

Midsession Reversal Task with Pigeons: Parallel Processing of Alternatives Explains Choices

Aaron P. Smith¹, Thomas R. Zentall^{1*}, and Alex Kacelnik²

1: Department of Psychology, University of Kentucky, Lexington, KY 40506-0044, USA

2: Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

* Corresponding author:

Thomas R. Zentall

Department of Psychology

University of Kentucky

Lexington, KY 40506-0044

Phone: 859-257-4076

Fax: 859-323-1979

Email: zentall@uky.edu

Abstract

Most models of choice assume a “tug of war” (ToW) between options present at the time of the choice, arguing that preferences are built in this process, and implying that adding options increases delay to act. In contrast, the Sequential Choice Model (SCM) proposes that choices are driven by parallel expression of the mechanisms that control action in sequential encounters, without comparative deliberation at choice time. Only the SCM predicts choice preferences based on latencies to respond in single-option encounters. SCM further predicts that latencies to choose should either be the same or shorter than those in sequential encounters. We contrasted these models using a midsession reversal task with pigeons. Responses to one alternative (S1) were rewarded in the first half of each session and those to the other (S2) in the second half. Single-option (sequential) and two-option (choice) trials were intermingled. In choice trials subjects strongly preferred S1 early in the session, showed intermediate preferences towards the mid-session, and preferred S2 late. These preferences were all predicted by changes in latency towards the presently negative alternative (S2 early and S1 late) in single-option trials. Latency towards presently positive stimuli were minimal throughout, in both single and two option trials, with no evidence of an evaluation time cost of choice. The ability to predict choice preference from latencies in sequential encounters and the absence of a choice delay support the SCM against ToW models, consistently with results from other protocols and species.

Keywords: choice, sequential choice model, race model, diffusion model, midsession reversal, pigeons, sequential vs. simultaneous encounters.

Midsession Reversal Task with Pigeons: Parallel Processing of Alternatives Explains Choices

Choice and the cognition behind it are core topics in behavioral, cognitive, and decision studies, and one important unresolved question is whether animals possess cognitive mechanisms specifically adapted to support optimal choices. Using an experimental protocol in which the outcomes of responding to either of two stimuli vary systematically over the course of each session, we ask whether pigeons deploy dedicated choice mechanisms or simply respond to each stimulus with the mechanisms that elicit action when they encounter alternatives individually.

We use our results to address and compare hypotheses addressing choice mechanisms. Some of these hypotheses argue that decision-takers construct preferences through relative value judgements at the time of choosing, while others support the view that choice behavior is explained by mechanisms driving responding in sequential encounters, with no relative judgements being made at choice time. Our treatment of the problem is within a specific experimental context, but the distinction between these two classes of model, that have been called ‘Tug of War’ (ToW) and ‘Sequential Choice Model’ (SCM) respectively (Kacelnik Vasconcelos, Monteiro, & Aw, 2011) is of widespread significance for the study of choice and preferences. More information on the distinction between theories about the determination of preferences can be found for humans in Shafir, Simonson, and Tversky, (1993) and for non-humans in Kacelnik et al. (2011).

Scholars from different behavioral sciences, including microeconomics, behavioral economics, experimental psychology and behavioral ecology, often deal with similar phenomena, but build theoretical models that differ greatly in foundations, structure, and predictions. For instance, assumptions of rational choice in microeconomics (viz. Gravelle and Rees 2004) imply the need to rank all available options, such that the greater the number of alternatives, and/or the harder to discriminate their outcomes, the longer it should take to act.

It is common for researchers to place emphasis on the computational or time-consuming costs of relative evaluations, and to generate predictions accordingly. For instance, Schwartz and Cheek (2017) have

argued that because of such costs, "... too much choice can lead to paralysis, bad decisions, and dissatisfaction with even good decisions", and the ABC group led by Gigerenzer predicate that information-frugal heuristics drive choices more efficiently than exhaustive evaluations (Gigerenzer, Hertwig, & Pachur, 2011; Gigerenzer, Todd, & The ABC Research Group, 1999; Simon, 1991). These various approaches share the expectation that, for a given level of performance, time to act must increase with number and similarity (i.e. discrimination difficulty) of the alternatives in the choice set. This also applies to the second of the two systems that, in Kahneman's 'Dual Process Account of Mind' (viz. Kahneman 2011) deal with difficult, complex decisions, but not with his 'System 1' which is meant to control decisions when speed is at a premium and there is no deliberation. It would make sense for system 1 to process alternatives in parallel, as implemented in the SCM.

There are some empirical reasons to argue that preferences are meaningful only in the context of choice: in humans, expressed preference differs as a function of how the options are presented and questions asked, rather than exclusively on the options' properties. For instance, in choosing between lotteries, subjects can be nudged to give greater weight to the probabilities or to the outcome magnitudes, sometimes reversing their preference depending on how the choice is framed (Slovic 1995). Since preference depends on how the question is put, then it could be inferred that some cognitive evaluation of differences occurs when facing the decision, and this may be expected to add some time to the generation of action. We don't have equivalent data for birds, but if they too built their preferences at choice time, then their choices should not be predictable from observations made when subjects face single options, thus contradicting the main SCM prediction.

The widespread assumption of ToW models, that choosing must involve cognitive operations with corresponding time costs, has a long history in psychophysics, where it is quantified by the Hick-Hyman Law (Hick, 1952; Hyman, 1953) that states that time to reach decisions increases proportionally with the logarithm of the number of options in the choice set. Starting from the assumption that choice must incur such a time cost, Pelé and Sueur (2013) discuss how 'Race' and 'Diffusion' models handle speed-accuracy trade-offs (their nomenclature maps to our labels of SCM and a specific version of ToW respectively). Their hypothesis is that

when agents are exposed to a set of alternatives, they accumulate evidence supporting each member of the choice set. In their race model, action takes place as soon as support for one alternative reaches some threshold, while in their diffusion model, action occurs when the difference in support between the alternatives exceeds some critical value. Pelé and Sueur favour the diffusion model because it can adjust the optimal speed of decision making for a required level of accuracy, a tradeoff that the race model does not address. Their less-preferred race model is structurally equivalent to that proposed by Shapiro et al. (2008) and further examined by Kacelnik et al. (2011). However, like most other researchers of choice behavior, Pelé and Sueur do not discuss what the different models predict when comparing time to act between sequential and simultaneous encounters with the same options. Shapiro et al. (2008) and Kacelnik et al. (2011) point out that the Sequential Choice Model (SCM) does not predict an increase in delay to act when the subject faces two alternatives as opposed to a single one. In contrast, the diffusion model concurs with all other ToW models in implicitly predicting that when comparing action towards a single alternative versus choosing between two of them, the time to act should increase in the latter, as it should take longer for the difference in evidence between alternatives to reach the action threshold.

Here we will argue that the idea that choice must necessarily involve a time-consuming comparative evaluation underlying the construction of preference at the time of choice is not empirically supported in a variety of experimental systems (Kacelnik et al. 2011; Shapiro, Siller, & Kacelnik, 2008). According to the SCM, behavior in sequential (one-option) encounters provides sufficient information to predict behavior when subjects face a choice between multiple options, thus opposing the view that preference is constructed when the subject is facing multiple options and choosing between them. ToW models do not explicitly address this prediction, as they focus only on choices. Further, according to the SCM, making choices should not involve extra time, because the options are processed in parallel. SCM's central argument is close in spirit to the 'ecological rationality' idea influential in human decision theories (Gigerenzer et al, 2011; Gigerenzer et al. 1999; Smith, 2002). The SCM is a mechanistic hypothesis, but its normative foundation is that the mechanisms that

drive action are adapted to natural circumstances in which simultaneous choices are rare, and consequently animals act in choice contexts with mechanisms evolved to control responding towards opportunities met sequentially. In sequential encounters, the decision is whether to pursue the current opportunity or to continue searching in the environment. Rate maximizing models in classical Optimal Foraging Theory (Stephens & Krebs, 1986) are based on such conceptualizations, but because their origin is purely normative, they do not address delays to act, which optimal foraging models implicitly assume to be negligible.

According to the SCM, the time taken to start pursuing an opportunity is a noisy variable that reflects the stimulus' (subjective) value relative to the background. The value of each option is thus not exclusive of simultaneous choice. By taking longer on average to act when the value of a stimulus relative to its background is lower, the agent increases the probability of detecting a preferable alternative. For each kind of stimulus encountered on its own, an observer can record a frequency distribution of latency to pursue, whose central value, and most likely its spread, should be inversely related to the background-dependent benefits associated with pursuing that prey.

The SCM does not include any special mechanism dedicated to simultaneous choice: when two stimuli are present simultaneously, each elicits sampling of the same frequency distribution of latencies that each stimulus would elicit when alone, and action is directed towards the alternative that elicits the shorter latency. It is, thus, a 'race' model, but it makes unique predictions.

First, the probability of choosing each stimulus in a simultaneous set should be predictable from the latency distributions of the same set of stimuli when each is encountered alone. This would of course be impossible if preference were a meaningless concept outside the context of choice (Slovic, 1995).

Second, because latency to respond in simultaneous encounters is not due to computing differences or ratios between evidence favouring the different options, but to the fact that the shortest latency leads to action, latencies to respond in simultaneous encounters are predictable from the latency distributions in sequential encounters. In an important contrast with ToW models, this implies that behavior in sequential

encounters has greater informative richness than in simultaneous choices, as the latter cannot predict the former. According to the SCM, choice results from cross-censorship between latency distributions. Thus, latencies to act in choice encounters should not be longer than those for each stimulus when encountered on its own, and if anything should be shorter.

Third, inconsistent with the prediction of the ToW's model, latencies to act should not increase when the choice is more difficult, namely when the options are more similar to each other. An extreme version of the conventional ToW would leave Buridan's ass starving between two identical haystacks (viz. Lamport 2012), as when options are equal it would take the animal an infinite time to choose. An ass governed by the SCM would choose one of the haystacks randomly, and would respond faster than when facing a single haystack.

In our experiment, half-way through each daily session, pigeons (*Columba livia*) faced a reversal in the reward contingencies of two stimuli. This procedure has been used with other variants, and is known as a midsession reversal (Cook & Rosen, 2010; Rayburn-Reeves & Cook, 2016; Rayburn-Reeves, Molet, & Zentall, 2011; McMillan & Roberts, 2012; McMillan, Sturdy, & Spetch, 2015; Smith, Beckmann, & Zentall, 2017)]. The common feature of most variants (although not that used by Cook & Rosen, 2010) is presenting the subject with trials showing two stimuli (S1 and S2). In the first half of each session, only responding to stimulus S1 leads to a food reward, while in the second half of the session the opposite is true, so that food is delivered if and only if the subject chooses S2. The typical finding with this procedure is that pigeons choose S1 for most of the first half of each session but begin to choose S2 prior to the reversal (they make anticipatory errors) and they continue to make errors shortly after the reversal (they make perseverative errors), before shifting to an almost exclusive preference for S2.

To ascertain the source of these errors, McMillan, Sturdy, and Spetch (2015) trained pigeons on a single-stimulus go/no-go version of the midsession reversal task, namely with only what we call sequential encounters. They found that pigeons rarely failed to respond to a stimulus when responding to that stimulus was rewarded, but sometimes did respond to a stimulus when such responses to it were not rewarded,

especially in a region near the middle of the session. While such responses are ‘errors’ from the perspective of a fully informed observer, they may be the outcome of an optimal strategy when available information about reward probability is constrained, as we discuss later. MacMillan et al. (2015) reported that the discrimination ratio obtained from these single-stimulus trials correlated reasonably well with the choice data obtained in other experiments, and noticed that the Sequential Choice Model would predict such a correlation, but they did not address the issue of latency to respond on those choice trials. In the present study, we interspersed single option encounters with choice trials, and measured all relevant latencies, so as to have the necessary information to compare the predictions of Tug of War and Sequential Choice models. The contrast between predictions of these families of models allows us to further discuss the evidence for existence of special mechanisms adapted for simultaneous alternatives.

Method

Subjects

Ten experimentally non-naïve White Carneau pigeons originally purchased from the Palmetto Pigeon Plant (Sumter, SC) were used in the experiment. The pigeons had prior experience making simultaneous and successive color discriminations. They were approximately 8-12 years old at the start of training and were maintained on a restricted diet of 80-85% of their *ad libitum* weight. The pigeons were housed in individual wired cages measuring 28 × 38 × 30.5 cm on a 12-hour light-dark cycle (lights off 7 pm) with free access to grit and water. The pigeons were cared for in accordance with University of Kentucky animal care guidelines. PSA Assurance #D16-00217 (A3336-01).

Apparatus

The experiment was conducted in a standard LVE/BRS (Laurel, MD) test chamber measuring 34 cm high, 30 cm from the response panel to the back wall, and 35 cm across the response panel. Response panels had three response keys, each 2.5 cm in diameter and 6.0 cm apart. The bottom edge of the response keys was 24 cm from the wire-mesh floor. A 12-stimulus inline projector behind each key projected one of three stimuli

(white, yellow, or blue) onto the response keys. A center-mounted feeder located 10 cm above the floor allowed access to mix grain when raised, and was then illuminated by a 28 V 0.04 A lamp. White noise was generated from outside the chamber and the procedure was controlled by a computer in an adjacent room running Med-PC IV (Tatham, & Zurn, 1989) with a 10-ms resolution.

Procedure

Pretraining. Pigeons were first trained using an autoshaping procedure in which one of two stimuli (yellow or blue) appeared on either the left or right response key, or a white stimulus appeared on the center key. The stimuli remained until either 30 s had elapsed or a pecking response was made, whichever came first, after which the stimulus went off and the feeder was raised for 2 s. Training continued until a peck to each stimulus occurred on 95% of the trials for two consecutive sessions.

Training/testing. Each session was composed of 80 trials organized in 16 blocks of 5 trials each, 4 single-stimulus trials and 1 choice trial. All trials began with a white orienting stimulus appearing on the center key. On single-stimulus ('sequential') trials, a response to the center key produced either S1 or S2 randomly on the right or left response key. The stimulus remained on until either a peck occurred or 5 s elapsed. If the stimulus was the currently rewarded one (S1 for the first 8 five-trial blocks and S2 for the last 8 five-trial blocks), the stimulus was turned off and the feeder was raised (made available) for 2 s. If the stimulus was not the currently rewarded one, or 5 s had elapsed without a response, the stimulus was turned off without reinforcement. On choice trials, a response to the center key produced both S1 and S2 randomly on the left and right-side keys. Both stimuli remained on until a 'choice' response to one of them was made. If the currently rewarded stimulus was chosen, both stimuli were turned off and reinforcement was delivered as described earlier; if the incorrect stimulus was chosen, both stimuli were turned off without reinforcement.

Trials were separated by a dark (house lights off) intertrial interval (ITI) that varied in duration, such that the time between pecking the orienting key and the next trial's onset was always 10 s. Pigeons received 80 sessions with this procedure.

Data Analysis

To ensure that stability had been reached, all analyses were conducted on the data from the last 10 sessions of training. In order to appropriately handle the repeated time-series design of the current experiment, grouped data were analyzed using linear and non-linear mixed effects models (Young, Clark, Goffus, & Hoane, 2009) using the nlme package in R (Pinheiro, Bates, DebRoy, Team RC, 2016). In all mixed effects models, subject was treated as a random effect while specific model parameters (defined below) were fixed factors. In the case of fitting data on a subject-by-subject level, we used non-linear regression in MATLAB (The Mathworks; Natick, MA, USA) for the specific models (defined below). Model selection was based on utility for describing the trends within the data and Akaike information criterion (Wagenmakers, & Farrell, 2004). SCM predictions (detailed below) and figures were averaged over sessions (in the case of individual data) or averaged over both sessions and subjects (in the case of group data including all subjects).

Results

Latencies to respond are shown in Figure 1, separately for S1 (Top) and S2 (Bottom). We call single-option trials in which responding did not occur within 5 s (see frequency in Figure S1) no-go, but assign them a latency of 5 s for computations. This underestimates potential latencies, but recognizes that a response could (and probably would) have occurred sometime beyond the guillotine of 5 s imposed by our procedure. No-go trials happened frequently early in the session for stimulus S2 and late in the session for stimulus S1 (Figure S1), and did not occur when the stimuli were being rewarded. Due to this assignment, the real variation in latency length as a function of block number must be more step-wise than it looks in Figure 1. The figure shows that before midsession for S1, and after midsession for S2, latencies were constantly very short, at just over 1 s. At no point in the session were latencies in choice trials longer than in single-option trials, as would have been expected if choosing involved time-consuming cognitive operations to evaluate the difference between the options, as predicated by ToW models.

Confirming this visual impression, a linear mixed effects analysis including only the blocks in which all individuals contributed data to choice trials (i.e., had less than exclusive preference for the alternative stimulus) yielded no effect of trial type (single vs. choice) on latencies, $F(1, 14.88) = 1.43$, $p = .251$, no interaction between trial type and block, $F(1, 6.30) = 3.32$, $p = .116$, and no three-way interaction, $F(1, 7.38) = 0.87$, $p = .380$. Furthermore, a model not allowing latencies to differ between single and choice trial types was found to be 4531.46 times more likely (using corrected Akaike Information Criterion; Burnham, Anderson, & Huyvaert, 2011) than one allowing the two trial types to differ. Visual inspection of all blocks for which there were at least some choices for the relevant stimulus shows that mean latencies were shorter in choice trials than in single-option trials. This latency shortening (predicted by SCM) can only be judged qualitatively, because, as Table S1 shows, there were few choices for the presently negative stimulus (S1 after midsession and S2 before midsession, identified by grey symbols in Figure 1), when the shortening effect is more pronounced.

The SCM can be implemented for the present study through the following assumptions: (i) when only one stimulus is present, latency to respond is a noisy decreasing function of the animal's estimated probability \hat{P}_{si} that a response to stimulus i will produce food; (ii) when more than one stimulus is present, a 'choice' response is elicited towards the stimulus evoking the shortest latency, sampled from the same functions as when only one option is present; (iii) For each stimulus, in the trial immediately after a successful response, $\hat{P}_{si} \sim 1$, and consequently latency to respond is minimal (this is a sort of Win-Stay, or 'follow the reward' assumption); (iv) For an animal experienced with the midsession protocol, success and failure have different effects for each stimulus: after an unsuccessful response, \hat{P}_{s1} decreases while \hat{P}_{s2} increases as a function of time (or trials) since the last success. Assumptions i, ii, and iii are suitable for many animal choice experiments, and is where the contrast with the alternative hypotheses lies. Assumption iv is the only one specific to the midsession protocol.

For S1, early in the session all responses are rewarded, hence \hat{P}_{s1} remains constantly high until midsession (assumption iii). After midsession, time since the last successful S1 response starts to increase and

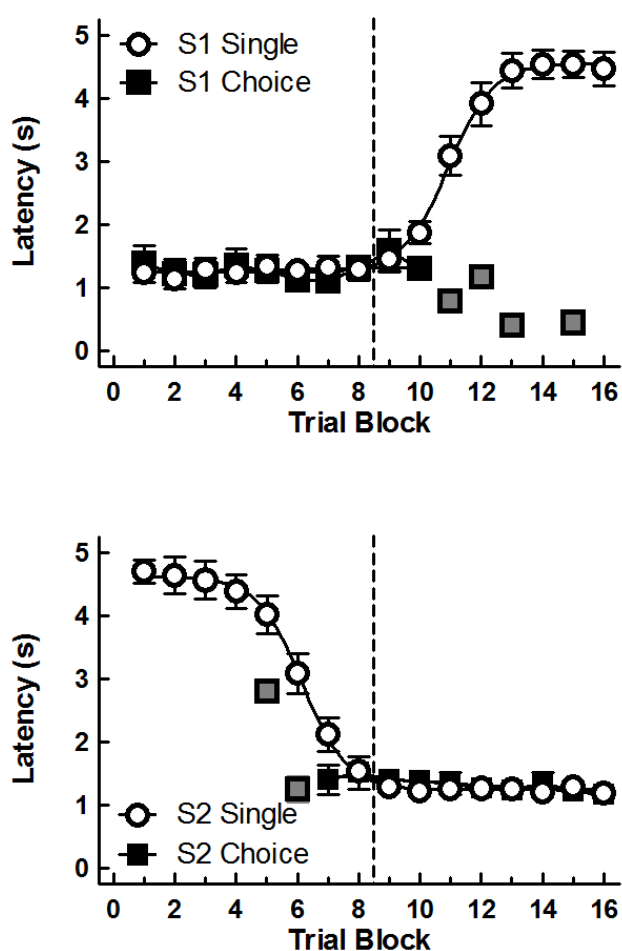


Figure 1. Mean (\pm SEM) latency to respond to S1 (Top) and S2 (bottom) over the last 10 sessions of the experiment, as a function of trial block. The plot shows latencies in single stimulus trials (open circles) and in choice trials in which the stimulus was chosen (filled squares). Grey squares identify blocks in which not all individuals contributed data (as some of them showed exclusive preference for the alternative stimulus). Number of individuals contributing to each choice data point are presented in table S1. The dotted line indicates the point of the reversal.

as a consequence \hat{P}_{s1} gradually drops (assumption iv). This causes latency to respond to S1 to be short and flat until midsession, and to gradually increase after that. For S2, early in the session the last rewarded response dates from the previous session, and as session time proceeds, \hat{P}_{s2} gradually increases and consequently latencies to S2 gradually drop (assumption iv). After midsession all responses to S2 produce food, hence \hat{P}_{s2} remains constantly high and latencies to S2 are short and constant (assumption iii). This interpretation is consistent with all details of Figure 1. Mean (\pm SEM) proportion 'Go' response to each stimulus on single-stimulus trials, averaged over the last 10 sessions as a function of trial block are presented in Supplementary Figure 1.

A behaviorist narrative could prescind from postulating probability estimates, since they are only observable through behavior, but the rationale for using a reward probability estimate as an intervening variable is that behavioral sensitivity to probability of success is to be expected from a normative (reward maximizing) perspective. Assuming the presence of probability estimates is useful to analyze sequential foraging encounters in the wild, to address partial reinforcement versions of the present experiment, and to generalize our rationale to any protocol in which information about reward probability is available to subjects. The variation in response latencies to each alternative when encountered on its own through the session remains the key element. In the next section we turn to choice trials.

Simultaneous trials

The proportion of responses to S1 in choice trials is shown in Figure 2. Consistently with other midsession reversal experiments (e.g. Rayburn-Reeves et al. 2011, Rayburn-Reeves & Cook 2016), proportion of choices of S1 started high and gradually diminished as sessions progressed. The data, plotted as proportion of responses to S1 as a function of trial block, are well described by a decreasing, symmetric, sigmoidal function (Figure 2 top, and caption) with inflexion at the approximate midsession point. Figure 2 (bottom) shows that the probability of choosing the presently unrewarded stimulus (i.e. the probability of error) peaks at the midsession

point, and includes both anticipatory errors (S2 before midsession) and perseverative errors (S1 after midsession). See fitting equation in the figure's caption.

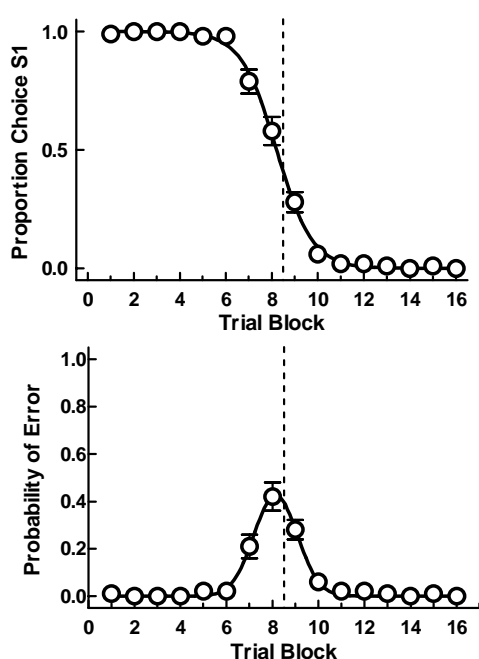


Figure 2. Top. Mean \pm SEM proportion of responses to S1 in choice trials (where both S1 and S2 are presented) as a function of trial block, averaged over the last 10 sessions of training. The smooth curve is a sigmoidal fit to
$$= \frac{1}{1 + \exp(-s(x-b))}$$
, where x is the block number, b is the fitted indifference point, and s is the fitted parameter that determines the slope ($s = -1.44$, $SEM = .13$; $b = 8.19$, $SEM = 0.14$). **Bottom.** Number of errors (responses to the currently unrewarded stimulus) made as a function of trial block, pooled across the last 10 sessions for all pigeons. The solid line represents a best fitting Gaussian function $(Y = A * \exp(-0.5 \left(\frac{(X-Mean)}{SD}\right)^2))$, where A is the maximal number of errors at the mean of the function (Mean), ($A = 4.23$; Mean = 8.14, SD = 0.94). The dotted line indicates the point of the reversal.

In order to test the validity of the SCM predictions one needs to compare the data in Figure 1 (latencies) with those in Figure 2 (proportion of choices). As Figure 1 shows, latency to respond in choice trials was short

and constant for both stimuli throughout the session, hovering just above 1s, similarly to the latencies to respond to the presently rewarded stimulus in sequential trials. In contrast with sequential trials, in choice trials there was no correspondence between latency to respond and time in the session or difficulty to choose: latency to produce either anticipatory or perseverative ‘errors’ were on the same level as latencies to respond correctly, even close to the midsession, where probability of choosing correctly was close to 50%. To quantify potential trends, the slopes of the latencies to respond to S1 and S2 in choice trials were fitted using linear mixed effects models ($s = -.01$, $SEM = 0.01$ and $s = 0.01$, $SEM = 0.02$ respectively); neither slope was significantly different from zero, ($p = .703$ and $p = .091$ respectively), although allowing for S1 and S2 slopes to differ increased by 8.63 times the likelihood of the model, according to the corrected Akaike Information Criterion; (Burnham, Anderson, & Huyvaert, 2011).

Comparison of Sequential and Simultaneous trials

To determine if latencies on choice trials differed from those in single stimulus trials, a linear mixed effects model using the fixed nominal variables of Trial Type (sequential vs. choice) and Stimulus (S1 vs. S2) was used for S1 and S2 latencies. To assure appropriate sampling sizes (see table S1), although Figure 1 shows all available data, statistical analysis of S1 latencies only included trial blocks 1-10 (namely up to two 5-trial blocks after the reversal), while analysis of S2 latencies only included trial blocks 7-16 (namely from two 5-trial blocks prior to the midsession).

Although Figure 1 shows that average latencies in choice trials were shorter than those in sequential trials, the statistical analyses, which used the reduced data sets, yielded no significant effects of trial type, block, or interactions (all $ps \geq .116$), confirming the lack of systematic differences between sequential and simultaneous trial types that is apparent from visual inspection of the data in the rewarded halves of the session. In the session regions when a stimulus is not rewarded (S2 for the first half and S1 for the second half), the number of choices of the non-reinforced stimulus decreases, but latencies to that stimulus remain below the corresponding single-stimulus latency (grey symbols in Figure 1). The scarcity of such choices prevents the

statistical confirmation of this shortening (the number of subjects contributing to the choice of S1 and S2 in each block of trials is presented in Supplementary Table 1). Correlation r^2 and linear regression slope and intercept estimates for SCM predictions against observed data for each subject appear in Supplementary Table 2.

Discussion

Testing different models

Construction of preference. The prevalent idea that choice results from a comparative cognitive evaluation of options' differences, and that preference is not a meaningful concept outside specific choice contexts are not supported by our data. Latency to respond in single-stimulus trials (namely in the absence of opportunity to choose) yields a rather direct measure of the subject's valuation of each stimulus, and follows an orderly and predictable pattern: long latencies when the stimulus is unlikely to yield reward and short latencies when the opposite is true. These latencies are sufficient to predict preference in choice trials, indicating that whatever it is that drives choices, that entity is also present outside the context of choice, rather than being constructed when a choice is elicited.

Rational choice and 'tug of war' models. A common property of the *tug of war* family of models is that choices depend on an evaluation of differences between the members of the choice set. As a consequence, according to these models, time to respond to any target should be longer when a competing alternative is added, and latency to act should increase when the evaluation of relative attractiveness is more difficult, as the options are closer to each other. None of these expectations was fulfilled.

Visual inspection of latencies in no-choice trials against latencies in choice trials in Figure 1 shows that, if anything, latencies in no-choice trials were longer than those in the latter, the opposite of the ToW models' expectation but in agreement to the predictions of the SCM. The statistical reliability of the trend towards shorter latencies in choice than in no-choice trials, however, is hard to quantify, because the sample size for choice trial latencies is controlled by the subjects' preferences, resulting in small sample sizes for the less-

preferred stimulus in choice trials. The statistical analysis using the blocks where all subjects had contributed data failed to detect significant differences. The data thus provide no reason to postulate the existence of an evaluation time cost, which is the main testable prediction of the ToW family of models. The SCM, instead, adds the connection between latencies in no-choice trials and preference in choice trials, as we discuss next.

Sequential choice model. According to the SCM (see assumption ii above), latencies in no-choice trials should predict preference in choice trials. To test this prediction, in each 5-trial block we averaged the latencies to respond in the two single-stimulus trials with S1 and in the two single-stimulus trials with S2; if the averaged S1 latency yielded a shorter value than that of S2, then S1 was selected as the predicted preferred alternative, and vice versa for S2. The predictions for each subject across blocks were then averaged over the last 10 sessions and compared with the observed choice in that block. The assignment of a 5-s latency to trials where the response did not occur within that time (no-go trials) underestimates latency, and hence overestimates predicted probability of a stimulus being chosen in blocks where the proportion of no-go trials is substantial, resulting in less extreme predicted preferences than the logic of the model implies.

The aggregate of the predictions for all subjects over the last 10 training sessions is presented in Figure 3, superimposed with the observed choices. The observed and predicted functions are similar in shape and in value, with the predicted one, less steep than that observed. To analyze the data at an individual level, we correlated observed versus predicted choices, and ran a linear regression for each subject, as shown in Table S2 and Figure S2. The correlation between SCM predictions and observed choices was significant for all 10 subjects, and explained between 82% and 98% of the within-subject variance.

If our implementation of the SCM were to perfectly predict the observed choices, one would expect linear regression intercepts of 0 and slopes of 1. In fact, all 10 regressions had positive intercepts and slopes lower than 1 for all subjects, implying that the model's predicted preferences were significantly less extreme than the data for the whole population ($p < 0.01$, sign test). Several comments and caveats are appropriate to judge the meaning of these correlations.

First, as noticed before, the way we computed SCM predictions generates less extreme preferences as an artifact of the 5 s censorship in latency duration imposed by the protocol, so that the underlying

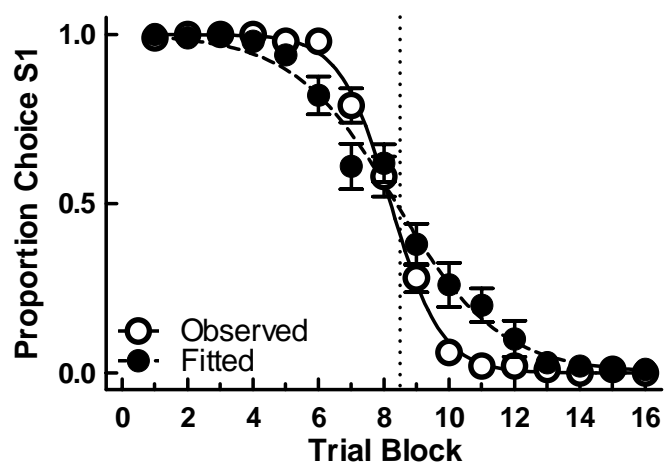


Figure 3. Mean (\pm SEM) proportion choice of S1 for both the observed data and that predicted by the sequential choice model averaged over the last 10 sessions as a function of trial block. The curves are fits with the sigmoidal function used in Figure 1. *Note: the dotted line indicates the point of the reversal.*

correspondence is weakened by the fact that the pigeons were not allowed to express and record latencies longer than 5s. Allowing longer latencies would have led to predicting sharper choice functions, and thus would make the model's predictions closer to the data.

Second, the fit of the model is maximal at the extreme of the sessions, where latencies do not overlap and, then, as predicted by the SCM, choices are absolute. In the intermediate part of the sessions, the distributions of latencies overlap and preferences are partial, again as predicted, but the fit is less tight, and other factors may explain some of the residual variance in those regions. A related but separate issue is that SCM targets the relation between value as expressed by latencies and choice, and does not by itself explain valuation. To see this point, consider a hypothetical individual that shows shorter latencies for the 'wrong' stimulus, and also prefers the wrong stimulus: this supports the SCM as a model of decision processes.

Finally, as explained above, we implemented the SCM by predicting choices using only the sign of the difference between the mean latencies for the two S1 trials and the two S2 trials, ignoring the magnitude of this difference. This is appropriate because for each block we make binary predictions for choosing either stimulus, but it misses information about the strength of the predicted preference for the aggregate results. It is worth noticing that SCM is the only model that makes predictions for choice based on no-choice trials, hence the results cannot be used to compare the goodness of fit between the two categories of model.

A point of interest is that the change in latencies in single-stimulus trials with the currently unrewarded stimulus (S2 before midsession and S1 after midsession) described by Figure 1 are mainly responsible for the change in preference through the session, while latencies in the currently rewarded stimuli (S1 before midsession and S2 after midsession) remain always stable and short, and thus cannot explain changes in preference. This is consistent with the report by McMillan et al. (2015) using only single-stimulus trials that session effects seemed to depend on variation in responding to the currently unrewarded stimulus. These authors explain their results in terms of the failure to inhibit 'error' responses.

We surmise that a reasonable alternative interpretation is that latency in responding to each stimulus is driven by the estimated probability that a response will be rewarded. Such an estimate is constantly high after a rewarded response (our assumption iii) but, for an animal experienced with the protocol, it is modified as a function of available information, which in the midsession protocol is fundamentally time since that option yielded a reward (our assumption iv). In summary: estimated probability of reward accounts for systematic variations in latency to respond, independently of the presence or absence of an alternative option, and these latencies explain choices when the two stimuli are simultaneously present. Crucially, there is no signal in the data suggesting a deliberative comparison between options at the time of choice.

General conclusion

The absence of any evidence for a temporal cost of choice, together with the evidence that on single-stimulus trials, subjective value of each alternative is expressed through latency to respond, is sufficient to integrate single-stimulus and choice results. This finding allows for a safe rejection of the idea that it is not possible to identify preferences unless a choice is elicited. Similarly, the fact that adding an alternative does not add any delay in responding argues strongly against models that propose comparative evaluation of alternatives at the time of choice.

An important caveat, however, is that our study used only one dimension (probability of reward). It remains possible that when stimuli signal outcomes that differ in more than one dimension, such as magnitude, probability, or delay of outcomes, choice might not be as strongly predictable from behavior towards stimuli encountered alone. This is a testable proposition, but results such as those reported by Shapiro et al. (2008) that tested combinations of two dimensions (amount and delay to reward) still favor the parallel account.

A further arena to compare the different models in the future may be brain imaging. In humans, imaging studies suggest that different areas of the brain are recruited for comparative and sequential choices (Kolling, Behrens, Mars, Rushworth, 2012). The SCM may apply to limited aspects of human decision making, underlying what Kahneman (2011) calls fast decisions governed by System 1. These are decisions based on

431 unconscious predispositions, heuristics and intuitions that allow us to respond quickly, without taking the time
432 to make the detailed evaluations of the options that are left to decisions made by System 2. From this
433 perspective, most laboratory procedures study choice in non-humans addressing the equivalent of the animals'
434 System 1.
435

References

- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113(2), 409–432. doi: 10.1037/0033-295X.113.2.409
- Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65(1), 23-35.
- Cook, R. G., & Rosen, H. A. (2010). Temporal control of internal states in pigeons. *Psychonomic Bulletin & Review*, 17, 915-922.
- Gigerenzer, T., Hertwig, R., & Pachur, T. (Eds.), (2011). *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.
- Gigerenzer, T. Todd, P. M., & The ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gravelle, H., & Rees, R. (2004). *Microeconomics* (3rd edition) Prentice Hall, Harlow, UK
- Hick W. E., (1952). On the Rate of Gain of Information. *Quarterly Journal of Experimental Psychology*, 4, 11–26.
- Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, 45, 188–196.
- Kacelnik, A., Vasconcelos, M., Monteiro, T. & Aw, J. (2011). Darwin’s “tug-of-war” vs. starlings’ “horse-racing”: how adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology*. 65, 547–558. DOI 10.1007/s00265-010-1101-2
- Kahneman, D. (2011) *Thinking, fast and slow*. Farrar, Straus & Giroux.
- Kolling, N., Behrens, T. E., Mars, R. B., Rushworth, M.F. (2012). Neural Mechanisms of Foraging. *Science*, 336(6077):95-8. doi: 10.1126/science.1216930.
- Lamport, L. (2012). Buridan’s Principle. *Foundations of Physics*, 42, 1056-1066.
<https://doi.org/10.1007/s10701-012-9647-7>

- 460 McMillan, N., & Roberts, W. A. (2012). Pigeons make errors as a result of interval timing in a visual, but not a
 461 visual-spatial, midsession reversal task. *Journal of Experimental Psychology: Animal Behavior Processes*,
 462 38, 440–445. Doi:10.1037/a0030192
- 463 McMillan, N., Sturdy, C. B. & Spetch, M. L. (2015). When is a choice not a choice? Pigeons fail to inhibit incorrect
 464 responses on a go/no-go midsession reversal task. *Journal of Experimental Psychology: Animal Learning*
 465 *and Cognition*, 41, 255-265.
- 466 Pelé, M., & Sueur, C. (2013). Decision-making theories: linking the disparate research areas of individual and
 467 collective cognition. *Animal Cognition*, 16, 543–556.
- 468 Pinheiro, J., Bates, D., DebRoy, S., Team R C (2016). nlme: Linear and nonlinear mixed effects models R.
 469 Retrieved from <http://cran.r-project.org/web/packages/nlme>
- 470 Rayburn-Reeves, R. M., & Cook, R. G. (2016). The organization of behavior over time: Insights from mid-session
 471 reversal. *Comparative Cognition & Behavior Reviews*, 11, 103–125. Doi:10.3819/ccbr.2016.110006
- 472 Rayburn-Reeves, R. M., Molet, M., & Zentall, T. R. (2011). Simultaneous discrimination reversal learning in
 473 pigeons and humans: Anticipatory and perseverative errors. *Learning & Behavior*, 39, 125-137.
- 474 Schwartz, B., & Cheek, N. (2017). Choice, freedom, and well-being: Considerations for public policy. *Behavioural*
 475 *Public Policy*, 1(1), 106-121. Doi:10.1017/bpp.2016.4
- 476 Shafir, E., Simonson, I., & Tversky, A. 1993. Reason-based choice. *Cognition*, 49, 2, 11-36.
- 477 Shapiro, M. S., Siller, S, & Kacelnik, A. (2008). Simultaneous and sequential choice as a function of reward delay
 478 and magnitude: normative, descriptive and process-based models tested in European starling (*Sturnus*
 479 *vulgaris*). *Journal of Experimental Psychology: Animal Behavioral Process*, 34. 75–93.
- 480 Simon, H. (1991). Bounded rationality and organizational learning. *Organizational Science*, 2, 125-134.
- 481 Slovic, P. The construction of preference. (1995). *American Psychologist*, 50 (5) 364-371.
- 482 Smith, A. P., Beckmann, J. S., & Zentall, T. R. (2017). Mechanisms of midsession reversal accuracy: memory for

- 483 preceding events and timing. *Journal of Experimental Psychology: Animal Learning and Cognition*, 43,
 484 62-71.
- 485 Smith, V. L. (2002). *Constructivist and Ecological Rationality in Economics Nobel Prize Lecture*. Les Prix
 486 Nobel. The Nobel Prizes, Editor Tore Frängsmyr, [Nobel Foundation], Stockholm.
- 487
- 488 Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton University Press, Princeton.
- 489 Tatham, T.A. & Zurn, K.R. (1989). The Med-PC experimental apparatus programming system. *Behavioral*
 490 *Research Methods, Instruments, and Computers*, 21, 294 – 302.
- 491 Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin &*
 492 *Review*, 11(1), 192-196.
- 493 Young, M. E., Clark, M., Goffus, A., & Hoane, M. R. (2009). Mixed effects modeling of Morris water maze data:
 494 Advantages and cautionary notes. *Learning and Motivation*, 40(2), 160-177.

495

Animal Care Note

496

The pigeons were cared for in accordance with University of Kentucky animal care guidelines. PSA

497

Assurance #D16-00217 (A3336-01).

498

SUPPLEMENTARY MATERIALS

Supplementary Table 1.

Number of subjects contributing choices to each of the stimuli in each block across the last 10 sessions.

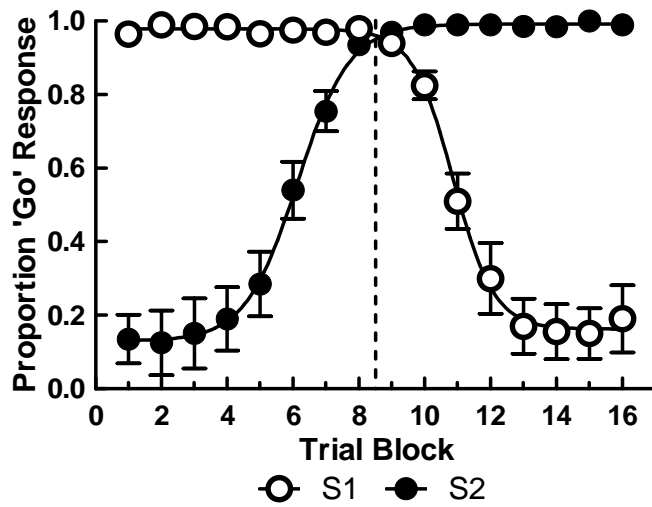
Block	S1	S2
1	10	0
2	10	0
3	10	0
4	10	0
5	10	2
6	10	2
7	10	8
8	10	10
9	10	10
10	5	10
11	2	10
12	2	10
13	1	10
14	0	10
15	1	10
16	0	10

Supplementary Table 2.

Correlation r^2 and linear regression slope and intercept estimates for SCM predictions against observed data for each subject. Significant r^2 values indicate that SCM predictions significantly correlated with the observed data, while significant slopes and intercepts indicate significant deviation from a perfect prediction with a slope of 1 and intercept of 0. All subjects show slope < 1 , and intercept > 0 , implying that predicted preferences are significantly less steep than those observed (see text).

Subject	r^2	slope	intercept
1	0.98*	$1.00 \pm (0.04)$	$0.01 \pm (0.03)$
25	0.82*	$0.76^* \pm (.10)$	$0.14^* \pm (0.06)$
723	0.92*	$0.93 \pm (0.07)$	$0.00 \pm (0.05)$
725	0.84*	$0.75 \pm (0.09)^*$	$0.22 \pm (0.06)^*$
728	0.98*	$0.95 \pm (0.03)$	$0.02 \pm (0.02)$
1869	0.82*	$0.86 \pm (0.11)$	$0.15 \pm (0.07)$
19338	0.96*	$0.92 \pm (0.05)$	$0.02 \pm (0.03)$
19824	0.83*	$0.68^* \pm (0.08)$	$0.19^* \pm (0.05)$
19836	0.91*	$0.90 \pm (0.08)$	$0.02 \pm (0.05)$
19849	0.90*	$0.89 \pm (0.08)$	$0.04 \pm (0.05)$

* indicates $p < .05$

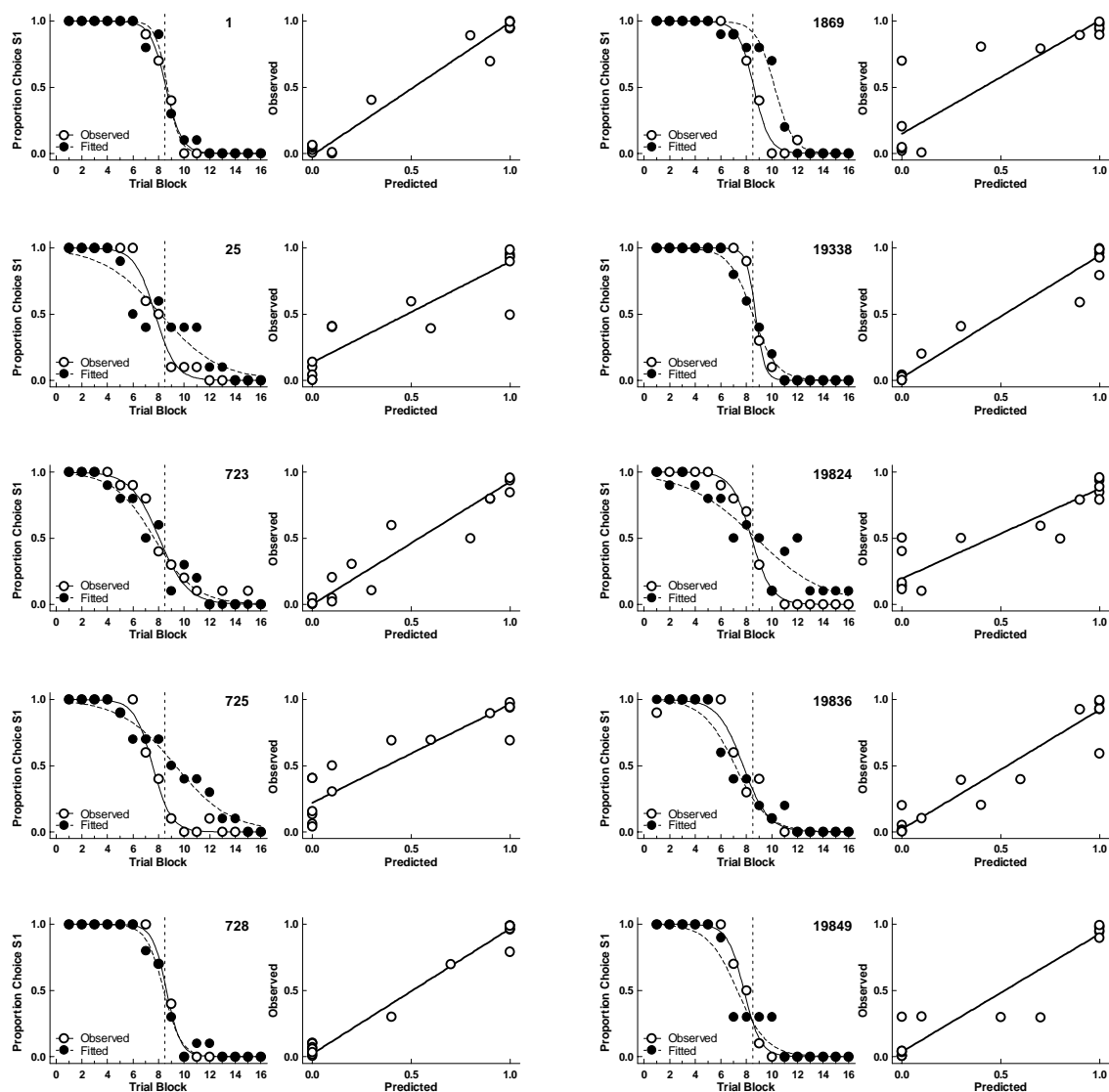


Supplementary Figure 1. Mean (\pm SEM) proportion 'Go' response to each stimulus on single-stimulus trials,

averaged over the last 10 sessions as a function of trial block. The fitted line follows equation $= \frac{1}{1 + \exp(-s(x-b))}$.

Slope parameters: $s=1.61$, SEM 0.31 for S1 and $s=1.26$, SEM=0.23 for S2.

531



532

533 **Supplementary Figure 2.** Individual data. Columns one and three show the mean (\pm SEM) choice of S1 for both
 534 actual (observed) and predicted data, both fitted with the Equation described in Figure S1. Columns 2 and 4
 535 show observed Vs. predicted regression of choice proportions of choices for S1 with each point corresponding
 536 to one block, averaged over the last 10 sessions. *Note: data points are jittered along the y-axis to aid*
 537 *visualization, while the line of best fit is for non-jittered data.*