

# Accepted Manuscript

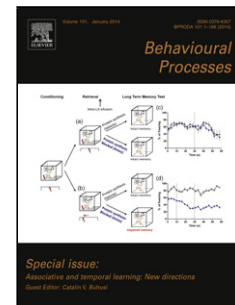
Title: Ultimate explanations and suboptimal choice

Authors: Marco Vasconcelos, Armando Machado, Josefa N.S. Pandeirada

PII: S0376-6357(17)30606-X  
DOI: <https://doi.org/10.1016/j.beproc.2018.03.023>  
Reference: BEPROC 3642

To appear in: *Behavioural Processes*

Received date: 11-12-2017  
Revised date: 21-3-2018  
Accepted date: 28-3-2018



Please cite this article as: Vasconcelos M, Machado A, Pandeirada JNS, Ultimate explanations and suboptimal choice, *Behavioural Processes* (2018), <https://doi.org/10.1016/j.beproc.2018.03.023>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Ultimate explanations and suboptimal choice

Marco Vasconcelos<sup>1,2</sup>, Armando Machado<sup>3</sup>, and Josefa N. S. Pandeirada<sup>1,4</sup>

<sup>1</sup>Department of Education and Psychology, University of Aveiro, Portugal, <sup>2</sup>CESAM, University of Aveiro, Portugal, <sup>3</sup>School of Psychology, University of Minho, Portugal, <sup>4</sup>CINTESIS.UA, University of Aveiro

## Author Note

Josefa N. S. Pandeirada was supported by the Portuguese Foundation for Science and Technology (grant numbers FCOMP-01-0124-FEDER-029610-PTDC/MHC-PCN/5274/2012 and IF/00058/2012/CP0172/CT0002).

Correspondence concerning this article should be addressed to Marco Vasconcelos, Department of Education and Psychology, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal. Email: mvasconcelos@ua.pt

## Highlights

- The optimizing action of natural selection does not imply optimal behavior
- Breaches of rationality are expected due to a number of constraints
- A model incorporating such constraints can account for sub-optimal choice

- Deviations from rationality are tools to understand the adaptive significance of behavioral mechanisms.
- Existing mechanisms bear the footprint of the selective pressures of the animals' characteristic ecology.

## **Abstract**

Researchers have unraveled multiple cases in which behavior deviates from rationality principles. We propose that such deviations are valuable tools to understand the adaptive significance of the underpinning mechanisms. To illustrate, we discuss in detail an experimental protocol in which animals systematically incur substantial foraging losses by preferring a lean but informative option over a rich but non-informative one. To understand how adaptive mechanisms may fail to maximize food intake, we review a model inspired by optimal foraging principles that reconciles sub-optimal choice with the view that current behavioral mechanisms were pruned by the optimizing action of natural selection. To move beyond retrospective speculation, we then review critical tests of the model, regarding both its assumptions and its (sometimes counterintuitive) predictions, all of which have been upheld. The overall contention is that (a) known mechanisms can be used to develop better ultimate accounts and that (b) to understand why mechanisms that generate suboptimal behavior evolved, we need to consider their adaptive value in the animal's characteristic ecology.

*Keywords:* long-term rate; mechanism; natural selection; rationality; sub-optimal choice; ultimate explanation.

## Ultimate explanations and suboptimal choice

### 1. Introduction

Ideally, to understand behavior, we must explain how it is generated and why it occurs. The former leads to proximate explanations that deal with the social, psychological, and biological processes underpinning behavior; the latter leads to ultimate explanations that deal with the fitness consequences of behavior (Mayer 1963; Tinbergen 1963). Put differently, to fully understand behavior, we need to unravel both its mechanisms (the “how” question) and its evolutionary functions (the “why” question). Despite common misunderstandings, these accounts are complementary, each contributing key ingredients for a complete explanation of behavior (for contemporary treatments, see Bateson and Laland 2013; Dawkins 2014; Scott-Phillips, Dickins, and West 2011). Yet, they are not independent because answers to one type of question (proximate/ultimate) should restrict the range of viable answers to the other type of question (ultimate/proximate).

The vast majority of research in the psychological sciences aims at identifying the mechanism(s) underlying some cognitive or behavioral feat and thus proximate questions are almost second nature. Ultimate questions are somewhat more unusual. What is meant by ‘fitness consequences’ or ‘evolutionary function’? Since Darwin’s (1859) key observation and the subsequent merger with Mendelian genetics (Fisher 1930; Haldane 1932; Wright 1931), it has become apparent that natural selection will produce organisms seemingly designed to maximize their inclusive fitness (Hamilton 1964, 1970; see also Gardner 2009; Grafen 2007). In other words, inclusive fitness is the currency that natural selection maximizes—its *maximand* (Grafen 2006). Hence, any ultimate explanation must explain how is it that a particular behavior contributes to inclusive fitness more than other possible variants.

A practical difficulty lies in measuring fitness. It is usually not possible to measure directly the fitness consequences of a particular behavior. The alternative is the use of proxies. For instance, optimal foraging theory seeks to explain the foraging strategies used by animals in terms of the maximization of the rate of energy gain (Stephens and Krebs 1986; see also Stephens, Brown, and Ydenberg 2007). In Charnov's (1976b, 1976a) seminal models, the assumption is that foraging animals make decisions as if maximizing their long-term rate of energy gain computed as the ratio of expected energy obtained from food to the expected time spent foraging. Many other proxies to measure the fitness consequences of foraging are possible, including the net rate (the net energy gain per unit of time) or efficiency (energy gain per unit of energy spent). Notwithstanding, the assumption is always the same: proxy maximization correlates positively with fitness maximization.

If behavior and the decision-making processes involved in behavioral allocation were sculpted by natural selection throughout evolutionary history, then one might (naively) expect most, if not all, behavior to be well adjusted to whatever demands are imposed on the organism. Yet, mounting evidence suggests that irrational behavior, defined as failures to maximize a particular benefit or as showing inconsistent preferences (e.g., violations of transitivity), are much more common than previously suspected. Striking deviations from rationality principles have been found in a variety of taxa, including humans (e.g., Baron 2008; Doyle et al. 1999; Gigerenzer, Hertwig, and Pachur 2011; Gilovich, Griffin, and Kahneman 2002; Goldstein and Gigerenzer 2002; Kalenscher et al. 2010; Tversky 1969; Tversky and Kahneman 1974, 1986) and nonhumans (e.g., Bateson 2002; Bateson, Healy, and Hurly 2002; Hurly and Oseen 1999; Latty and Beekman 2011; Shafir 1994; Shafir, Waite, and Smith 2002; Waite 2001). Although these observations are sometimes mitigated by reports of strict rationality (e.g., Monteiro, Vasconcelos, and Kacelnik

2013), they are at first glance difficult to reconcile with the view that cognitive and behavioral systems were pruned by the optimizing action of natural selection.

To avoid or minimize the paradox, one could assume that the costs of such deviations are negligible, but, in most instances, this is not the case. We propose instead that we consider two constraints. First, existing mechanisms may preclude natural selection from reaching (near) optimal solutions. Because naturally selected systems are not built from scratch, the current “raw material” necessarily constrains all evolutionary possibilities. Therefore, ultimate and proximate explanations although separate are not independent: functional issues constrain which mechanisms evolve, and existing mechanisms constrain the range of feasible solutions. Known mechanisms should inform functional explanations (e.g., McNamara and Houston 2009; Kacelnik 2012; Vasconcelos, Fortes, and Kacelnik 2017).

Second, ‘irrational’ behavior can be caused by *adaptive* decision processes when the domain of selection and the domain of testing mismatch in key ingredients (Stevens and Stephens 2010; Houston, McNamara, and Steer 2007). To elaborate, decision processes are expected to perform well in the environment an animal is likely to encounter (i.e., the environment in which it evolved; cf. Seth 2007), but may be suboptimal when the animal is placed in an environment that does not mirror the main features of the environments to which it is adapted (Haselton et al. 2009). Recognizing this gap is an essential step towards understanding the adaptive significance of behavioral mechanisms.

To illustrate how these two constraints can help us understand breaches of rationality, we discuss in detail an experimental protocol in which animals systematically incur substantial foraging losses by preferring a lean but informative option over a richer but non-informative one.

## 2. Suboptimal choice

The outcome of most foraging cycles is uncertain. Yet, in some instances, reliable cues indicate whether success or failure is forthcoming. The distance between predator and prey may reach a threshold beyond which the prey will certainly escape or it may reach a different threshold below which the predator will certainly succeed. Consider then the following scenario: a forager can choose between two options that differ in the probability of reinforcement (the probability of a successful cycle) and in the duration of uncertainty about the outcome of the cycle. The *Informative* option ends in success rarely (with probability  $p_{info}$ ) but it informs the forager about the forthcoming outcome immediately after choice by displaying either a positive discriminative stimulus (say a Red key light) or a negative discriminative stimulus (say a Green key light) for  $t$  s, with probabilities  $p_{info}$  and  $1 - p_{info}$ , respectively. We represent these contingencies by  $S_{R, 1.0}$  (i.e., Red signals impending food) and  $S_{G, 0.0}$  (i.e., Green signals the absence of food). The *Non-informative* option, on the other hand, ends in reinforcement more often (with probability  $p_{non-info}$ , where  $p_{non-info} > p_{info}$ ) but once chosen it displays for  $t$  s one of two non-discriminative stimuli, say a Blue or a Yellow key light, each ending in reinforcement half of the times, such that the outcome remains uncertain until it happens. We represent these contingencies by  $S_{B, 0.5}$  and  $S_{Y, 0.5}$ .

Figure 1 depicts the procedure originally developed by Zentall and collaborators (e.g., Stagner and Zentall 2010) with the typical parameters ( $p_{info} = .2$ ,  $p_{non-info} = .5$ , and  $t = 10$  s). With these parameters, the probability of reinforcement in the Non-informative option is 2.5 times higher than in the Informative option, but both pigeons (*Columba livia*) and starlings (*Sturnus vulgaris*) show extreme preferences (usually  $> 90\%$ ) for the latter (e.g., Fortes, Vasconcelos, and Machado 2016; Fortes, Machado, and Vasconcelos 2017; Laude, Stagner, and Zentall 2014; Stagner and Zentall 2010; Stagner, Laude, and Zentall 2011; Stagner, Laude, and Zentall 2012;

Vasconcelos, Monteiro, and Kacelnik 2015; Zentall and Stagner 2011a, 2011b; Zentall et al. 2015; for reviews see McDevitt et al. 2016; Zentall 2014, 2016). To put this figure into perspective, an animal choosing the Informative but suboptimal option 90% of the time would be foregoing 54% of the maximum attainable benefit and would even be losing 12% relative to random choice. Figure 2 shows typical findings with pigeons and starlings.

The failure to maximize *food intake* is unquestionable. The meaning of such failure, however, is not straightforward. Some researchers frame it as a violation of the foundational principles of optimal foraging theory (e.g., McDevitt et al. 2016; Zentall 2016), perhaps because they assume that according to optimality theory, animals should choose optimally or rationally under all circumstances. This assumption misinterprets the entire optimality framework and it may stem from the erroneous premise that animals must compute optima in order to act (Vasconcelos et al. 2017). Were this the case, a brief overview of the animal learning and behavior literature would suffice to debunk the entire optimality enterprise, for animals frequently fail to maximize their rewards: The so-called self-control paradigm has shown that animals often prefer a shorter but sooner reward over a larger but later reward even when the latter yields a higher rate of intake (e.g., Green et al. 1981; Mazur 1987; Mazur 2000); the pervasive “matching law” (Herrnstein 1961) deviates from reinforcement maximization in a variety of experimental preparations, though not all (Herrnstein and Heyman 1979; Herrnstein 1990).

How then should researchers handle instances of sub-optimal preferences? In our view, they should be seen as tools to understand the adaptive significance of the mechanisms generating such preferences (McNamara and Houston 2009). To answer the ‘why question’ we assume that the optimization agent is natural selection (not the organism itself) and therefore that the current behavioral and cognitive mechanisms are those that outperformed other alternatives in past



generations. Consequently, on the average, these mechanisms ought to be adaptive. How do we reconcile this prediction with the ostensibly ‘irrational’ preferences shown in Figure 2? We must consider both the statistical features of the environment to which the animal is adapted and the well-established mechanisms we know animals possess.

### 3. A functional model

According to classical optimal foraging theory (Stephens and Krebs 1986), evolved behavioral mechanisms ought to maximize the ratio of expected gains to expected time (used here as a proxy for fitness) because such mechanisms outperformed their contemporaneous variants over the generations. This is equivalent to saying that these mechanisms *minimize lost opportunity* because time spent pursuing or handling one alternative cannot be used to search for alternatives. Consider, then, the following scenario: an animal searches for food and, after  $s$  seconds, finds item  $i$ ; the animal stops searching and pursues item  $i$  for  $t$  seconds, after which it either catches the item with probability  $p$  and consumes it (with a handling time,  $h$ ) or the item escapes with probability  $1 - p$  and the cycle restarts. If the animal chooses exclusively item  $i$  (and assuming energy content equal to unity), the long-term rate of return is given by

$$R_i = \frac{p_i}{s + p_i \times (t + h) + (1 - p_i) \times t} \quad (1)$$

Applying the same rationale to the experimental situation depicted in Figure 1,  $p_i$  is the overall probability of reinforcement in a given option,  $s$  corresponds to the inter-trial interval (ITI),  $t$  is the duration of the terminal links (which, in this case, is always the same), and  $h$  is the time to consume the reward (also the same for both options). Because (a) the rate of return increases monotonically with  $p_i$  and (b)  $p_i$  is higher for the Non-informative than for Informative option (0.5

vs. 0.2, respectively), Equation 1 predicts a preference for the former whereas animals prefer the latter.

Yet, despite appearances, an animal behaving as if maximizing the currency shown in Equation 1 would not be maximizing its long-term rate of return or, equivalently, minimizing lost opportunity in this task. Under natural circumstances, an animal would not pay the cost of waiting for certain no reward but would use that time to search for alternatives. In other words, the experimental task imposes a waiting period for no reward in the presence of  $S_{G, 0.0}$ , but animals were sculpted by natural selection to use the information about no impending reward to redirect their foraging efforts. By continuing to engage with a stimulus signaling the absence of reward, they would decrease their rate of return because the average time to capture an item would increase. In practice, this means that animals should disengage from  $S_{G, 0.0}$  and ignore the time spent in its presence, but actively wait in the presence of the other terminal stimuli ( $S_{R, 1.0}$ ,  $S_{B, 0.5}$  and  $S_{Y, 0.5}$ ). Equation 1 assumes that animals pay the opportunity cost in the presence of  $S_{G, 0.0}$ ; we propose instead that animals disengage from this stimulus as they would in their natural environment. Theoretically, the learning mechanism could take into account the forced delays in the computation of rate of food intake – in which case Equation 1 would apply as it is. But we argue that deprived of the corresponding selection pressures, the learning mechanisms did not evolve that flexibility.

Another feature of the experimental task that deviates from natural foraging scenarios is the ITI ( $s$  in Equation 1). In the natural environment, animals start searching and only then experience the search time (which should thus be included in rate computations for any item; e.g., Kacelnik and Todd 1992; Kacelnik and Brunner 2002; Todd and Kacelnik 1993); conversely, in experimental preparations, the animal experiences the ITI before the trial starts and after the outcome. Extensive research has shown that in the latter case the ITI is not credited to a specific

option (e.g., Logue, Smith, and Rachlin 1985; Mazur 1989; Spetch et al. 1994; Stephens and Anderson 2001). In fact, animals learn the consequences of their actions by their outcomes and therefore they should be sensitive to delays between actions and outcomes but relatively insensitive to other intervals (Bateson and Kacelnik 1996). In the wild, animals do not experience delays between food and the next foraging opportunity; they are free to start searching immediately.

Even though the first optimality models ignored known mechanisms, we now know that such mechanisms should inform functional accounts (e.g., McNamara and Houston 2009; Kacelnik 2012; Vasconcelos et al. 2017). Therefore, considering the constraints related to the ITI and the time spent in the presence of  $S_{G, 0.0}$ , and assuming for simplicity  $h = 1$ , the rate of return of the Informative option becomes

$$R_{info} = \frac{p_{info}}{s + p_{info} \times (t+1) + (1-p_{info}) \times t} = \frac{p_{info}}{p_{info} \times (t+1)} = \frac{1}{t+1} \quad (2)$$

where  $p_{info}$  is the probability of reward in the Informative option (and equals the probability of  $S_{R, 1.0}$  being presented). Neither the ITI ( $s$ ) nor the time spent in the presence of  $S_{G, 0.0}$  feature in the equation. In practice, the animal perceives the Informative option as always leading to  $S_{R, 1.0}$  and therefore food. For the Non-informative option, the ITI is also excluded but, because the delivery of food is not signaled reliably, both terminal link durations are considered. According to Equation 1, and considering that  $P(S_{B, 0.5}) + P(S_{Y, 0.5}) = 1$ , the rate of return in the Non-informative option becomes

$$R_{Non-info} = \frac{p_{non-info}}{s + p_{non-info} \times (t+1) + (1-p_{non-info}) \times t} = \frac{1}{\frac{t}{p_{non-info}} + 1} \quad (3)$$

where  $p_{non-info}$  is the probability of reward in the non-informative option. Assuming, as in the typical task, that  $p_{info} = .2$ ,  $p_{non-info} = .5$  and  $t = 10$  s, the rate for the Informative option is now

higher than the rate for the Non-informative option (approximately 0.09 vs. 0.05 reinforcers/s, respectively) as required.

None of the assumptions and simplifications we have made to arrive at this result is trivial. We have resorted to a sort of reverse engineering by first identifying a puzzling result; next we considered the matches and mismatches between the experimental task and the organisms' natural environment in order to identify the key components of the task; and finally we proposed how the puzzling result could emerge from real-world mechanisms operating in the artificial laboratory settings. Nonetheless, one adaptive explanation does not rule out the existence of many others. To dispel the ghost of mere speculation (the proverbial *story-telling*), we must move beyond the explanation of known results and, if possible, (a) test the assumptions of the model and (b) generate and test novel predictions.

Some of the assumptions of the model are amenable to testing. Currently, two related, yet distinct, assumptions have been tested:

- A1. The model assumes that the subjects learn the signaling properties of each terminal stimulus. This assumption is paramount. If animals do not learn that some signals ( $S_{R, 1.0}$  and  $S_{G, 0.0}$ ) predict the outcome whereas others ( $S_{B, 0.5}$ ,  $S_{Y, 0.5}$ ) do not, we would have no basis to presuppose that animals disengage from  $S_{G, 0.0}$ .
- A2. The core assumption of the model is that in the wild animals use the information conveyed by stimuli signaling the absence of food to move away and search for a patch or prey elsewhere thus increasing their probability of success. Similarly, in the laboratory, they disengage from the  $S_{G, 0.0}$  but have nowhere else to search in the usually small operant chamber. In other words, they behave as if that information were usable when in fact it is not. Therefore, even though animals cannot physically

withdraw from the ‘patch’, they should nonetheless show some signs of disengaging from  $S_{G, 0.0}$  but not from the other stimuli.

The model (cf. Equations 2 and 3) also leads to a number of predictions, some of them counterintuitive:

- P1. Preference for the Informative option should not depend on its associated probability of reinforcement ( $p_{info}$  in Equation 2). Because  $S_{G, 0.0}$  is not taken into account, the Informative option is perceived as always leading to  $S_{R, 1.0}$  independently of the objective probability, unless it is zero. If the ancillary assumption that  $p_{non-info} < 1$  is met, then sub-optimal choice should always surface.
- P2. Preference for the Informative option should be independent of the duration of  $S_{G, 0.0}$ . If our analysis is correct, animals were shaped by natural selection to behave as if  $S_{G, 0.0}$  did not affect the value of the Informative option (cf. Equation 2), which means that not only the probability (see previous prediction) but also the duration of  $S_{G, 0.0}$  is irrelevant *vis-à-vis* preference.
- P3. Predictability and the duration of uncertainty are intertwined in the protocol depicted in Figure 1: The Informative option informs about the outcome and eliminates uncertainty immediately whereas the Non-informative option does not disambiguate the outcome and maintains uncertainty until the outcome occurs. Thus far, we have mainly focused on predictability but, according to the model, the duration of uncertainty is as important a driving wheel of suboptimal choice as predictability. If, for instance, the Informative option did not eliminate uncertainty immediately the animal would have no reason to disengage from  $S_{G, 0.0}$  and the time under uncertainty

should be included in the rate computations, reducing the profitability of this option. Thus, if the moment when information is conveyed by the Informative option is delayed by  $t$  s (roughly equalizing uncertainty duration across options but maintaining predictability as in the typical task), preference should follow the objective probabilities of reward.

In what follows, we briefly summarize the main findings of a set of experiments testing the assumptions and predictions of the model. These experiments are part of an ongoing research program aimed at clarifying the proximate and ultimate sources of ‘irrational’ behavior.

### 3. Testing the assumptions and predictions of the model

#### 3.1. *Animals learn the outcomes that follow each terminal stimulus (A1)*

This assumption can be tested in a number of ways. For instance, one can analyze the responses to each terminal stimulus or one can give animals explicit choices between the stimuli. Regarding responses to the terminal stimuli, the model assumes that animals engage with (and presumably respond to) stimuli that are intermittently or always followed by food ( $S_{R, 1.0}$ ,  $S_{B, 0.5}$ ,  $S_{Y, 0.5}$ ) but disengage from (and do not respond to) stimuli never followed by food ( $S_{G, 0.0}$ ). The top panel of Figure 3 shows the cumulative number of responses starlings gave to each terminal link as a function time (Vasconcelos et al. 2015, Experiment 2, control group). As expected, animals showed a relatively constant rate of responding to  $S_{R, 1.0}$ ,  $S_{B, 0.5}$ , and  $S_{Y, 0.5}$ , with the rate for the perfect predictor of food ( $S_{R, 1.0}$ ) about twice that found for the unreliable predictors ( $S_{B, 0.5}$ , and  $S_{Y, 0.5}$ ). Importantly, responding to  $S_{G, 0.0}$  was mostly absent, also as expected (see also Hinnenkamp, Shahan, and Madden 2017). These findings are in line with findings in the

autosshaping literature showing that animals typically engage with stimuli that predict food but not with stimuli that predict its absence (e.g., Brown and Jenkins 1968; Hearts and Jenkins 1974; Wasserman, Franklin and Hearst 1974).

Regarding binary choices between terminal stimuli, the model presupposes that animals learn the outcomes that follow each signal and therefore preferences between terminal stimuli should reflect the impending outcomes. After learning the typical task, Vasconcelos and colleagues (2015; Experiment 1) gave starlings the opportunity to repeatedly choose between cross-option terminal links (namely,  $S_{R, 1.0}$  vs.  $S_{B, 0.5}$ ,  $S_{R, 1.0}$  vs.  $S_{Y, 0.5}$ ,  $S_{G, 0.0}$  vs.  $S_{B, 0.5}$ ,  $S_{G, 0.0}$  vs.  $S_{Y, 0.5}$ ). The bottom panel of Figure 3 shows that the average preferences (filled circles) were in agreement with the probability of reward signaled by each option. In other words, they preferred  $S_{R, 1.0}$  but avoided  $S_{G, 0.0}$ .

In sum, both the rate of responding and the preference measure indicate that animals do learn the signaling properties of the terminal stimuli as presupposed by the model.

### 3.2. *Animals disengage from $S_{G, 0.0}$ but not from the other terminal stimuli (A2)*

The response profiles shown in the top panel of Figure 3 suggest that animals disengage from  $S_{G, 0.0}$  but not from the other terminal stimuli. Yet, this is only indirect evidence because pecking and not pecking may mean only that the response is under stimulus control. An animal not pecking at  $S_{G, 0.0}$  may remain involved in the task.

Previous unsystematic observations suggest that animals not only refrain from pecking but move away from  $S_{G, 0.0}$  (see also Dinsmoor et al. 1982; Rand 1977). Fortes and colleagues (2017; Experiment 2) tested directly the assumption by allowing pigeons to move freely in a large operant chamber (cf. top panel of Figure 4) while recording their location. For simplicity, we assume that

all birds had the Non-informative option on the left (floor position = 1) and the Informative option on the right (floor position = 6). The bottom panel of Figure 4 shows the average position of the birds as a function of time into the terminal link, both when the terminal links were 10-s long and when they were 20 s. Overall, it is clear that pigeons always remained near the chosen option except when  $S_{G,0.0}$  was presented. In such instances, they moved away from the stimulus, usually to the center of the chamber, and waited for the next trial. In sum, when given the opportunity to behave as in their natural foraging environment, pigeons behaved as presupposed by the model. These findings bear a close resemblance to those in the sign- and goal-tracking literature in that animals spatially adjust their behavior according to the signaling properties of the cue (e.g., Holland 1980; Silva, Silva and Pear 1992; Wasserman, Franklin and Hearst 1974).

Fortes and colleagues (2017) also studied the conditions under which pigeons would actively escape from a trial by pecking an escape key that cancelled the ongoing trial and initiated the ITI. To that end, they resorted to a prey model (Charnov 1976a; see also Krebs et al. 1977) modification of the model presented here and found that (a) the rate of escaping depended on the parameters, as the model predicted, and (b) that pigeons escaped from  $S_{G,0.0}$  but not from the other terminal or initial stimuli, as the model assumed. The rate of escaping from  $S_{G,0.0}$  was not as high as expected but that may have been caused by the asymmetry in response requirements: escaping required a peck whereas accepting required no peck, which may have biased pigeons toward not escaping (for other examples, see Lea 1979; Mazur 2007). Briefly, then, pigeons move away from  $S_{G,0.0}$  virtually always and often actively terminate it, thus upholding another model core assumption. Yet, physical removal from the  $S_{G,0.0}$  is not a prerequisite for animals to disengage. When the  $S_{G,0.0}$  is unavoidable animals will nonetheless disengage and show suboptimal preferences (Stagner, Laude and Zentall 2011; see also Laude, Stagner, and Zentall 2014).



### 3.3. Changing the reward probability of the Informative option does not change preference (P1)

This prediction was tested in both pigeons (Fortes et al. 2016; Experiment 1) and starlings (Vasconcelos et al. 2015; Experiment 1). Despite some procedural differences, such as the number of trials completed per day (420 in starlings vs. 120 in pigeons), the main ingredients of the task were the same between species:  $p_{non-info} = .5$ ,  $t = 10$  s, and the probabilities of seeing  $S_{B, 0.5}$  and  $S_{Y, 0.5}$  were .2 and .8, as usual. What varied systematically in both experiments was  $p_{info}$  or equivalently the probability of seeing  $S_{R, 1.0}$  (and consequently  $S_{G, 0.0}$ ). In both experiments,  $p_{info}$  was reduced by decreasing the proportion of  $S_{R, 1.0}$  trials in an orderly (starlings) or pseudorandom (pigeons) fashion from the typical .2 down to zero. The parameters used for starlings were .2, .15, .10, .05, and .0; the experiment with pigeons did not include .15 but included .025. The model predicts a preference for the Informative option except when  $p_{info} = 0$ .

Figure 5 shows the average preference for the Informative option as a function of  $p_{info}$  for both pigeons and starlings. As predicted, both species continued to prefer the Informative option even when its profitability was dramatically low, reversing preference only when  $p_{info} = 0$ . Taking  $p_{info} = .05$  as an example, when all within-trial delays are included in the rate computations (and assuming  $h = 1$ ), the Informative and Non-informative options yielded 0.286 and 2.857 reinforcers per minute, a nine-fold difference favoring the latter. However, when the time waiting for certain no reward is excluded, the rates become 5.455 and 2.609 reinforcers per minute for all  $p_{info}$  except  $p_{info} = 0$ . In other words, the Informative option is always 1.9 times richer than the Non-informative option. The slight modulation of preference by  $p_{info}$  (note that preference decreased slightly as  $p_{info}$  decreased) is probably caused by the minimum engagement time necessary to identify the

increasingly frequent  $S_{G, 0.0}$  before disregarding it. Altogether, the results of the two experiments corroborate the prediction (see also Roper and Zentall 1999).

Findings from other laboratories upheld the prediction, too. For instance, Gipson and colleagues (2009) used  $p_{info} = .5$  and  $p_{non-info} = .75$ , instead of the typical .2 vs. .5. If animals disengage from the stimulus signaling the absence of food, the effective choice would have been between a stimulus signaling certain reinforcement and a stimulus signaling reinforcement with probability .75. Smith and Zentall (2016), on the other hand, used  $p_{info} = .5$  and  $p_{non-info} = 1.0$  and therefore the effective choice would have been between two stimuli signaling certain reinforcement. Consistent with the predictions, pigeons preferred the Informative option in the former case and were indifferent in the latter.

### 3.4. Changing the duration of $S_{G, 0.0}$ does not change preference (P2)

Fortes and colleagues (2016; Experiment 2) tested this prediction in pigeons using a titration procedure. All parameters were initially set to the usual values but the duration of  $S_{G, 0.0}$  could increase, decrease or stay the same depending on the preference observed in the previous 10 choice trials. If the pigeon chose the Informative option on two trials or less, said duration decreased by 5 s; if it chose the Informative option on eight trials or more, the duration increased by 5 s; otherwise, the duration remained the same.

Figure 6 shows the  $S_{G, 0.0}$  duration in effect per block of trials for each bird. Initially, all birds started by preferring the Non-informative option, which caused the  $S_{G, 0.0}$  duration to decrease to zero. This initial preference for the Non-informative option is actually quite typical as Figure 2 illustrates (see also, for example, Laude et al. 2014; Zentall and Stagner 2011b). Yet, with continued exposure to the procedure, pigeons began to choose the Informative option more often,

causing a continuous increase in the  $S_{G,0.0}$  duration up to a maximum between 195 and 210 s. This upper limit was not an experimental finding (i.e., preference for the Informative option did not decrease to indifference with long  $S_{G,0.0}$  durations); rather, it was an experimenter imposed upper limit to maintain session duration within reasonable limits. Our guess is that the  $S_{G,0.0}$  duration would continue to increase otherwise (for other, less extreme examples, see Mazur 1995; Spetch et al. 1994; Pisklak et al. 2015).

Taking the 200 s  $S_{G,0.0}$  duration as an example and assuming as before  $h = 1$ , the Non-informative option was objectively 35 times better than the Informative option (the rates are 2.857 and 0.074 reinforcers per minute, respectively). Yet, the model predicts that the  $S_{G,0.0}$  duration is inconsequential and therefore the Informative option should remain 1.9 times richer than the Non-informative option. The results confirm the prediction.

### *3.5. Equalizing the duration of uncertainty across options eliminates preference for the Informative option (P3)*

Vasconcelos and colleagues (2015; Experiment 2) tested this prediction in starlings by maintaining the predictability of the outcome with each option unchanged and equalizing the duration of uncertainty across options. To that end, they moved the delay  $t$  (10 s) from the terminal link ( $S_{R,1.0}$  and  $S_{G,0.0}$ ) to the initial link of the Informative option such that (a) the trial duration remained the same, (b) predictability remained the same ( $S_{R,1.0}$  was followed by food but  $S_{G,0.0}$  was not), but (c) the duration of uncertainty was now roughly the same as in the Non-informative option; the Non-informative option remained unchanged.

Figure 7 depicts the modified procedure (top panel) and the average preference for the Informative option (bottom panel) with the original and the modified procedure. As expected,

preference for the Informative options develops with the original procedure (filled dots) but declines to and remains at floor with the modified one (unfilled dots).

### 3.6. Summary of findings

We have proposed a model inspired by optimal foraging theory that reconciles sub-optimal choice with the view that current behavioral mechanisms were pruned by the optimizing action of natural selection. To that end, and based on known psychological mechanisms and the statistical structure of typical foraging cycles, we assumed that, in the task depicted in Figure 1, animals learn the outcomes that follow each terminal stimulus (A1) and that they disengage from the stimulus signaling the absence of food ( $S_{G,0,0}$ ) but not from the other terminal stimuli (A2). These assumptions, in turn, led to specific and sometimes counterintuitive predictions. Specifically, given that  $S_{G,0,0}$  is disregarded immediately upon appearance, preference should be unaffected by its frequency (P1) and the delay it imposes (P2); also, even though predictability and time under uncertainty are confounded in the original procedure (cf. Figure 1), both are necessary for the effect to emerge. In fact, sub-optimal choice should vanish when the time under uncertainty is equalized across options (P3).

Overall, both the assumptions and the predictions passed the empirical tests. Animals do learn the outcomes of each terminal stimulus and choose appropriately between them. In addition, if given the opportunity, they will move away from the stimulus signaling certain no reward. Regarding the predictions, animals are indeed unaffected by the frequency of  $S_{G,0,0}$  as well as by the delay it imposes. At the extreme, when  $S_{G,0,0}$  appeared on 95% of the Informative trials (and therefore only 5% of these trials ended in reward), the animals still preferred the Informative option but the Non-informative option was objectively nine times better. When the duration of  $S_{G,0,0}$  was

200 s, they also still preferred the Informative option but the Non-informative option was objectively 35 times better. In the latter case, the animals systematically chose the Informative option, but 80% of such choices led to more than three minutes waiting for no reward. Thus, animals do disengage from the stimulus signaling the absence of food and perceive the Informative option as always leading to food (for a similar proposal and additional evidence, see Smith & Zentall, 2016). Finally, we showed that besides predictability, uncertainty duration also plays a significant role in the emergence of sub-optimal choice in this task.

#### 4. Concluding remarks

Failures to maximize and some other well known ‘irrationalities’ are sometimes invoked to question research programs focusing on the adaptive value of behavior and its underpinning mechanisms. We propose instead that such irrationalities reflect the ‘misbehavior’ of *evolved mechanisms* under certain circumstances and are therefore valuable tools to understand the adaptive significance of such mechanisms.

Our claim is that ultimate accounts such as those underlying optimality models do not presuppose that the behavior is optimal in every circumstance. Given the complexities of the natural environments, natural selection endowed the organisms not with specific mechanisms to deal optimally with each and every particular situation, but with general mechanisms adjusted to their typical ecology that perform well on the average (e.g., Fawcett et al. 2014; Fawcett, Hamblin, and Giraldeau 2013; McNamara and Houston 2009). Yet, under atypical circumstances, the animal will rely on the mechanisms that usually work, potentially generating maladaptive behaviors. The task we have discussed here is a paradigmatic example. Animals resort to mechanisms that are usually effective, but given that the task comprises features that deviate significantly from the

statistical structure of the natural environments, the outcome is sub-optimal. We showed how an optimality approach informed by known evolved mechanisms is able to, not only account for sub-optimal preferences, but also generate paradoxical predictions that subsequent experimentation has proven correct. The same general approach has been used to tackle many other issues in the animal learning and behavior literature, including intertemporal choice (e.g., Kacelnik 2003), risk-sensitive foraging (e.g., Bateson and Kacelnik 1998; Kacelnik and Bateson 1997), and even the close interplay between timing processes and foraging (e.g., Brunner, Kacelnik and Gibbon 1992; Kacelnik and Brunner 2002).

Both humans and non-humans often violate the principles of economic and biological rationality (*sensu* Kacelnik 2006). Sometimes they violate core principles (such as transitivity or regularity); sometimes they fail to maximize some benefit such as when they choose impulsively. These observations do not imply that breaches of rationality are the norm or that the principles of rationality are normatively inadequate. They simply remind us that, in order to understand why the mechanisms underpinning these deviations evolved, first we need to consider the adaptive value of such mechanisms in the animals' characteristic ecology. Existing mechanisms bear the footprint of the selective pressures of such ecology.

## References

- Baron, J. 2008. *Thinking and deciding*. Cambridge University Press, Cambridge.
- Bateson, M. 2002. Context-dependent foraging choices in risk-sensitive starlings. *Animal Behaviour*, 64: 251-260. doi:10.1006/anbe.2002.3059
- Bateson, M., Healy, S.D. and Hurly, T.A. 2002. Irrational choices in hummingbird foraging behaviour. *Animal Behaviour*, 63: 587-596. doi:10.1006/anbe.2001.1925
- Bateson, M. and Kacelnik, A. 1996. Rate currencies and the foraging starling: the fallacy of the averages revisited. *Behavioral Ecology*, 7: 341-352. doi:10.1093/beheco/7.3.341
- Bateson, M., & Kacelnik, A. (1998). Risk-sensitive foraging: Decision making in variable environments. In: R. Dukas (Editor), *Cognitive Ecology* (pp. 297-341). Chicago, IL: University of Chicago Press.
- Bateson, P. and Laland, K.N. 2013. Tinbergen's four questions: an appreciation and an update. *Trends in Ecology & Evolution*, 28: 712-718. doi:10.1016/j.tree.2013.09.013
- Brown, P. L. and Jenkins, H. M. 1968. Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 11: 1-8. doi:10.1901/jeab.1968.11-1
- Brunner, D., Kacelnik, A. and Gibbon, J. 1992. Optimal foraging and timing processes in the starling, *Sturnus vulgaris*: effect of inter-capture interval. *Animal Behaviour*, 44, 597-613. doi:10.1016/S0003-3472(05)80289-1
- Charnov, E.L. 1976a. Optimal foraging: Attack strategy of a mantid. *The American naturalist*, 110: 141-151.
- Charnov, E.L. 1976b. Optimal foraging: The marginal value theorem. *Theoretical Population Biology*, 9: 129-136. doi:10.1016/0040-5809(76)90040-x
- Darwin, C. 1859. *The origin of species*. John Murray, London.
- Dawkins, M.S. 2014. Tribute to Tinbergen: Questions and How to Answer Them. *Ethology*, 120: 120-122. doi:10.1111/eth.12186
- Dinsmoor, J.A., Mueller, K.L., Martin, L.T. and Bowe, C.A. 1982. The acquisition of observing. *Journal of the Experimental Analysis of Behavior*, 38: 249-263. doi:10.1901/jeab.1982.38-249
- Doyle, J.R., O'Connor, D.J., Reynolds, G.M. and Bottomley, P.A. 1999. The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store

- purchases. *Psychology and Marketing*, 16: 225-243. doi:10.1002/(SICI)1520-6793(199905)16:3<225::AID-MAR3>3.0.CO;2-X
- Fawcett, T.W., Fallenstein, B., Higginson, A.D., Houston, A.I., Mallpress, D.E.W., Trimmer, P.C. and McNamara, J.M. 2014. The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*, 18: 153-161. doi:10.1016/j.tics.2013.12.012
- Fawcett, T.W., Hamblin, S. and Giraldeau, L.A. 2013. Exposing the behavioral gambit: The evolution of learning and decision rules. *Behavioral Ecology*, 24: 2-11. doi:10.1093/beheco/ars085
- Fisher, R.A. 1930. *The genetical theory of natural selection*. Clarendon, Oxford.
- Fortes, I., Machado, A. and Vasconcelos, M. 2017. Do pigeons (*Columba livia*) use information about the absence of food appropriately? A further look into suboptimal choice. *Journal of Comparative Psychology*, 131: 277-289. doi:10.1037/com0000079
- Fortes, I., Vasconcelos, M. and Machado, A. 2016. Testing the boundaries of “paradoxical” predictions: Pigeons do disregard bad news. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42: 336-346. doi:10.1037/xan0000114
- Gardner, A. 2009. Adaptation as organism design. *Biology Letters*, 5: 861-864. doi:10.1098/rsbl.2009.0674
- Gigerenzer, G., Hertwig, R. and Pachur, T. 2011. *Heuristics: The foundations of adaptive behavior*. Oxford University Press, Oxford, UK.
- Gilovich, T., Griffin, D. and Kahneman, D. 2002. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press, Cambridge, UK.
- Gipson, C. D., Alessandri, J. J. D., Miller, H. C. and Zentall, T. R. 2009. Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning & Behavior*, 37: 289-298. doi:10.3758/lb.37.4.289
- Goldstein, D.G. and Gigerenzer, G. 2002. Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109: 75-90. doi:10.1037/0033-295x.109.1.75
- Grafen, A. 2006. Optimization of inclusive fitness. *Journal of Theoretical Biology*, 238: 541-563. doi:10.1016/j.jtbi.2005.06.009
- Grafen, A. 2007. The formal Darwinism project: A mid-term report. *Journal of Evolutionary Biology*, 20: 1243-1254. doi:10.1111/j.1420-9101.2007.01321.x



- Green, L., Fisher, E.B., Perlow, S. and Sherman, L. 1981. Preference reversal and self control: Choice as a function of reward amount and delay. *Behaviour Analysis Letters*, 1: 43-51.
- Haldane, J.B.S. 1932. *The causes of evolution*. Harper's, New York.
- Hamilton, W.D. 1964. The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7: 1-16. doi:10.1016/0022-5193(64)90038-4
- Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature*, 228: 1218-1220. doi:10.1038/2281218a0
- Haselton, M.G., Bryant, G.A., Wilke, A., Frederick, D.A., Galperin, A., Frankenhuis, W.E. and Moore, T. 2009. Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, 27: 733-763. doi:10.1521/soco.2009.27.5.733
- Hearst, E. and Jenkins, H.M. 1974. *Sign-tracking: The stimulus-reinforcer relation and directed action*. Psychonomic Society, Austin, TX.
- Herrnstein, R.J. 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4: 267-272. doi:10.1901/jeab.1961.4-267
- Herrnstein, R.J. 1990. Rational choice theory: Necessary but not sufficient. *American Psychologist*, 45: 356-367. doi:10.1037/0003-066X.45.3.356
- Herrnstein, R.J. and Heyman, G.M. 1979. Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior*, 31: 209-223. doi:10.1901/jeab.1979.31-209
- Hinnenkamp, J.E., Shahan, T.A. and Madden, G.J. 2017. How suboptimal is suboptimal choice? *Journal of the Experimental Analysis of Behavior*, 107: 136-150. doi:10.1002/jeab.239
- Holland, P.C. 1980. Influence of visual conditioned stimulus characteristics on the form of Pavlovian appetitive conditioned responding in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 6: 81-97. doi:10.1037/0097-7403.6.1.81
- Houston, A.I., McNamara, J.M. and Steer, M.D. 2007. Do we expect natural selection to produce rational behaviour? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362: 1531-1543. doi:10.1098/rstb.2007.2051
- Hurly, T.A. and Oseen, M.D. 1999. Context-dependent, risk-sensitive foraging preferences in wild rufous hummingbirds. *Animal Behaviour*, 58: 59-66. doi:10.1006/anbe.1999.1130

- Kacelnik, A. 2003. The evolution of patience. In: G. Loewenstein, D. Read, and R. Baumeister (Editors), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 115-138). New York: Russell Sage Foundation.
- Kacelnik, A. 2006. Meanings of rationality. In: S. Hurley and M. Nudds (Editors), *Rational Animals?* (pp. 87-106). Oxford University Press, Oxford, New York.
- Kacelnik, A. 2012. Putting mechanisms into behavioral ecology. In: P. Hammerstein and J.R. Stevens (Editors), *Evolution and the mechanisms of decision making* (pp. 21-38). MIT Press, Cambridge, MA.
- Kacelnik, A. and Bateson, M. 1997. Risk-sensitivity: crossroads for theories of decision-making. *Trends in Cognitive Sciences*, 1, 304-309. doi:10.1016/S1364-6613(97)01093-0
- Kacelnik, A. and Brunner, D. 2002. Timing and foraging: Gibbon's Scalar expectancy theory and optimal patch exploitation. *Learning and Motivation*, 33: 177-195. doi:10.1006/lmot.2001.1110
- Kacelnik, A. and Todd, I.A. 1992. Psychological mechanisms and the Marginal Value Theorem: Effect of variability in travel time on patch exploitation. *Animal Behaviour*, 43: 313-322. doi:10.1016/S0003-3472(05)80226-X
- Kalenscher, T., Tobler, P.N., Huijbers, W., Daselaar, S.M. and Pennartz, C.M.A. 2010. Neural signatures of intransitive preferences. *Frontiers in Human Neuroscience*, 4, 49. doi:10.3389/fnhum.2010.00049
- Krebs, J.R., Erichsen, J.T., Webber, M.I. and Charnov, E.L. 1977. Optimal prey selection in the great tit (*Parus major*). *Animal Behaviour*, 25: 30-38. doi:10.1016/0003-3472(77)90064-1
- Latty, T. and Beekman, M. 2011. Irrational decision-making in an amoeboid organism: Transitivity and context-dependent preferences. *Proceedings of the Royal Society B: Biological Sciences*, 278: 307-312. doi:10.1098/rspb.2010.1045
- Laude, J.R., Stagner, J.P. and Zentall, T.R. 2014. Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, 40: 12-21. doi:10.1037/xan0000010
- Lea, S.E. 1979. Foraging and reinforcement schedules in the pigeon: Optimal and non-optimal aspects of choice. *Animal Behaviour*, 27: 875-886. doi:10.1016/0003-3472(79)90025-3
- Logue, A.W., Smith, M.E. and Rachlin, H. 1985. Sensitivity of pigeons to preinforcer and postreinforcer delay. *Animal Learning & Behavior*, 13: 181-186. doi:10.3758/BF03199271

- Mayr, E. 1963. *Animal species and evolution*. Harvard University Press, Cambridge, MA.
- Mazur, J.E. 1987. An adjusting procedure for studying delayed reinforcement. In: M.L. Commons, J.E. Mazur, J.A. Nevin and H. Rachlin (Editors), *Quantitative analyses of behavior: The effect of delay and of intervening events on reinforcement value*. (Vol. 5, pp. 55–73). Erlbaum, Hillsdale, NJ.
- Mazur, J.E. 1989. Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior*, 51: 87-99. doi:10.1901/jeab.1989.51-87
- Mazur, J.E. 1995. Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *Journal of the Experimental Analysis of Behavior*, 63: 139-150. doi:10.1901/jeab.1995.63-139
- Mazur, J.E. 2000. Tradeoffs among delay, rate, and amount of reinforcement. *Behavioural Processes*, 49: 1-10. doi:10.1016/S0376-6357(00)00070-X
- Mazur, J.E. 2007. Choice in a successive-encounters procedure and hyperbolic decay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 88: 73-85. doi:10.1901/jeab.2007.87-06
- McDevitt, M.A., Dunn, R.M., Spetch, M.L. and Ludvig, E.A. 2016. When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, 105: 23-40. doi:10.1002/jeab.192
- McNamara, J.M. and Houston, A.I. 2009. Integrating function and mechanism. *Trends in Ecology and Evolution*, 24: 670-675. doi:10.1016/j.tree.2009.05.011
- Monteiro, T., Vasconcelos, M. and Kacelnik, A. 2013. Starlings uphold principles of economic rationality for delay and probability of reward. *Proceedings of the Royal Society B: Biological Sciences*, 280. doi:10.1098/rspb.2012.2386
- Pisklak, J.M., McDevitt, M.A., Dunn, R.M. and Spetch, M.L. 2015. When good pigeons make bad decisions: Choice with probabilistic delays and outcomes. *Journal of the Experimental Analysis of Behavior*, 104: 241-251. doi:10.1002/jeab.177
- Rand, J.F. 1977. Behaviors observed during S- in a simple discrimination learning task. *Journal of the Experimental Analysis of Behavior*, 27: 103-117. doi:10.1901/jeab.1977.27-103
- Roper, K. L. and Zentall, T. R. 1999. Observing behavior in pigeons: The effect of reinforcement probability and response cost using a symmetrical choice procedure. *Learning and Motivation*, 30: 201-220. doi:10.1006/lmot.1999.1030

- Scott-Phillips, T.C., Dickins, T.E. and West, S.A. 2011. Evolutionary theory and the ultimate–proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science*, 6: 38-47. doi:10.1177/1745691610393528
- Seth, A.K. 2007. The ecology of action selection: Insights from artificial life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362: 1545-1558. doi:10.1098/rstb.2007.2052
- Shafir, S. 1994. Intransitivity of preferences in honey bees: Support for 'comparative' evaluation of foraging options. *Animal Behaviour*, 48: 55-67. doi:10.1006/anbe.1994.1211
- Shafir, S., Waite, T.A. and Smith, B.H. 2002. Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, 51: 180-187. doi:10.1007/s00265-001-0420-8
- Silva, F.J., Silva, K. and Pear, J.J. 1992. Sign- versus goal-tracking: Effects of conditioned-stimulus-to-unconditioned-stimulus distance. *Journal of the Experimental Analysis of Behavior*, 57: 17-31. doi:10.1901/jeab.1992.57-17
- Smith, A. P. and Zentall, T. R. 2016. Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42: 212-220. doi:10.1037/xan0000092
- Spetch, M.L., Mondloch, M.V., Belke, T.W. and Dunn, R. 1994. Determinants of pigeons' choice between certain and probabilistic outcomes. *Animal Learning & Behavior*, 22: 239-251. doi:10.3758/BF03209832
- Stagner, J.P., Laude, J.R. and Zentall, T.R. 2011. Sub-optimal choice in pigeons does not depend on avoidance of the stimulus associated with the absence of reinforcement. *Learning and Motivation*, 42: 282-287. doi:10.1016/j.lmot.2011.09.001
- Stagner, J.P., Laude, J.R. and Zentall, T.R. 2012. Pigeons prefer discriminative stimuli independently of the overall probability of reinforcement and of the number of presentations of the conditioned reinforcer. *Journal of Experimental Psychology: Animal Behavior Processes*, 38: 446-452. doi:10.1037/a0030321
- Stagner, J.P. and Zentall, T.R. 2010. Suboptimal choice behavior by pigeons. *Psychonomic Bulletin and Review*, 17: 412-416. doi:10.3758/PBR.17.3.412

- Stephens, D.W. and Anderson, D. 2001. The adaptive value of preference for immediacy: When shortsighted rules have farsighted consequences. *Behavioral Ecology*, 12: 330-339. doi:10.1093/beheco/12.3.330
- Stephens, D.W., Brown, J.S. and Ydenberg, R.C. 2007. *Foraging: Behavior and ecology*. University of Chicago Press, Chicago.
- Stephens, D.W. and Krebs, J.R. 1986. *Foraging theory*. Princeton University Press, Princeton, NJ.
- Stevens, J.R. and Stephens, D.W. 2010. The adaptive nature of impulsivity. In: G.J. Madden and W.K. Bickel (Editors), *Impulsivity: The behavioral and neurological science of discounting* (pp. 361–387). American Psychological Association, Washington, DC.
- Tinbergen, N. 1963. On aims and methods of Ethology. *Zeitschrift für Tierpsychologie*, 20: 410-433. doi:10.1111/j.1439-0310.1963.tb01161.x
- Todd, I.A. and Kacelnik, A. 1993. Psychological mechanisms and the marginal value theorem: Dynamics of scalar memory for travel-time. *Animal Behaviour*, 46: 765-775. doi:10.1006/anbe.1993.1254
- Tversky, A. 1969. Intransitivity of preferences. *Psychological Review*, 76: 31-48. doi:10.1037/h0026750
- Tversky, A. and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185: 1124-1131. doi:10.1126/science.185.4157.1124
- Tversky, A. and Kahneman, D. 1986. Rational choice and the framing of decisions. *Journal of Business*, 59: S251-S278.
- Vasconcelos, M., Fortes, I. and Kacelnik, A. 2017. On the Structure and Role of Optimality Models in the Study of Behavior. In: J. Call (Editor), *APA Handbook of Comparative Psychology* (Vol. 2, pp. 287-307). American Psychological Association, Washington, DC.
- Vasconcelos, M., Monteiro, T. and Kacelnik, A. 2015. Irrational choice and the value of information. *Scientific Reports*, 5, 13874. doi:10.1038/srep13874
- Waite, T.A. 2001. Intransitive preferences in hoarding gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, 50: 116-121. doi:10.1007/s002650100346
- Wasserman, E. A., Franklin, S. R. and Hearst, E. 1974. Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of Comparative and Physiological Psychology*, 86: 616-627. doi:10.1037/h0036171
- Wright, S. 1931. Evolution in mendelian populations. *Genetics*, 16: 97-159.

- Zentall, T.R. 2014. Suboptimal choice by pigeons: An analog of human gambling behavior. *Behavioural Processes*, 103: 156-164. doi:10.1016/j.beproc.2013.11.004
- Zentall, T.R. 2016. Resolving the paradox of suboptimal choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42: 1-14. doi:10.1037/xan0000085
- Zentall, T.R., Laude, J.R., Stagner, J.P. and Smith, A.P. 2015. Suboptimal Choice by Pigeons: Evidence that the Value of the Conditioned Reinforcer Rather than its Frequency Determines Choice. *Psychological Record*, 65: 223-229. doi:10.1007/s40732-015-0119-2
- Zentall, T.R. and Stagner, J. 2011a. Maladaptive choice behaviour by pigeons: an animal analogue and possible mechanism for gambling (sub-optimal human decision-making behaviour). *Proceedings of the Royal Society B: Biological Sciences*, 278: 1203-1208. doi:10.1098/rspb.2010.1607
- Zentall, T.R. and Stagner, J. 2011b. Sub-optimal choice by pigeons: Failure to support the Allais paradox. *Learning and Motivation*, 42: 245-254.

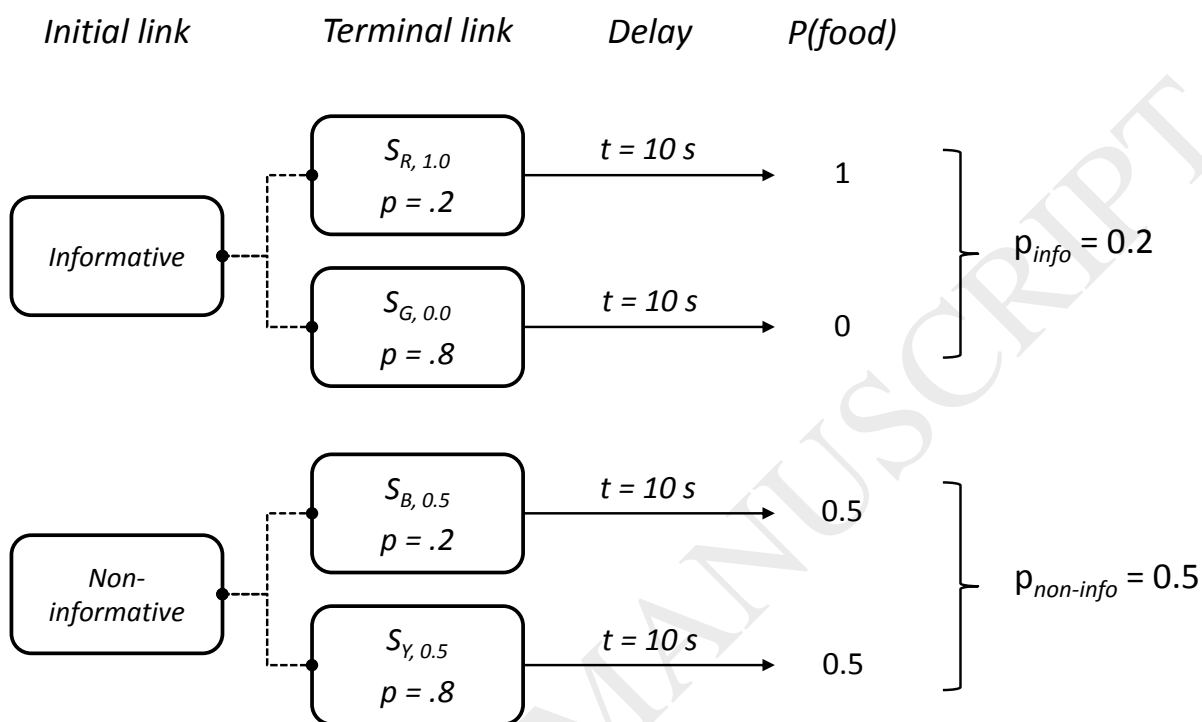
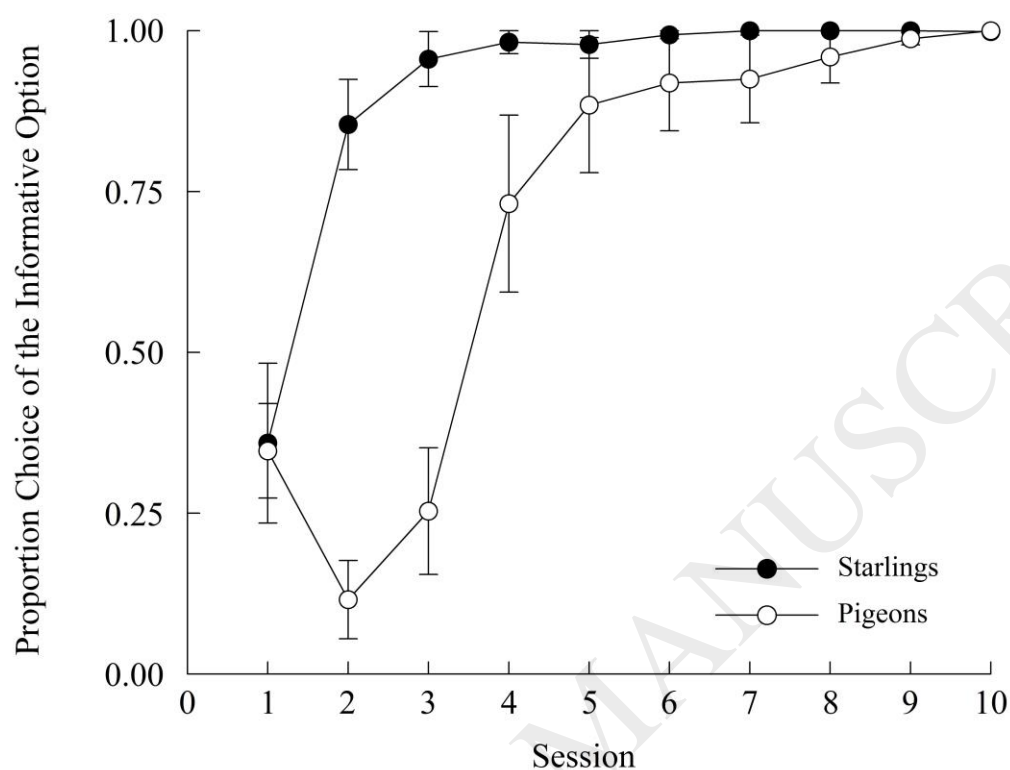
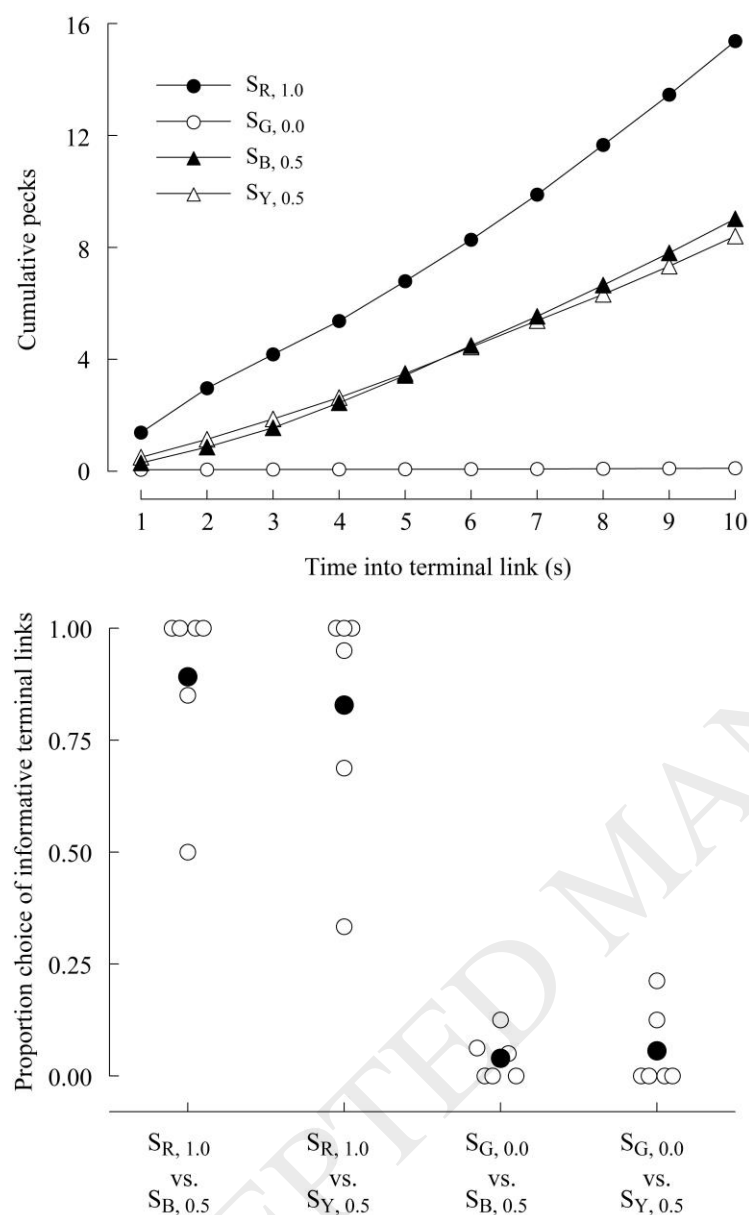


Figure 1. Schematic of events and outcomes in the typical sub-optimal choice task. Dotted lines indicate no delay between consecutive events and solid lines indicate a delay of  $t$  s (typically 10).



*Figure 2.* Mean proportion of choices of the Informative option across sessions both in starlings ( $n = 6$ ; data from Vasconcelos et al. 2015, Experiment 1) and in pigeons ( $n = 8$ ; data from Fortes et al. 2017; first 10 sessions of Experiment 1). The error bars are the standard error of the mean.





*Figure 3.* Top panel: Average cumulative pecks to each terminal stimulus as a function of time into the stimulus in starlings ( $n = 6$ ). Data from Vasconcelos et al. (2015; Experiment 2, control group). Bottom panel: Average preference between terminal stimuli in starlings ( $n = 6$ ). Filled and unfilled dots represent the average and individual data, respectively. Data from Vasconcelos et al. (2015; Experiment 1).

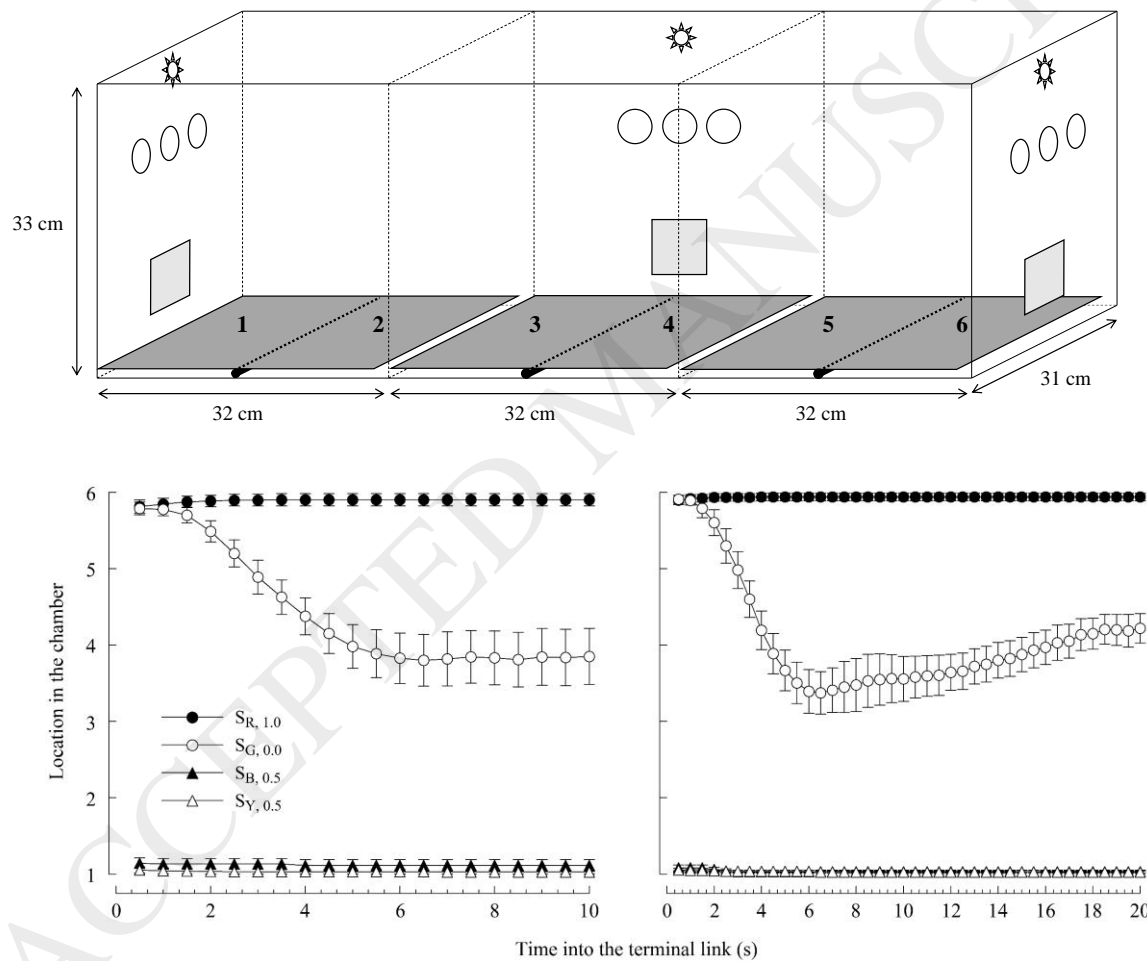
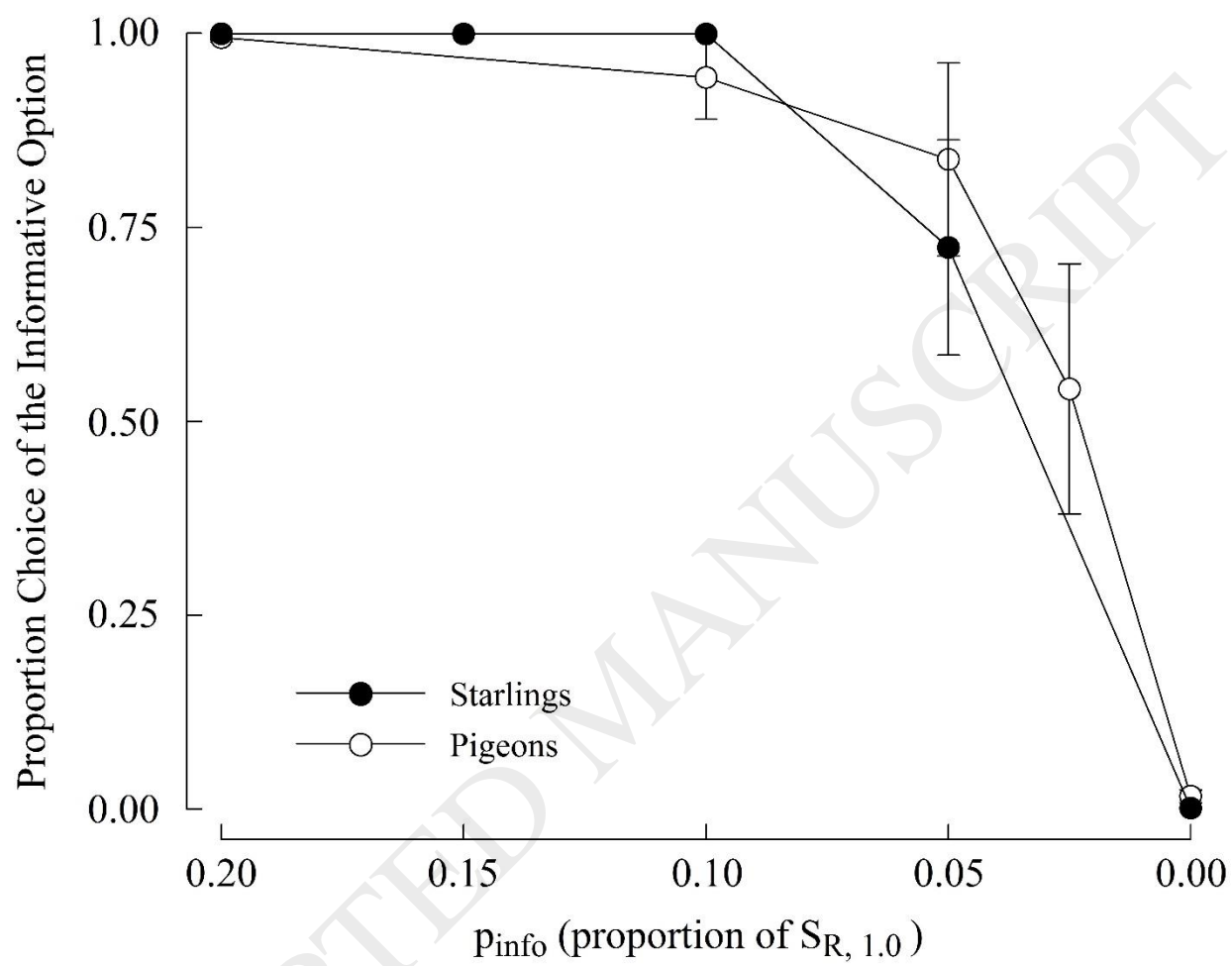
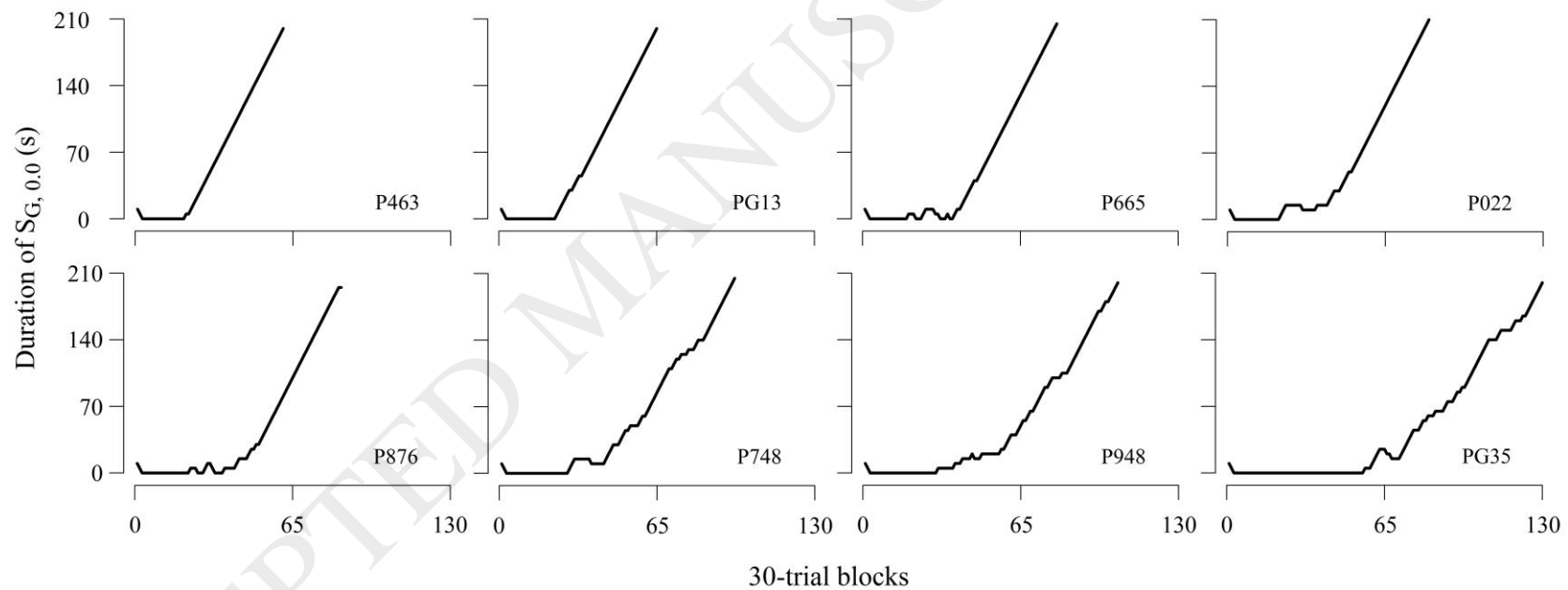


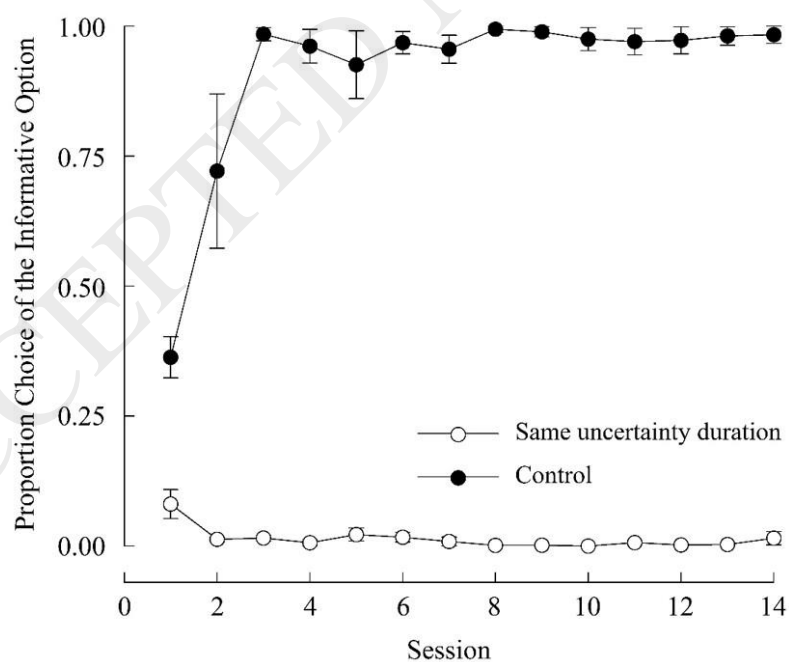
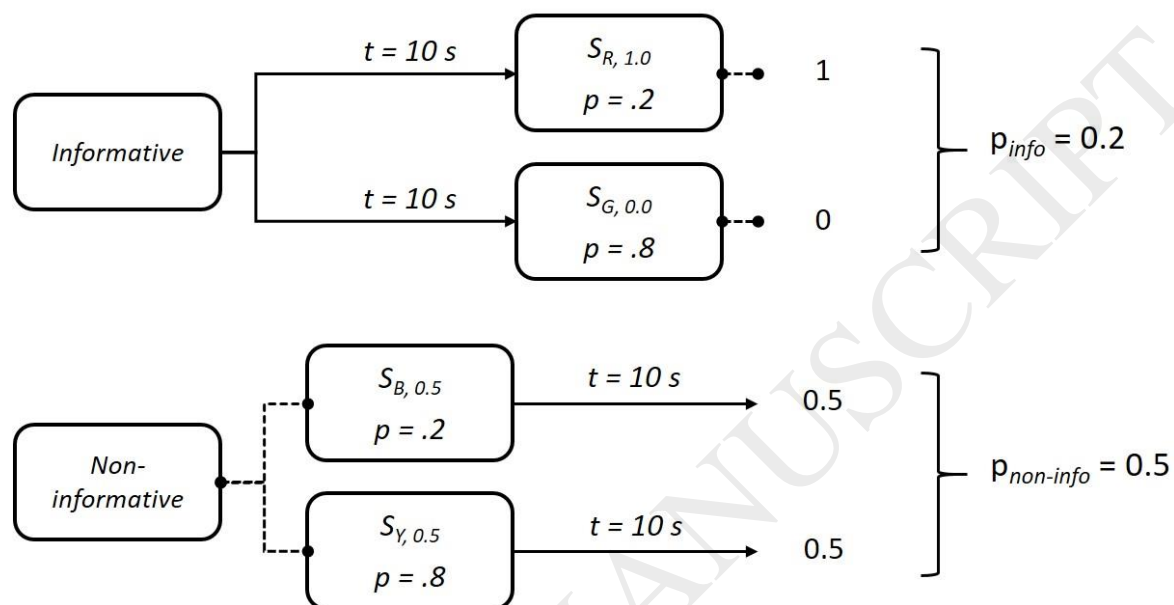
Figure 4. Top panel: Diagram of the long chamber used by Fortes et al. (2017; Experiment 2). Bottom panel: Average location of the pigeons ( $n = 4$ ) during the 10-s (left) and the 20-s (right) terminal stimulus, respectively. The error bars are the standard error of the mean. Data from Fortes et al. (2017; Experiment 2).



*Figure 5.* Mean proportion of choices of the Informative option, at stability, as a function of the probability of reinforcement in that option. The error bars are the standard error of the mean. Starlings' ( $n = 6$ ) and pigeons' ( $n = 6$ ) data are from Vasconcelos et al. (2015; Experiment 1) and Fortes et al. (2016; Experiment 1), respectively.



*Figure 6.* Duration of the stimulus signaling the absence of reward ( $S_{G,0,0}$ ) in each 30-trial block of the adjusting-delay condition in Experiment 2 of Fortes et al. (2016). Within each 30-trial block, the first 20 trials were forced-choice trials (10 with the Informative and 10 with the Non-informative option, interspersed) and the last 10 trials were free-choice trials with the two options available. Copyright 2017 by the American Psychological Association. Adapted with permission.



*Figure 7.* Top panel: Schematic of the modified procedure used by Vasconcelos et al. (2015; Experiment 2). Dotted lines indicate no delay between consecutive events and solid lines indicate a delay of 10 s. Bottom panel: Mean proportion of choices of the Informative option across sessions in the protocol equalizing uncertainty duration ( $n = 6$ ) and the typical protocol ( $n = 6$ ). The error bars are the standard error of the mean. Data from Vasconcelos et al. (2015; Experiment 2).