality approaches used in behavioural ecology, and highlight the need for functional and mechanistic enquiring when apparent violations of such principles are observed.

1. Introduction

The study of rationality principles of choice has prospered across disciplines (viz. behavioural ecology, experimental psychology and economics) and taxa (viz. slime moulds, humans, birds and insects [1–7]). This ubiquity reflects the importance of rationality principles for any self-consistent theory of decision-making. Unfortunately, theoretical interpretation of the significance of such research can be marrred by semantic difficulties, because rationality has different meanings in different fields. To mitigate these difficulties, one of us [8] introduced the terms PP-rationality, B-rationality and E-rationality to refer to the different meanings in psychology–philosophy, biology and economics, respectively. Here, we disregard PP-rationality, focusing on the other two, with emphasis on E-rationality, because B- and E-rationality are more amenable to experimental testing in non-human subjects, and because the criteria for E-rationality are fully specified and are shared across disciplines.

E-rationality is equivalent to the maximization of a stable utility function (itself defined as the maximand of E-rational agents’ behaviour), and this requires compliance with some deducible principles of choice, of which the main are completeness, transitivity and regularity [9–11]. Completeness requires well-defined preferences across all members of choice sets. Transitivity implies non-circularity of preferences; given A > B and B > C, then C > A is a violation (the symbol > is used to express ‘is preferred to’, and to discuss strength of preference). Matters would be simple if decision-makers showed exclusive preferences between options (as might be expected from rigorous maximization of outcomes), but frequently they do not, and the existence of partial preferences leads to the
formulation of elaborate principles concerning the strength of preferences. Luce’s [12] choice axiom defines degrees of stochastically transitive behaviour: weak stochastic transitivity (i.e. A ∼ C), moderate stochastic transitivity (i.e. A ∼ C being at least as strong as the weakest of A ∼ B and B ∼ C) and strong stochastic transitivity (i.e. A ∼ C at least as strong as the strongest of A ∼ B and B ∼ C). Regularity states that the overall proportion of choices for an option cannot increase when the choice set is enlarged [13]. Luce’s principle of independence of irrelevant alternatives (IIA) [12] further requires that the introduction of a new option into any choice set should not change relative preferences between pre-existing options.

Most experiments use sets of three options [3–5,14–17], establishing transitivity with binary combinations within the set, and/or evaluating the effects of introducing a third option. Added (irrelevant) options (e.g. decoys) are often equal or inferior to the focal options, probably to increase the likelihood of exposing violations of regularity, as well as of IIA [1,3,18–20]. Transitivity is the most scrutinized principle, and has been examined in a few non-human studies [1,15,16,21], whereas tests of IIA are rarer [5,22]. For example, Navarick & Fantino [23] discuss violations of transitivity in the concurrent chain procedure. The results are empirically robust, but Houston et al. [24] subsequently proved that rate maximization can, in principle, be compatible with such violations owing to properties of the concurrent chains procedure, and hence would be rational after all. Nevertheless, Houston [25,26] later wrote that rate maximization might not explain the concurrent chain findings, arguing, and we agree, that rate maximization is itself violated in other experimental protocols.

Apparent irrationalities can result from unplanned state-dependence effects [20,27–30], and they do not pose theoretical challenges to optimality approaches. However, state-dependence cannot explain all observations of apparently irrational behaviour [14]. Violations of rationality that are not procedural artefacts tend to be attributed to suitable mechanisms of choice deduced post hoc from the observed choices or to the subjects’ inferring future opportunities from the options they face [31], but the former does not solve the functional issues and the latter is a form of state-dependence. Prominent examples of mechanistic accounts are the inferred use of comparative, as opposed to absolute choice processes. They include algorithms where the decision-maker counts the dimensions in which one option beats the alternative, rather than weighing the perceived differences in a common currency [1,3,6,18,32] and irrational context-dependence [2,4,5,14,19,33,34]. However, if choice mechanisms caused widespread breaches of rationality in problems of ecological relevance, the adaptationalist perspective in behavioural ecology and the normative logic of microeconomics would be severely threatened.

To accommodate apparent irrationalities, evolutionary theoreticians argue that there are, in fact, no violations at the level of state-dependent choice strategies, but considering preferences out of context can generate intransitive (but state-dependent and artefactual) choices [35]. When alternatives differ in more than one dimension simultaneously, apparent irrationalities can also emerge from nonlinearities in the combination of dimensions such as delay, probability, amount, energetic concentration of rewards or perceived risk into a single utility metric. With this in mind, this study aims at testing qualitatively and quantitatively the principles of transitivity and IIA, while (i) controlling for possible state-dependence artefacts and (ii) using alternatives that vary in only one dimension, so as to make a priori predictions more transparent. We designed two experiments to test transitivity across binary-choice sets (i.e. choices where two options are simultaneously available) as well as evaluating the effect of introducing a third option to the set (i.e. creating trinary choices where three different options were simultaneously available). We used multiple options (figure 1a), combined in binary and trinary sets according to the principle under test (figure 1b). In contrast to previous studies, here we added options that were inferior, intermediate and superior to one or both options in each binary set. We chose two dimensions of reward, delay (exp. 1) and probability (exp. 2), each for different reasons. Delay to food is perhaps the only manipulation where it is possible to test whether animals have accurate representations of the properties of each option. Extensive previous work on interval timing shows that animals (including starlings) learn accurately the time when rewards are due, hence it is possible to relate preferences to knowledge, avoiding confounding whether behaviour is due to the agent’s available information or to its strategic preferences (this is crucial for mechanistic accounts of context-dependent effects [30,36]). Probability of reward, in turn, has special interest because humans are notoriously unable to treat probability rationally, especially when communicated verbally rather than by experience. Irrationalities such as the Allais paradox, or over- and undervaluation of extreme probabilities are common, and have been extensively discussed [37–39]. Little is known about non-human perception of probability.

2. Methods

(a) Subjects and apparatus

Subjects were 20 adult wild-caught European starlings (Sturnus vulgaris; Natural England licences nos. 20083718 and 20093673). Eight birds initiated exp. 1 and 12 exp. 2, of which five and 10 completed the experiment, respectively. Birds had experience pecking at keys but not with the contingencies and symbols

Figure 1. Experimental design. (a) Options used in exp. 1 (left) and exp. 2 (right). (b) Example of trials used to test stochastic transitivity (left) and independence of irrelevant alternatives (right). The black arrows show comparisons relevant to testing the principles identified next to each arrow.
used in this study. Prior to testing, birds were socially housed in outdoor aviaries with ad libitum food and water. For the experiment, subjects were housed in pairs in indoor cages (1350 × 784 × 800 mm (l × w × b)) and kept under a 12:12 L:12 D cycle with gradual transitions at dawn and dusk. They were confined individually to one section of their cage for tests, but after each session, they had 4 h of ad libitum food and social interaction with the cage-partner. The experiments were run from April to June (exp. 1) and from October to December 2010 (exp. 2). On completion, the birds were reintroduced into communal aviaries and eventually released into the wild. We followed animal welfare regulations of Oxford University and maintenance and experimental protocols adhered to guidelines for the use of animals in research from the Association for the Study of Animal Behaviour/Animal Behaviour Society [40]. The experimental cages had two side areas that could be split with sliding panels from a shared middle section. Each side area had attached a working panel with three sections, each with a centrally placed response key. The central section of the side panels also had a food hopper connected to a pellet dispenser (Campden Instruments) containing 20 mg BioServ precision dustless pellets. Contingencies were controlled by custom software running on the Microsoft Windows operating system and attached to an Animal Behaviour Environment Test System (Campden Instruments) via WhiskerServer.

(b) Preliminary training

After training to peck keys and obtain food, the birds received either response-initiated fixed interval training (exp. 1) or probability training (exp. 2). Trials started with an attention key lighting up. In exp. 1, pecking the central attention key triggered a delay, and the first peck after the programmed delay caused the key light to switch off and two food pellets to be delivered. In exp. 2, a single peck to a flashing side key was followed by reinforcement with a programmed probability.

(c) Trial structure

Both experiments included training and testing phases. There were three types of trials: single-option, choice and peak trials. Single-option trials began with a flashing attention key (location randomly assigned). Pecking at the flashing key turned it OFF, and caused one of the three available keys (randomly assigned) to begin flashing (700 ms ON, 300 ms OFF). A peak to the illuminated flashing key either turned the key steadily ON and initiated a delay to food (exp. 1) or turned the key OFF and delivered food with a certain probability (exp. 2). Choice trials offered simultaneous choices between two or three options. All combinations of two and three different options were used. Trials began with a flashing attention key at a random location. A peak to it turned it OFF and began the flashing of two (binary choices) or three (trinary choices) randomly located keys (each displaying a different symbol). A peak to one of the flashing keys initiated its programmed contingency, and turned the unselected key(s) OFF. Peak trials were similar to single-option trials and were used to determine the animals’ knowledge of the delay to reward associated with each option in exp. 1. In these trials, the chosen option remained ON for three times its typical delay and then extinguished without reward. Successive trials were separated by a 45 s (exp. 1) or 30 s (exp. 2) inter-trial interval (ITI).

(d) Particulars of experiment 1

There were five different options that differed in delay to food reward (6, 9, 12, 15 or 18 s; figure 1a), each signalled by a different colour/symbol combination. The experiment lasted for 35 days. After 7 days of no-choice training (150 daily single-option trials, 30 per option), the birds were randomly assigned to two groups, and experienced 7 days of choice training wherein a number of choice trials with binary or trinary presentations of the five options were interspersed among single-option trials (90 single-option trials, 18 per option and 60 choice trials, six per pair or trio combination, daily). One group faced first all possible binary combinations of the different options (10 binary choices), whereas the other group faced all possible trinary combinations (10 trinary choices). After the first choice-training period, no-choice training resumed for 7 days before the choice-training contingencies were reversed between groups and the procedure repeated. Finally, birds were tested for 7 days, with both binary- and trinary-choice trials interspersed between no-choice trials (75 single-option trials, 15 per option, 60 binary-choice trials and 60 trinary-choice trials). In summary, one group received the sequence ‘training → binary → training → trinary → training → binary + trinary’, whereas the other the sequence ‘training → trinary → training → binary → training → binary + trinary’. If a bird failed to peck any of the flashing keys for 120 s after onset, then the key was turned off and the ITI started, followed appropriately by the next trial.

(e) Particulars of experiment 2

Birds learned about six options each signalling one of five different probabilities of food reward (viz. 0.8, 0.7, 0.5, 0.3 and 0.2, with 0.5 signalled by two options; figure 1b). They first encountered single-option trials for 12 days (414 single-option trials daily, 69 per option). Afterwards, birds were tested for 30 days with single-option trials interspersed with choice trials of all possible binary (15), and trinary (20) combinations of different options (216 single-option trials, 36 per option, 90 binary-choice trials, six per pair and 120 trinary-choice trials, six per trio). The attention key was removed during the last 7 days of the experiment to eliminate possible influence of the location of the attention key on the subsequent choice. Trials proceeded as in exp. 1, but the programme did not advance to the next trial unless the animal responded.

(f) Data analysis

All proportion data were successfully normalized, using an arcsine square-root transformation [41]. An alpha level of 0.05 was adopted for all statistical comparisons. For exp. 1, only the test sessions (i.e. last 7 days) were considered in the analysis (see the electronic supplementary material, dataset S1). For exp. 2, the last 3 days of testing were used (see the electronic supplementary material, dataset S2). In both experiments, we set an experimental criterion of 80 per cent completion of all daily trials, for a session to be included in the analysis. Five birds were excluded from all analyses after repeatedly failing to reach the experimental criterion (four birds, three from exp. 1 and one from exp. 2) or due to showing a strong side preference (one bird from exp. 2).

Testing for stochastic transitivity requires comparisons of strength of preferences in binary choices. Given inter-individual variation, within each transitivity set (TS), we identified the pair with higher statistical significance (smaller p-value) as the stronger preference, and used it to test strong stochastic transitivity, whereas the other pair was used to test moderate stochastic transitivity. Analyses of quantitative variations in relative preferences are viable only between options that are chosen at least once; hence, we discarded a minority of cases when preferences were absolute. This occurred three times of 150 possible ones in exp. 1 (distributed in two subjects) and 24 times of 600 in exp. 2 (one subject missed 5, one 4, three 3, two 2 and two 1). Weighted least-squares linear regressions were calculated both for the average and individual data. Number of choices
between every two options in each trinary-choice combination was used as weighing factor.

(g) Notation
To facilitate description and visualization of the data, in the figures and parts of the text option values are replaced by letters, from A to E in exp. 1 (A > B > C > D > E) and from A to F in exp. 2 (A > B > C = D > E > F), where A is the most profitable option.

3. Results

(a) Stochastic transitivity
We define a ‘TS’ as a combination of all possible binary choices between three different options. Figure 2 shows the proportion of choices observed for the better option (shortest delay or highest probability) in all binary-choice combinations of ‘TSs’ composed only of contiguous options for exps. 1 and 2 (figure 2a,b, respectively; electronic supplementary material, tables S1 and S2 show all binary-choice combinations). To test for stochastic transitivity, we compared preference for the best option in the extreme pair (red bars) of each ‘TS’ with preference in the intermediate pair (blue bars).

An inspection of figure 2 (and electronic supplementary material, tables S1 and S2) reveals that the requirement for moderate stochastic transitivity was met for all ‘TSs’, with preference for each extreme pair being at least as strong as the weakest of those found in intermediate pairs. In fact, preference in the extreme pair was significantly stronger than the weakest preference found in the intermediate pairs in seven of 10 ‘TSs’ in exp. 1 (smallest significant t(4) = 2.953, p = 0.042) and in 19 of 20 ‘TSs’ in exp. 2 (smallest significant t(9) = 2.505, p = 0.034). With respect to strong stochastic transitivity, in exp. 1, 10 of 10 comparisons were in the required direction (binomial p = 0.002), but these differences were not individually significant (all ts(4) ≥ 0.057 and ≤ 2.392, n.s.). In exp. 2, 17 of 20 comparisons were also in the required direction (there was one tie; binomial p = 0.001) and of those eight were statistically significant while none was significant in the opposite direction, which would have been necessary to prove a violation (all t(9) ≥ 2.462, p ≤ 0.036). The results thus confirm statistically reliable strong stochastic transitivity for both delay and probability of reward.

(b) Independence of irrelevant alternatives
The IIA principle encompasses regularity, so we use it to examine both, enlarging choice sets with inferior, intermediate and superior alternatives. IIA requires demonstrating no change in relative proportion, namely the absence of an effect, and is thus statistically more demanding. We calculated the relative proportion of choices for the better of two options in trinary-choice trials as a function of the proportion of choices between the same two options when met in binary-choice trials (figure 3a,b for exps. 1 and 2, respectively) as follows:

$$\text{RelProp}_{A_B C} = \frac{A_{ABC} - B_{ABC}}{A_{ABC} + B_{ABC}}. \quad (3.1)$$

where RelProp_{A_B C} is the relative proportion of choices for option A versus option B in trinary [ABC] trials, and A_{ABC} and B_{ABC} are the number of choices for A and B in those trials. We display the effect of adding an option (e.g. C) to a binary choice (e.g. [AB]), plotting on the abcissa relative preference in binary choices (e.g. A over B in [AB]), and on the ordinate relative preference in trinary choices (e.g. relative preference for A over B in [ABC]). For IIA to hold, values should fall along the diagonal. With the exception of two outliers, this is the case regardless of the quality of the added third alternative (individual compliance with IIA is shown in the electronic supplementary material, figure S1).
We did not use alternatives differing in more than one dimension simultaneously. Such tests are complicated because the scaling of utility (or preference) to physical dimensions probably includes nonlinearities when different properties are traded against each other [1,3,6,18,32]. Demonstrations of full rationality in unidimensional choices such as those shown here and results obtained in transitive inference experiments across multiple species that are consistent with these results [42–48] suggest that rational choice, rather than its opposites, is widespread, and should be the foundation from which to interpret observations of logically inconsistent behaviour.

Our view is that most reports of apparent violations of logical decision principles in non-human studies result from failing to follow preconditions for their validity, such as constancy of the agent’s physiological or informational state, or lack of satiety effects from the commodity [10,27,29–31,35,36]. Variations in preference following variations in subjects’ energetic reserves, or when the testing conditions allow subjects to infer differences in their circumstances, are perfectly consistent with evolutionarily normative, rationality-based theories—as is clear from Houston et al. [31] and Houston [26]. This matters because if it were convincingly shown that when necessary conditions are controlled, logical principles do not apply to decision processes, the foundation of normative modelling in behavioural biology would melt away. This is relevant to decision-making across multiple taxa, including humans, and highlights the value of integrating decision research across economics, psychology and biology.

This research was supported by grant no. BB/G007144/1 from the UK Biotechnology and Biological Sciences Research Council (BBSRC) to A.K. T.M. was supported by a doctoral grant from the Portuguese Foundation for Science and Technology and a Pembroke College Graduate Scholarship. M.V. was successively funded by a Marie Curie Intra European Fellowship within the Seventh European Community Framework Programme, the abovementioned BBSRC grant and a post-doctoral grant (no. SFRH/BPD/79311/2011) from the Portuguese Foundation for Science and Technology. T.M., M.V. and A.K. jointly designed the experiments, T.M. collected the data, T.M. and M.V. processed and analysed the data, T.M., M.V. and A.K. jointly wrote the manuscript. We thank Alan Grafen and Justine Aw for their comments on earlier versions of this manuscript, and Simon and Jan Parry for trapping our participants.