



Universidade do Minho
Escola de Psicologia

Valeria Vanessa González Díaz **Integration and use of information in the suboptimal choice task**

Valeria Vanessa González Díaz

**Integration and use of information in the
suboptimal choice task**

UMinho | 2020

fevereiro de 2020



Universidade do Minho
Escola de Psicologia

Valeria Vanessa González Díaz

**Integration and use of information in the
suboptimal choice task**

Tese de Doutoramento
Doutoramento em Psicologia Básica

Trabalho efetuado sob a orientação do
Professor Doutor Armando Machado,
Professor Doutor Marco Vasconcelos

e sob orientação da
Professora Doutora Montserrat Comesaña

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição-NãoComercial-Compartilhalgal
CC BY-NC-SA

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

“Suppose that we are wise enough to learn and know—and yet not wise enough to control our learning and knowledge, so that we use it to destroy ourselves? Even if that is so, knowledge remains better than ignorance. It is better to know—even if the knowledge endures only for the moment that comes before destruction—than to gain eternal life at the price of a dull and swinish lack of comprehension of a universe that swirls unseen before us in all its wonder. That was the choice of Achilles, and it is mine, too.”

— Isaac Asimov —

ACKNOWLEDGEMENTS

I would like to thank my advisors, Professors Armando Machado, Marco Vasconcelos and Montserrat Comesaña. Armando, for all the feedback, for taking the time to think twice my ideas; but especially for your company every day when my health was at stake. Marco, for give me the opportunity to be here, for all the feedback, the last minutes meetings, and for teach me to slow down. Montserrat, for your time and good disposition to help me.

Thank you, colleagues at the Learning and Behavior Lab for the help, for the distractions, for the CC meetings and the friendship. Thank you, Renata for your invaluable support! For hold me every time I cried. Thank you, Catarina and your parents, for be there when I needed most. To the last losers, thank you Cristina for being a big sister, for all the advice and efforts to help me. Thanks, Harry for the coffees, music and movie recommendations that distracted me on rainy days. And Alex! For everything, for your company on those long days in the lab, for always be available to cover me, for the discussions about experiments, data analysis, ideas and life; for your patience, for listening, for caring.

Thank you, Professor Aaron Blaisdell, for welcoming me into your lab. I am deeply grateful for everything you shared, for your availability, and the motivating discussions. My time in your lab was a reminder of how much I love to do research. Thank you, Mary and Ben for all our talks, for the help and support; you made me always feel welcome. Thank you, to all the students, especially Eric, Liz, Hannah and Rachel. I would also like to thank Professor Kate Wassum and your team for welcoming and taught me much in such a short period of time.

Gracias, Paulo y Marcela, por darme un hogar cuando más lo necesité, por las risas y la buena comida. Gracias a mi familia, mamá y papá, Caro, Ale, Lili y Leo, por dejarme ir a soñar lejos, por aceptar verme crecer al otro lado del mundo, y comprender que la distancia física no cambia el amor.

Gracias a mis amigos, la familia elegida, los de Chile, los que emigraron, los del mundo. Ustedes saben quiénes son. Por estar presentes, por las risas, los abrazos y el amor. Thank you, Wren, for your love and support even when I could not stand myself.

Finally, I would like to thank the Portuguese Foundation for Science and Technology (FCT) for the scholarship that made all this possible (ref. PD/BD/114368/2016) and the University of Minho and CIPsi for the facilities.

STATEMENT OF INTEGRITY

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration. I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, February 3rd, 2020.

Full name: Valeria Vanessa González Díaz

Signature:

A handwritten signature in black ink, appearing to read 'Valeria Vanessa González Díaz', written over a horizontal line.

INTEGRAÇÃO E USO DA INFORMAÇÃO NA TAREFA DE ESCOLHA SUBOPTIMA

Os animais vivem em ambientes altamente incertos, e qualquer informação que reduza a incerteza é considerada valiosa uma vez que os ajuda a prever eventos futuros e a comportarem-se de acordo com os mesmos. A presente dissertação explora uma situação em que os pombos podem trocar comida por informação: usando um procedimento de esquemas encadeados concorrentes, os pombos escolhem entre duas alternativas sendo que cada uma delas leva a dois sinais possíveis que, após um atraso, são seguidos por comida com uma determinada probabilidade. As alternativas diferem (1) na probabilidade global de reforço (2) na probabilidade de reforço de cada sinal. Nesta tarefa, os animais preferem a alternativa mais informativa mesmo quando esta é globalmente menos reforçada; a maioria dos modelos teóricos explicam este fenómeno assumindo que os pombos sobre-estimam os sinais que são sempre reforçados e ignoram os sinais não reforçados. O objetivo desta dissertação foi avaliar sistematicamente duas das variáveis que tem sido sugeridas como responsáveis por esta escolha subótima. Propomos um novo modelo, a hipótese Delta-Sigma, que considera que a diferença na probabilidade de reforço entre os sinais, Delta (Δ); e o rácio das probabilidades globais de reforço das alternativas, Sigma (Σ), podem explicar o comportamento dos pombos. Na experiência 1, avaliamos o papel de Δ e na experiência 2, o papel de Σ . Os resultados foram consistentes com a hipótese Delta-Sigma, demonstrando que os pombos preferem o Δ maior e que a preferência é modulada por Σ . Nas experiências 3 e 4, testamos a independência da construção do Δ , demonstrando que Δ iguais construídos com diferentes probabilidades de reforço podem ter valores diferentes. Os resultados são parcialmente explicados pela nossa hipótese e sugerem que existem outros fatores que podem afetar a preferência. Finalmente, as experiências 5 e 6 propõe um mecanismo sobre o que é aprendido acerca do sinal não reforçado. Os resultados demonstraram que este sinal desenvolveu propriedades inibitórias e que estas propriedades estão correlacionadas com o desenvolvimento da preferência subótima. No geral, os resultados aqui apresentados são a primeira avaliação sistemática das variáveis sugeridas como responsáveis pela preferência subótima.

Palavras-chave: escolha, esquemas encadeados concorrentes, informação, pombos

INTEGRATION AND USE OF INFORMATION IN THE SUBOPTIMAL CHOICE TASK

Animals live in highly uncertain environments, and any information that reduces uncertainty is considered valuable as it aids to predict upcoming events and behave accordingly. The present dissertation explores a situation in which pigeons can trade food for information: using a concurrent-chain procedure, pigeons choose between two alternatives, each one of them leads to two possible signals that, after a delay, are followed by food with a certain probability. The alternatives differed in (1) the overall probability of reinforcement (2) each signals' individual probability of reinforcement. In this task, animals prefer the most informative alternative even when this leads to less reinforcement overall; most theoretical accounts explain this phenomena assuming that pigeons overestimate signals that are always reinforced and ignore non-reinforced signals. The objective of this dissertation was to systematically evaluate two of the variables that have been suggested to be responsible for this suboptimal choice. We propose a new model, the Delta-Sigma hypothesis, which states that the difference in probability of reinforcement between the signals, Delta (Δ); and the ratio of the overall probability of reinforcement between alternatives, Sigma (Σ) can account for pigeons' behavior. In Experiment 1, we evaluated the role of Δ . In Experiment 2, we evaluated the role of Σ . The results were align with the Delta-Sigma hypothesis, finding that pigeons preferred the greater Δ and that the preference is modulated by Σ . In experiments 3 and 4, we tested the independency of the construction of the Δ , finding that not all equal Δ built with different probabilities of reinforcement have the same value. The results are partially explained by our hypothesis and suggest that might be other factors affecting preference. Finally, experiments 5 and 6 proposed a mechanism for the signal associated with absence of reinforcement. The results found that the signal developed inhibitory properties and that these are correlated with the development of the suboptimal preference. Overall, the results presented here are the first systematic evaluation of variables suggested as responsible of the suboptimal preference.

keywords: choice, concurrent-chain, information, pigeons

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	III
RESUMO	VI
ABSTRACT.....	VII
ABBREVIATIONS, ACRONYMS AND SYMBOLS.....	X
FIGURES	XI
TABLES	XV
CHAPTER I - INTRODUCTION.....	1
Observing response.....	2
The suboptimal choice task.....	4
Models of suboptimal choice	6
Zentall’s contrast-like account.....	7
Reinforcement Rate Model	7
Hyperbolic Discount Model.....	9
Temporal Information approach	10
The present dissertation	11
CHAPTER II - THE Δ - Σ HYPOTHESIS: HOW CONTRAST AND REINFORCEMENT RATE COMBINE TO GENERATE SUBOPTIMAL CHOICE	13
Abstract.....	14
Introduction.....	15
Experiment 1	22
Method	22
Results and Discussion	25
Experiment 2.....	28
Method	29
Results and Discussion	30
General Discussion	33
CHAPTER III - TESTING THE Δ - Σ MODEL: SAME DELTA WITH DIFFERENT PROBABILITIES OF REINFORCEMENT	41
Abstract.....	42
Introduction.....	43

Experiment 3	47
Method	47
Results and discussion	51
Experiment 4	55
Method	55
Results and discussion	57
General discussion	61
CHAPTER IV - THE ROLE OF INHIBITION IN THE SUBOPTIMAL CHOICE TASK	67
Abstract	68
Introduction	69
Experiment 5	72
Subjects	72
Materials and Methods	73
Procedure	74
Results	75
Experiment 6	78
Subjects	79
Materials and Methods	79
Procedure	80
Results	81
General discussion	89
CHAPTER V - CONCLUSIONS	92
Two high-order variables are responsible for the preference	93
The value of information	95
The mechanism of a never reinforced stimulus	96
REFERENCES	98

ABBREVIATIONS, ACRONYMS AND SYMBOLS

ANOVA – Analysis of Variance

β – Sensitivity to Delta

C – Sensitivity to Sigma

Δ – Delta, difference in probability of reinforcement within an alternative

FR – Fixed Ratio schedule

FT – Fixed Time

HDM – Hyperbolic Discount Model

IL – Initial link

ITI – Intertrial Interval

p 's – Probability of reinforcement of a terminal link

r 's – Probability of occurrence of a terminal link

RR – response rate

RRM – Reinforcement Rate Model

Σ – Sigma, overall probability of reinforcement

TL – Terminal link

FIGURES

CHAPTER I – INTRODUCTION

Figure 1. Wyckoff observing response procedure. After a 30 s delay, the pigeon had to peck the white key to get or not to get food reward depending on the trial type (left). If, meanwhile, the animal pressed the pedal, the key turned either red or green indicating that the final peck would be reinforced (middle) or not (right), respectively.

Figure 2. Suboptimal choice task. Left panel shows the procedure used by Kendall (1974). Right panel shows the procedure used by Stagner and Zentall (2010).

CHAPTER II – THE Δ - Σ HYPOTHESIS: HOW CONTRAST AND REINFORCEMENT RATE COMBINE TO GENERATE SUBOPTIMAL CHOICE

Figure 3. Suboptimal choice task. Left panel: Stagner and Zentall's (2010) procedure, in which animals chose between the suboptimal (i.e. Alternative 1) or optimal (i.e. Alternative 2) alternative. Once a choice was made, animals saw one of two possible terminal colors per alternative that remained on for 10s. Each color occurred 20% (S+ and S3) or 80% of the time (S- and S4). S+ was always reinforced, S- was never reinforced, and S3 and S4 are reinforced with a probability of .5. The overall probability of reinforcement was .2 and .5 for suboptimal and optimal alternative, respectively. Right panel: General task. Animals are confronted with a two-alternative concurrent-chain schedule. IL1 and IL2 correspond to the Initial links; r1 and r2 are the probabilities of occurrence of the terminal links in each alternative (TL1,1 and TL1,2 for Alternative 1, TL2,1 and TL2,2 for Alternative 2). p corresponds to the probability of reinforcement after each terminal link (p1,1, p1,2, p2,1, and p2,2 for TL1,1, TL1,2, TL2,1, and TL2,2, respectively); d is the terminal-link duration; Σ refers to the overall probability of reinforcement in each alternative.

Figure 4. Average preference for Alternative 1 across Δ diffs. The first five graphs show individual data and the last graph shows the average preference across pigeons. The black dots show preference when one (or two for the middle black dot) of the alternatives had stimuli perfectly correlated with the presence and absence of food. The grey dots show the other cases. Lines correspond to linear fittings of the data for both data sets. Error bars in the bottom right graph correspond to the standard error of the mean.

Figure 5. Average latencies to peck each alternative in Experiment 1. Latencies are shown by condition for each Δ diff. Dark gray and white bars refer to latencies for Alternatives 1 and 2, respectively. The error bars correspond to standard error of the mean.

Figure 6. Average preference for Alternative 1 as a function of $\Sigma 1$. The first six graphs show individual data and the bottom graph shows the average preference across pigeons. Lines correspond to the linear fitting of the data. Error bars in the bottom graph correspond to the standard error of the mean.

Figure 7. Average latencies to peck each alternative across conditions in Experiment 2. Gray and white bars refer to latencies for Alternative 1 and 2, respectively. The error bars correspond to the standard error of the mean.

Figure 8. Average fit of the Delta-Sigma hypothesis (solid line) for Experiments 1 (left panel) and 2 (right panel). The predicted preference was estimated for each animal across conditions, then the fits were averaged. Error bars are the standard error of the mean.

CHAPTER III – TESTING THE Δ - Σ MODEL: SAME DELTA WITH DIFFERENT PROBABILITIES OF REINFORCEMENT

Figure 9. Suboptimal choice task. Left panel: typical procedure, in which animals choose with one peck in the initial links, a cross for the suboptimal (i.e. Alternative 1) or a circle for the optimal (i.e. Alternative 2) alternative. After the choice, animals see with probability .2 or .8, one out of two terminal links signaled by color, each lasting for 10s. The terminal link Red is always reinforced, Green is never reinforced; Yellow and Blue are reinforced half of the time. The overall probability of reinforcement was .2 and .5 for suboptimal and optimal alternative, respectively. Right panel: General task. Animals are confronted to a concurrent-chain schedule with two links. IL1 and IL2 correspond to the Initial links where the choice is made. r_1 and r_2 are the probabilities of occurrence of the terminal links in each alternative. $p_{1,1}$ and $p_{1,2}$ are the probabilities of reinforcement for Alternative 1, $p_{2,1}$ and $p_{2,2}$ for Alternative 2. The difference of those correspond to $\Delta 1$ and $\Delta 2$. Finally, Σ_1 and Σ_2 refer to the overall probability of reinforcement of each alternative.

Figure 10. Choice data Experiment 3. The average preference for a $\Delta = .5$ on the last four sessions for each condition is presented by pigeon. The bottom panel shows the average preference for each condition. Error bars show SEM.

Figure 11. Averaged median latency data on forced trials of Experiment 3. Black dots show latency for the $\Delta = .5$, white dots show latencies for alternatives with $\Delta = 0$. Error bars correspond to the standard error of the mean.

Figure 12. Choice data Experiment 4. The average preference for a $\Delta = .5$ of the last four sessions for the two conditions is presented by pigeon. The bottom right panel shows the average preference. Error bars show SEM.

Figure 13. Averaged median latency data for forced trials in Experiment 4. Black dots show latencies for alternatives with $\Delta = .5$, white dots show latencies for alternatives with $\Delta = 0$. Error bars correspond to the standard error of the mean.

Figure 14. Preference for the Delta of .5 in experiments 3 and 4 by pigeon.

CHAPTER IV – THE ROLE OF INHIBITION IN THE SUBOPTIMAL CHOICE TASK

Figure 15. Procedure used by Laude, Stagner and Zentall (2014) to evaluate the inhibitory properties of the stimulus never reinforced (S-, black circle with a white vertical line). After the choice, each signal was followed by a different magnitude of reinforcement.

Figure 16. Left panel: Examples of stimuli presented during the Pre-training and Training phases of Experiment 5. Right panel: Examples of compound stimuli presented on probe trials during the Testing phase of Experiment 5. Color and orientation were counterbalanced across pigeons with the restriction that S+ and S3, and S- and S4 should have the same orientation to be able to create the compound stimuli.

Figure 17. Mean normalized RR for each block of five sessions of training (Upper panel) and testing (Lower panel) for each stimulus in Experiment 5. Error bars represent 95% confidence interval.

Figure 18. Mean normalized RR across the last two blocks of testing for each probe stimulus in Experiment 5. Error bars correspond to the 95% confidence interval.

Figure 19. Left panel: Structure of the suboptimal choice task used on choice trials during training in Experiment 6. The initial link stimuli were counterbalanced across pigeons. Terminal link stimuli could be two pairs of figures (diamonds, triangles, circles, or 4-point stars) each one filled with a different pattern and presented vertically (S+ and S3) or horizontally (S- and S4). In this example, the vertical pair of diamonds correspond to the S+, the horizontal pair of triangles correspond to the S-, the vertical pair of circles correspond to the S3, and the horizontal pair of stars correspond to the S4. Right panel: Examples of four possible compound stimuli used during testing in Experiment 6. All stimuli used black patterns over a white background. Stimuli were presented half of the time in each side of the screen within each test session.

Figure 20. Proportion of suboptimal preference during training (left panel) and testing (right panel) in blocks of five sessions by pigeon in Experiment 6. As with the data analysis, for Herriot blocks 2 and 3 of training, and 1 and 2 for test were averaged for a better comparison.

Figure 21. Mean normalized RR to each stimulus by 5-session block of training in Experiment 6. For Herriot, blocks 2 and 3 of training were averaged. Error bars correspond to the 95% confidence interval.

Figure 22. Mean normalized RR across all blocks of testing for each stimulus in Experiment 6. Error bars correspond to the 95% confidence interval.

Figure 23. Normalized RR for each element and compound on probe trials for each pigeon in Experiment 6. Solid colors depict elements, black for S+ and grey for S3. Stripes depict compounds, black and white for S+S-, and grey and white for S3S-.

Figure 24. The correlation between suboptimal preference and log RR to S- during the 6 blocks of training for each pigeon in Experiment 6. The bottom right panel shows the mean correlation across all pigeons. Dashed lines represent the best linear fitting. r = Pearson's correlation.

TABLES

CHAPTER II –THE Δ - Σ HYPOTHESIS: HOW CONTRAST AND REINFORCEMENT RATE COMBINE TO GENERATE SUBOPTIMAL CHOICE

Table 1. Experimental conditions of Experiment 1.

Table 2. Order of experimental conditions for each pigeon in Experiment 1.

Table 3. Experimental conditions of Experiment 2.

Table 4. Order of experimental conditions for each pigeon in Experiment 2.

Table 5. Failures and successes of sub-optimal choice models.

Table 6. Failures and successes of sub-optimal choice models.

Table 7. Suboptimal choice studies using the 3-stimuli procedure.

CHAPTER III – TESTING THE Δ - Σ MODEL: SAME DELTA WITH DIFFERENT PROBABILITIES OF REINFORCEMENT

Table 8. Suboptimal choice studies using the 3-stimuli procedure.

Table 9. Order of experimental conditions in Experiment 3.

Table 10. Stimulus assignments for Experiment 4.

Table 11. Order of experimental conditions in Experiment 4.

Table 12. Averaged, minimum and maximum obtained probability of reinforcement of each terminal link stimuli.

CHAPTER IV – THE ROLE OF INHIBITION IN THE SUBOPTIMAL CHOICE TASK

Table 13. Averaged rate of response for stimuli by block for each pigeon for Experiment 6.

CHAPTER I

INTRODUCTION

Observing response

When L. B. Wyckoff published his dissertation in 1969, he wanted to extend our understanding of discrimination learning, the process by which organisms learn to respond under one specific stimulus but not to another. To do that Wyckoff developed the following procedure (see Figure 1): pigeons were exposed to a series of 30-s trials in which a white key was presented. In some trials, a peck after the 30 s was reinforced with food; in other trials, pecks were never reinforced. At the same time that the white key was presented, pigeons had the option to press a pedal on the floor, which, when pressed, indicated the type of trial in effect: if the trial was going to be reinforced, the white key turned red, and if the trial was not going to be reinforced, the white key turned green (see Wyckoff, 1952). Therefore, the response of pressing the pedal was an ‘observing response’, which did not change the schedule of reinforcement but informed the pigeon about which trial was in effect. Wyckoff (1969) found that animals pressed the pedal significantly more when it produced stimuli informing which schedule was in effect than when it produced stimuli uncorrelated with the current schedule.

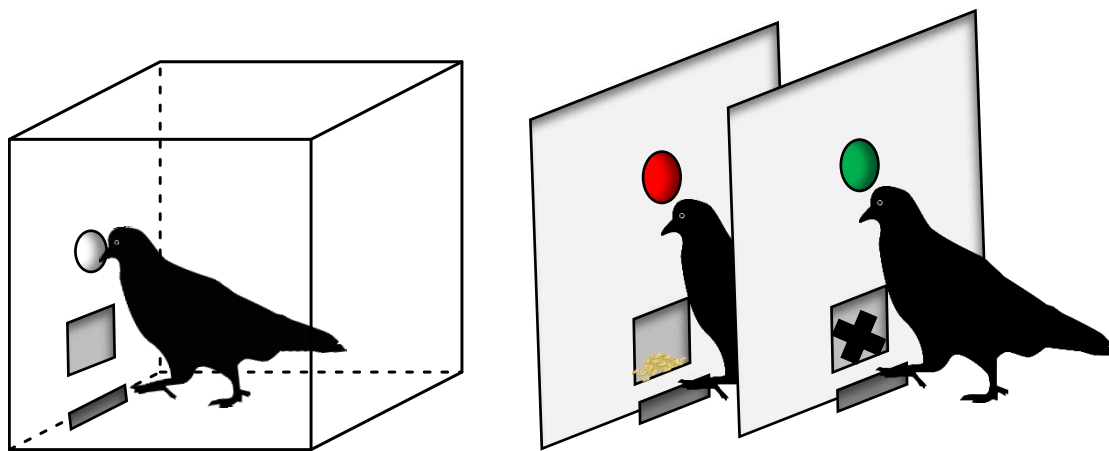


Figure 1. Wyckoff observing response procedure. After a 30 s delay, the pigeon had to peck the white key to get or not to get food reward depending on the trial type (left). If, meanwhile, the animal pressed the pedal, the key turned either red or green indicating that the final peck would be reinforced (middle) or not (right), respectively.

The concept of ‘observing response’ opened a line of research questioning whether and, in the affirmative, why animals would press a pedal that did not change the amount or quality of reinforcement. Apparently, animals do engage in the observing response. What is more, they perform observing response even when the access to reinforcement is response-independent (i.e., when the peck was not required to get food) (Browne & Dinsmoor, 1974).

One of the earliest explanations for this finding emphasized the value of ‘information’. Information itself ought to be a reinforcer because the observing response had not further consequence. The consequence of the response was a hue perfectly correlated with trial outcome. This idea, derived from information theory (Shannon & Weaver, 1949), holds that preference for information about presence and absence of food should be equally valuable. However, further research has found that information about the presence or absence of reinforcement is not equivalent (Dinsmoor, 1983). Indeed, it seems that animals prefer information yielding ‘good news’ (e.g., information about reinforcement) to information yielding ‘bad news’ (e.g., information about the absence of reinforcement or a punishment). For instance, Hirota (1972) found that pigeons spent more time pressing a pedal leading to a signal for sure reinforcement than they did when it led to a signal for no reinforcement. Similarly, Browne and Dinsmoor (1974) designed an experiment wherein the access to food was response-independent and the informative stimuli were diffuse lights illuminating the entire experimental box. The aim was to disentangle the observing response from any sign- or goal-tracking behavior. A tilting floor was used as the manipulandum for the observing response: walking to the left or right side of the chamber turned on the lights informing about the presence or absence of food, respectively. Therefore, the behaviors to obtain information about the presence or absence of food were mutually exclusive and literally orthogonal to the consummatory response.

The results indicated that pigeons prefer information about food over absence of food. Although interesting, these experiments do not indicate whether or not animals learn about bad news, they just seem to suggest that they prefer information about good news over bad news. Results coming from studies with humans have showed that humans preferred bad news over no news (Fantino & Silberberg, 2010; Lieberman, Cathro, Nichol, & Watson, 1997). Yet most studies have focused on the variables influencing the emission of the observing response rather than on the preference for informative or discriminative stimuli. Prokasy (1956) evaluated this preference by placing rats in the middle of an E-maze. Rats could choose the left or the right arm. After the choice, the rat had to wait 30 s before getting access to the goal box, which could be either baited or empty. The overall probability of reinforcement was equal in both arms, but one arm gave information: the walls were white when baited and black when not baited; the other arm had the same colors, but uncorrelated with the outcome. Under these contingencies, animals preferred the informative arm. It is important to notice that the rats were not producing the observing response, but choosing the

alternative that gave information about whether food would be delivered or not. This preference for the informative option was later replicated with pigeons (Bower, Mclean, & Meacham, 1966; Roper & Zentall, 1999).

The suboptimal choice task

In the choice procedures previously mentioned, the probability of reinforcement was unaffected by the observing response. Animals could choose to be informed about the outcome with minimal costs (usually one response). Kendall (1974) wondered if animals were willing to pay some cost in exchange for information. Using a concurrent-chain procedure, pigeons had to choose between two alternatives. As shown in Figure 2 (left panel), if Alternative 1 was chosen, half of the times the pecked key changed to a color stimulus, say green, that was always reinforced after a 15 s delay (S+); in the remaining half of choices, another color stimulus would appear, say red that after 15 s always ended without reinforcement (S-). On the other hand, if Alternative 2 was chosen, the pecked key turned to a third color stimulus, say yellow that was always followed by food after the 15 s delay elapsed (S3). The results showed a strong preference for Alternative 1. Given that the overall probability of reinforcement in Alternative 1 was .5 but 1.0 in Alternative 2, some authors called this preference ‘suboptimal’ because pigeons failed to maximize food intake. These findings gave some support to the notion that animals would be willing to pay some cost (i.e., give up some food) in exchange for information about the contingencies of reinforcement in effect.

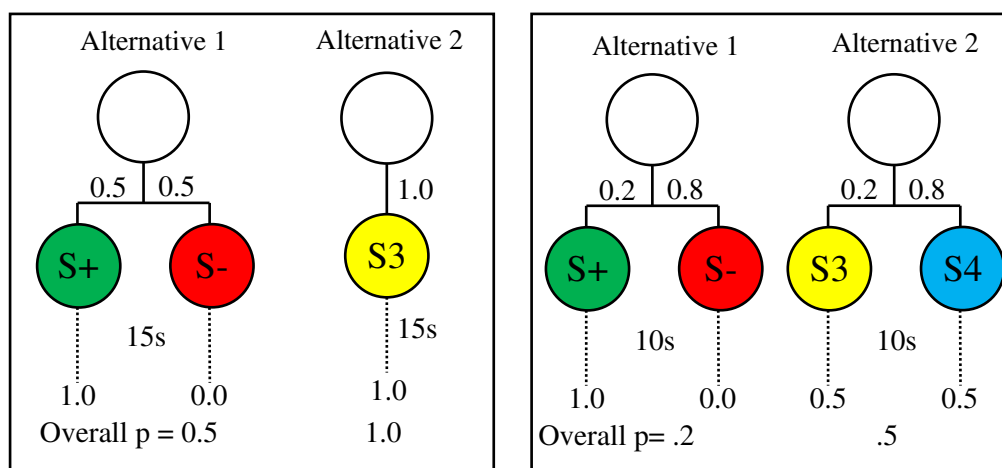


Figure 2. Suboptimal choice task. Left panel shows the procedure used by Kendall (1974). Right panel shows the procedure used by Stagner and Zentall (2010).

Subsequent studies have been inconsistent in their findings. Some have reported results similar to those reported by Kendall (Dunn & Spetch, 1990, Experiment 3), and others have been unable to replicate his findings (Dunn & Spetch, 1990, Experiment 2; Fantino, Dunn, & Meck, 1979). Recent research attempted to clarify the inconsistent results found with Kendall's procedure. A recent replication of the original experiment found indifference between alternatives, a result that is still suboptimal, but also the emergence of preference for Alternative 1 when training is extended (Case & Zentall, 2018). Perhaps the inconsistent of results stems from the fact that both available alternatives are informative: Alternative 1 led to stimuli (S+ or S-) perfectly correlated with the outcome but Alternative 2 also led to a stimulus (S3) that was always reinforced; in other words, S3 was associated with reinforcement with a 1.0 probability, just like the S+ in Alternative 1. Gipson, Alessandri, Miller and Zentall (2009) used the same task but decreased the overall probability of reinforcement in Alternative 2 from 1 to .75, and found a preference for Alternative 1.

In an effort to evaluate the suboptimal preference and the variables controlling this behavior, Stagner and Zentall (2010) used the procedure shown in Figure 2, right panel. This procedure is nowadays the task used most often to study suboptimal choice. Similar to Kendall's procedure, pigeons have to choose between two alternatives presented side by side. A single peck at the white key on the left side (Alternative 1), turns the right white key off, and the pecked key is replaced with one of two colors: with probability .2, the key turns green for 10 s and is always followed by reinforcement (S+); with the complementary probability (.8), the key turns red and ends without food after the same delay (S-). The overall probability of reinforcement for this alternative is thus .2. If, instead, the pigeon pecks the right, white key, the left key turns off and the right key changes to one of two colors: with probability of .2 to yellow, and with probability of .8 to blue, both of which end with reinforcement half of the time once the 10 s delay elapses. The overall probability of reinforcement for Alternative 2 is thus .5. As expected from previous findings, pigeons showed a strong and reliable preference for Alternative 1, a suboptimal preference. It is again worth to notice that in this case, a preference for the suboptimal alternative is also a preference for the informative option (Alternative 1 in our description).

In the last 20 years, great effort has been made to unravel the factors affecting animals' choice in this task. Some of these factors have been summarized by McDevitt, Dunn, Spetch and Ludvig (2016): 1) the contingency in the signals: the preference for the

leaner option (Alternative 1 in Figure 2) only occurs when that option has signals informing about the presence and absence of reinforcement, and the richer option does not have informative signals; 2) Initial response requirement: when the number of responses to choose increases, the preference for the informative alternative decreases; 3) The delay between signal and food: longer delays to food lead to increased preference for the informative and suboptimal option; 4) The contiguity between the choice and the signal: delaying the onset of the informative signals reduces the preference for that alternative; 5) the overall probability of reinforcement: a preference for the suboptimal alternative has been observed even when it gives significantly less food than the other alternative. However, it remains unclear if what matters is the absolute or relative difference between alternatives; and 6) Factors related to self-control: for instance, impulsive subjects appear to show stronger suboptimal preferences (for other reviews, see Zentall, 2015, 2016a, 2016b).

Most of these results have been obtained in experiments with pigeons or starlings, in which a reliable and strong suboptimal preference has been observed (Daniels & Sanabria, 2018; Vasconcelos, Machado, & Pandeirada, 2018; Zentall, 2016a). Attempts to replicate paradoxical choice in rats have yielded conflicting results. Failures led researchers to suggest intrinsic differences between species (Trujano, López, Rojas-Leguizamón, & Orduña, 2016; Trujano & Orduña, 2015). Positive results have been achieved when the task has been adapted to fit rats' perceptual features; for instance: use of sounds instead of lights as signals to reinforcement (Chow, Smith, Wilson, Zentall, & Beckmann, 2017; Ojeda, Murphy, & Kacelnik, 2018) or increasing the delay to reward (Cunningham & Shahan, 2019). Considering recent experiments, we surmise that suboptimal choice is a strong effect across species, and thus the mechanisms behind this effect are likely to be similar -if not the same.

Models of suboptimal choice

Suboptimal choice preference seems to be a robust effect that appears under many different conditions and species. In the last few years, great efforts have been made to understand the mechanisms underlying this apparently suboptimal behavior. Here, we review some of the models and hypotheses that have been proposed.

Zentall's contrast-like account

One of the earliest explanations described suboptimal choice as a case of contrast (Zentall, 2016a). Considering the procedure in Figure 2 right panel, when the animal chooses Alternative 1, its expectation is given by the overall probability of reinforcement in this option: .2 (this option is reinforced only 20% of the time). If the signal that is always followed by food (S+) appears, this expectation increases from .2 to 1.0, because this signal is always followed by reinforcement, thus leading to a contrast of +.8. The rest of the time, when the signal never followed by food (S-) appears, the expectation drops from .2 to 0, a contrast of -.2. In consequence, it is assumed, the contrast of S- has null or little impact in the choice. When the animal chooses Alternative 2 instead, the expectation is .5 and remains at .5 when the signal (S3 or S4) is lit because the probability of reinforcement is still .5, thus leading a contrast of 0.

This general explanation has no mathematical formalization, but it has worked as a fruitful framework to generate questions and interpret findings. Although interesting, the hypothesis does not address how the contrasts of each signal in an alternative are combined. Furthermore, Stagner, Laude, and Zentall (2012) found that when given a choice between two informative alternatives (i.e. both alternative containing an S+ and S-) that only differed in the overall rate of reinforcement (.2 for Alternative 1, .5 for Alternative 2), pigeons are indifferent. Even though the preference remained objectively suboptimal, if it was a matter of contrast, they should have preferred the alternative yielding the greater contrast (Alternative 1).

Reinforcement Rate Model

The Reinforcement Rate Model (RRM) is an ecological model based on optimal foraging theory. Optimality is the backbone of evolutionary approaches to behavioral sciences, thus the model postulates that animals should behave optimally under all conditions. The model proposes that the attractiveness of an option is a function of the long-term rate it affords. For instance, the animal may search for food and, after s seconds, finds prey i , pursues it for t time, and catches it with probability p and consumes it for h seconds. With probability $1-p$, the prey escapes and the cycle starts again. Thus, the rate of intake for the cycle involving prey i , is given by

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

Vasconcelos, Monteiro and Kacelnik (2015), analyzed the suboptimal choice procedure using the RRM. They assumed that p_i corresponds to the probability of reinforcement with each alternative, s corresponds to the intertrial interval (ITI), t is the delay of the signals, h is the time in which the food is available and consumed. Notice that t and h are equal across conditions. Because R_i monotonically increases with the value of p_i , Equation 1 predicts a preference for the alternative with higher overall rate of reinforcement (Alternative 2, in

Figure 2). However, the authors claimed that in natural settings, animals do not pay the cost of waiting when signals informing about the absence of food are available. That would be an unnecessary opportunity cost. Instead, organisms use that information to redirect their foraging efforts. Because animals may apply the same mechanisms in the laboratory settings, it is reasonable to conclude that animals may ignore signals for the absence of reinforcement (S-). In consequence, the S- stimulus does not influence the decision process. So, for Alternative 1, the suboptimal and informative option, the equation reduces to:

$$R_1 = \frac{1}{\frac{s}{p_{S+}} + t + h} \quad (2)$$

with p_{S+} corresponding to the probability of reinforcement of S+ (In Figure 2, right panel equal to 1).

On the other hand, for Alternative 2, the optimal and non-informative option, animals have to actively wait until the end of the trial because no reliable information about the trial outcome is available. Thus, the value of Alternative 2 is given by:

$$R_2 = \frac{1}{\frac{s+t}{p_{S3/S4}} + h} \quad (3)$$

With $p_{S3/S4}$ corresponding to the probability of reinforcement of the signals in alternative 2 (In Figure 2, right panel equal to .5). Applying equations 2 and 3 to the suboptimal choice task, the RRM predicts a preference for Alternative 1 (See Fortes, Vasconcelos, & Machado, 2016).

However, the model immediately encounters difficulties when confronted with probabilities of reinforcement in the S- different from 0. For instance, if the S- is now reinforced one time but it is not reinforced in another 19 trials, is the stimulus still ignored? Fortes, Pinto, Machado and Vasconcelos (2018) answered empirically to this question. They reported an experiment where the S- was followed by reinforcement in seldom occasions. As a consequence, the probability of reinforcement with the S- was no longer 0 and the case for ignoring this stimulus less tenable. They found that preference was delayed as more trials with the 'S-' signal were reinforced, but by the end of training it reached a similar, almost exclusive preference for Alternative 1. On a theoretical level, the authors modified the RRM adding a linear engagement function, in which the probability of reinforcement after a signal determines the likelihood of engagement (i.e., animals are more likely to pay attention to stimuli associated with a high probability of reinforcement). With this modification, the RRM can explain most of the literature using the suboptimal choice task, including conditions where the informative alternative has signals associated to a probability of reinforcement different from 1 and 0.

Hyperbolic Discount Model

The Hyperbolic Discount Model (HDM; Mazur, 1987, 1984) is a descriptive model stating that the current value of a delayed reward decreases hyperbolically with delay to reinforcement. Thus, the value is given by

$$V_i = \frac{S_i}{1+kD_i} \quad (4)$$

where S_i refers to the subjective value of the reinforcer if available immediately, k is a free parameter to estimate discount rate, and D_i is the delay between the choice and reward delivery. Variations in k are used to describe differences between individuals and between species.

For the HDM only delays in the presence of conditioned reinforcer, that is, stimuli associated with an appetitive outcome such as food, are taken into account. Delays in the presence of an S- do not affect the option's value. This means that the value of Alternative 1 in Figure 2 is affected only by the consequence and the delay associated to the S+. In practice, it is as if Alternative 1 always results in food. In choice situations (Alternative 1 vs. Alternative 2), the model predicts a preference for Alternative 1. (Mazur, 1996, 1995).

Temporal Information approach

Cunningham and Shahan (2018) proposed an informational approach to tackle suboptimal preference. They suggest that animals choose the suboptimal alternative because the signals predicting food provide relevant information about *when* and *where* food will be available. This proposal also explicitly rejects the notion that a stimulus that does not signal food could have any impact on choice (Blanchard, 1975).

Formally, the model adopts the temporal information approach to Pavlovian conditioning (Gallistel & Gibbon, 2000) and Shannon's entropy (Shannon & Weaver, 1949), to which the value of information corresponds to

$$H_i = \log_2 \frac{C}{t} \quad (5)$$

where i refers to the alternative, C corresponds to the Cycle time (i.e., the average time from food to food which is constant across conditions) and t to the Trial time (i.e., the delay to food when a signal with or without information about reinforcement is presented).

The value of each alternative is transformed to preference, using Equation 6:

$$p1 = \frac{V_1^a}{V_1^a + V_2^a} \quad (6)$$

where a is a scaling parameter, and V_1 and V_2 refer to the values of Alternatives 1 and 2, respectively.

Thus far, the proposal assumes that the suboptimal choice preference depends mainly in the value of the signals (i.e., $S+$, $S3$ and $S4$) following the choice response. However, it is known that the overall probability of reinforcement influences choice (Shahan & Cunningham, 2015). Indeed, we cannot call a preference suboptimal, if the animal does not perceive the differences in overall probability of reinforcement between alternatives. In a situation in which the alternatives only differed in this variable, the equation to calculate preference is as follows,

$$p1 = \frac{R_1^b}{R_1^b + R_2^b} \quad (7)$$

where R_1 and R_2 are the overall rate of reinforcement for alternative 1 and 2, respectively; b is a scaling parameter.

In summary, the temporal informational approach proposes two sources of control that compete to determine choice: signal value (Equation 6) and the relative rate of reinforcement (Equation 7). The influences of these two sources of control are weighed according to

$$p1 = w \frac{V_1^a}{V_1^a + V_2^a} + (1 - w) \frac{R_1^b}{R_1^b + R_2^b} \quad (8)$$

where w is a parameter, taking values between 0 and 1.

Although the model fits empirical data quite well, it also has several shortcomings, all detailed in Cunningham and Shahan (2018). Of relevance for the present dissertation is the assumption that the S- is ignored and does not directly influence the decision. This is in fact an assumption common to all models aforementioned. Furthermore, the approach cannot deal with probabilities of reinforcement for the terminal stimuli of a given alternative different from 1 and 0. For Alternative 2 (Figure 2, right panel), and even though two signals exist (S3 and S4), the model simplifies it to a single terminal stimulus reinforced 50% the time.

In summary, the models and approaches propose different mechanisms and reasons for suboptimal choice. Yet, some assumptions seem to trespass all of them. In particular, the disproportionate impact of the S+, and/or the null effect of the S- on choice, has been the cornerstone for most models developed so far.

The present dissertation

The main goal of the present dissertation was to further evaluate the role of some of the key variables that have been proposed to explain suboptimal choice. A subsidiary goal was to test the ability of a newly developed hypothesis to fit the empirical data. The dissertation is organized in five chapters as follows:

In Chapter I, a revision of the literature and models regarding suboptimal choice is presented. The key variables highlighted in literature are described, namely (1) the information given by the signals presented after the choice (i.e., S+, S-, S3 and S4), and (2) the overall rate of reinforcement.

Chapter II describes two experiments that parametrically manipulated each of the two variables mentioned above. We advance a new model, the Delta-Sigma hypothesis, which

proposes that (1) the difference or contrast in probability of reinforcement within the signals of each alternative (Delta) and the (2) ratio between the overall probabilities of reinforcement (Sigma) can explain suboptimal choice. Experiment 1, kept all variables constant, except the difference in probability of reinforcement within alternatives. On the contrary, Experiment 2 kept all variables constant except the ratio of the overall rate of reinforcement per alternative. In light of the results of these experiments, we tested the new hypothesis and compared its suitability against Zentall's contrast-like account, the RRM and Mazur's model (who mathematically gave similar predictions as RRM), and the Temporal Information approach. Finally, we extended the Delta-Sigma hypothesis to the existing literature.

Chapter III reports two experiments that extended the results of Chapter II, by testing a prediction of the Delta-Sigma hypothesis. According with the model, animals should assign the same value to two equal deltas, independently of the specific probabilities of reinforcement used to obtain them. In this study, animals were presented with a choice between two alternatives that differed in the Delta. One alternative had a constant Delta of 0 (i.e. both signals reinforced with the same probability) and the other varied with different deltas of .5, each built from different probabilities of reinforcement. The Delta-Sigma hypothesis was applied to the results, together with the same models used in Chapter II.

In Chapter IV, the role of S- is assessed. Bear in mind that most models assume that animals ignore the signal for the absence of food. However, it is not completely clear if 'ignore' means that the value of the signal does not take part on the decision process, or if the animals do not learn about the stimulus. In order to disentangle the role of the S-, two experiments were carried out in which conditioned inhibition and its relationship with the development of suboptimal choice was explored. Furthermore, we discussed the implications of the results for the models in the area.

Finally, Chapter V summarizes the empirical results and their theoretical implications for the suboptimal choice literature.

CHAPTER II ¹

**THE Δ - Σ HYPOTHESIS: HOW
CONTRAST AND REINFORCEMENT
RATE COMBINE TO GENERATE
SUBOPTIMAL CHOICE**

¹ This Chapter reproduces the submitted version of the paper:

González, V. V., Macías, A., Machado, A., & Vasconcelos, M. (Accepted). The Δ - Σ hypothesis: How contrast and reinforcement rate combine to generate suboptimal choice. *Journal of the Experimental Analysis of Behavior*.

Abstract

When given a choice between two alternatives, each offering food after the same delay with different but signaled probabilities, pigeons often prefer the low probability alternative. This preference is surprising because pigeons fail to maximize the rate of food intake; they exhibit a suboptimal preference. We advance a new explanation, the Δ - Σ hypothesis, in which the difference in probability of reinforcement within terminal links (Δ) and the overall reinforcement probability rate of each alternative (Σ) are the key variables responsible for such suboptimal preference. We tested the Δ - Σ hypothesis in two experiments. In Experiment 1, we manipulated the Δ s while maintaining constant all other parameters of the task, in particular the Σ s. We predicted a preference for the alternative with the larger Δ . In Experiment 2, we examined the effect of the overall reinforcement probabilities, the Σ s, while maintaining constant all other parameters of the task, in particular the Δ s. We predicted a preference for the larger Σ . The results of both experiments support the Δ - Σ hypothesis.

Given a choice between two alternatives, each offering food after the same delay but with different probabilities, pigeons may prefer the low probability alternative. This preference surprises us because, in a seemingly simple task, pigeons fail to maximize rate of food intake (Vasconcelos, Machado, & Pandeirada, 2018). The left panel of Figure 3 illustrates the typical task in which pigeons exhibit suboptimal performance. The reinforcement schedule is a two-alternative concurrent-chain. Two keys, one with a Plus figure and one with a Circle figure, correspond to the initial links. A single peck at the Plus key turns off the other key and replaces the Plus with a color, Red with probability 0.2, and Green with probability 0.8; the color initiates the terminal link. After a 10-s delay, if the key is Red, the pigeon always receives food; if the key is Green, the pigeon never receives food. Hence, the overall probability of receiving food given a peck at the Plus key equals 0.2. On the other hand, a single peck on the Circle during the initial link turns off the Plus key and replaces the Circle with a color, Yellow with probability 0.2, and Blue with probability 0.8. After a 10 s delay, and regardless of the key color, the pigeon receives food on half of the trials. Hence, the overall probability of receiving food given a peck at the Circle key equals 0.5. Pigeons and starlings consistently prefer the Plus (suboptimal) alternative to the Circle (optimal) alternative (Fortes, Machado, & Vasconcelos, 2017; Fortes et al., 2018, 2016; Stagner et al., 2012; Stagner & Zentall, 2010; Vasconcelos et al., 2015; Zentall & Stagner, 2011).

To compare different accounts of suboptimal choice, we need to identify the key variables in the task and to that end we need to introduce a new terminology. As the right panel of Figure 3 shows, each choice alternative comprises an Initial-Link (IL_1 or IL_2) and two Terminal-Links ($TL_{1,1}$ and $TL_{1,2}$, for alternative 1; $TL_{2,1}$ and $TL_{2,2}$ for alternative 2). The first subscript identifies the alternative and the second subscript identifies the terminal link within an alternative. The total number of links equals 6, and they are distinctly signaled, the two initial-links with geometric figures (e.g., Plus and Circle) and the four terminal-links with colors (e.g., Red, Green, Blue, and Yellow). When the pigeon chooses IL_1 , then with probability r_1 it enters terminal link $TL_{1,1}$ and with the complementary probability $1-r_1$ it enters terminal link $TL_{1,2}$. Similarly, when the pigeon chooses IL_2 , then it enters terminal links $TL_{2,1}$ or $TL_{2,2}$ with probabilities r_2 and $1-r_2$, respectively. All terminal links last d seconds. After d , the trial ends with food according to the probability of reinforcement associated with the terminal link, $p_{1,1}$ for $TL_{1,1}$, $p_{1,2}$ for $TL_{1,2}$, $p_{2,1}$ for $TL_{2,1}$, and $p_{2,2}$ for $TL_{2,2}$.

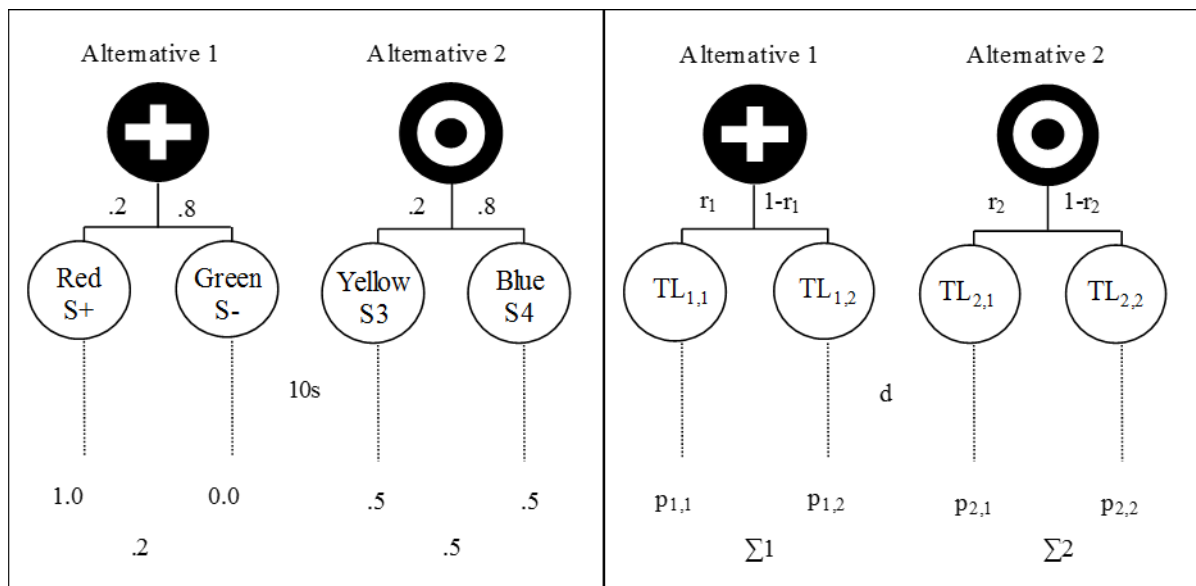


Figure 3. Suboptimal choice task. Left panel: Stagner and Zentall's (2010) procedure, in which animals chose between the suboptimal (i.e. Alternative 1) or optimal (i.e. Alternative 2) alternative. Once a choice was made, animals saw one of two possible terminal colors per alternative that remained on for 10s. Each color occurred 20% (S+ and S3) or 80% of the time (S- and S4). S+ was always reinforced, S- was never reinforced, and S3 and S4 are reinforced with a probability of .5. The overall probability of reinforcement was .2 and .5 for suboptimal and optimal alternative, respectively. Right panel: General task. Animals are confronted with a two-alternative concurrent-chain schedule. IL_1 and IL_2 correspond to the Initial links; r_1 and r_2 are the probabilities of occurrence of the terminal links in each alternative ($TL_{1,1}$ and $TL_{1,2}$ for Alternative 1, $TL_{2,1}$ and $TL_{2,2}$ for Alternative 2). p corresponds to the probability of reinforcement after each terminal link ($p_{1,1}$, $p_{1,2}$, $p_{2,1}$, and $p_{2,2}$ for $TL_{1,1}$, $TL_{1,2}$, $TL_{2,1}$, and $TL_{2,2}$, respectively); d is the terminal-link duration; Σ refers to the overall probability of reinforcement in each alternative.

The overall probability of reinforcement (Σ) associated with each alternative is the weighted average of the probabilities of reinforcement following each terminal link,

$$\Sigma_1 = r_1 \times p_{1,1} + (1-r_1) \times p_{1,2} \text{ and } \Sigma_2 = r_2 \times p_{2,1} + (1-r_2) \times p_{2,2}, \text{ for alternatives 1 and 2, respectively.}$$

The weights are the probabilities of entering each terminal link. In the standard task (left panel of Figure 3), alternative 1 has $p_{1,1}=1$ and $p_{1,2}=0$, which means that the color in the terminal link correlates perfectly with the outcome. Moreover, $r_1=.2$ and therefore $\Sigma_1 = .2$. Alternative 2 has $p_{2,1}=p_{2,2}=.5$, which means that color and outcome are uncorrelated and $\Sigma_2 = .5$.

To explain suboptimal choice, Roper and Zentall (1999) advanced a contrast-like account. When a pigeon chooses the IL_1 , its food expectation equals the overall probability of reinforcement associated with that alternative, $\Sigma_1 = .2$. However, when it encounters $TL_{1,1}$, its food expectation increases to $p_{1,1}=1$, a positive contrast of magnitude $+0.8$. When the pigeon

chooses the IL_2 , it expects food with probability $\sum_2 = .5$ and this expectation does not change with the terminal link stimuli because each is associated with the same food probability, $p_{2,1}=p_{2,2}=0.5$; there is no contrast in this alternative. If the value of an alternative depends on this form of local contrast, the IL_1 will have more value than the IL_2 ; hence the suboptimal choice. Critically, in Roper and Zentall's (1999) account, the negative contrast of magnitude -0.2 occasioned by entering the S- ($TL_{1,2}$), plays no role. As we shall see, subsequent accounts have assumed also that the S- does not affect preference.

Another attempt to explain suboptimal choice uses Mazur's (1996, 1995) hyperbolic discounting model. The model states that the current value of a delayed reward is an inverse function of its delay. Moreover, the value of the delayed reward accrues to the stimulus that signals the delay ($TL_{1,1}$). However, if a stimulus signals a delay never followed by reward ($TL_{1,2}$), the stimulus does not accrue value and it is ignored; functionally, it is as if the stimulus does not exist. Therefore, the value of IL_1 stimulus equals, not the average of the values of the $TL_{1,1}$ and $TL_{1,2}$, but the value of the $TL_{1,1}$ stimulus alone, that is, the value of food delayed 10 s. In contrast, because the $TL_{2,1}$ and $TL_{2,2}$ signal delays occasionally followed by food, the value of the IL_2 is the average of the values of the $TL_{2,1}$ and $TL_{2,2}$ stimuli, that is, the value of food delayed 20 s on average (it takes on average two trials to obtain food in the presence of the stimuli); hence the preference for the IL_1 , the suboptimal alternative.

Another model developed to explain suboptimal behavior is the Reinforcement Rate Model (RRM). Aiming at ultimate rather than proximate explanations, the model is based on optimal foraging considerations and suggests that animals behave as if following an information-seeking strategy (Fortes et al., 2016; Fortes et al., 2017; Vasconcelos et al., 2015). In their original proposal, $TL_{1,1}$ is a valuable stimulus because it provides information about the availability of food; animals attend to it. However, the $TL_{1,2}$ is a signal for no food, and because in the natural environment animals use this information to initiate another foraging bout, they never learn to attend and devalue it. As a consequence, in laboratory contexts animals simply ignore the $TL_{1,2}$. Given this ancillary assumption, RRM makes the same prediction as Mazur's model.

Cunningham and Shahan (2018) took also an informational approach. However, instead of information about whether food will be delivered, their model stresses information about when food will be delivered. Animals choose the suboptimal alternative because the

signals predicting food provide relevant information about when and where food will be available. The authors construe the animals as searching for and valuing information about the timing of food deliveries. Their proposal also explicitly accepts the notion that a stimulus that does not signal food has no impact on choice (see also Blanchard, 1975).

These accounts share three features. First, they all assume that animals ignore the stimulus never followed by food, the $TL_{1,2}(S-)$, in the typical pigeon task (Laude, Beckmann, Daniels, & Zentall, 2014; McDevitt et al., 2016; Stagner, Laude, & Zentall, 2011; Vasconcelos et al., 2015; Zentall, 2015). What each account understands by ‘ignoring’, though, is not explicitly defined (but see Fortes et al., 2017). Second, as a consequence of the first assumption, in all accounts the $TL_{1,1}(S+)$ has an uneven impact on choice. And third, they deal well with tasks in which the terminal links of alternative 1 correlate perfectly with the trial outcome ($p_{1,1} \rightarrow \text{food}$, $p_{1,2} \rightarrow \text{no food}$; for reviews, see Cunningham & Shahan, 2018; McDevitt et al., 2016; Vasconcelos et al., 2018; Zentall, 2015). When the correlation is not 1 and, in particular, when $p_{1,2} > 0$ and, in consequence, the $TL_{1,2}$ is occasionally reinforced, the models either do not apply – this condition falls outside their domain (e.g., Cunningham & Shahan, 2018; Roper & Zentall, 1999), or need further ancillary assumptions (RRM).

The diversity of accounts of suboptimal choice may reflect not only obvious differences in theoretical perspective or research tradition, but also the complexity of the task and our ignorance of how its multiple features affect preference. These features include the six different stimuli (2 ILs, 4 TLs), six probabilities (2 r ’s and 4 p ’s), the reinforcement schedule operating during the initial links, and the terminal link delays. In what follows, we propose the Delta-Sigma hypothesis, abbreviated $\Delta\text{-}\Sigma$, a hypothesis that suggests two higher-order variables as determiners of the value of an alternative and a decision rule mapping value to preference. This hypothesis intends to encompass a wider set of conditions, particularly those in which the p ’s in the suboptimal alternative differ from 0 and 1. The hypothesis combines the six independent probability variables into two. After presenting the hypothesis, we report two experiments that tested it.

In the $\Delta\text{-}\Sigma$ hypothesis, Δ refers to the difference between the two reinforcement probabilities associated with the terminal links of an alternative, $\Delta_1 = p_{1,1} - p_{1,2}$, and $\Delta_2 = p_{2,1} - p_{2,2}$. In the standard task (see Figure 3, left panel), $\Delta_1=1$ and $\Delta_2=0$; hence, $\Delta_1 > \Delta_2$. The hypothesis assumes that, all else equal, the value of an alternative varies directly with Δ ; greater differences between the two terminal link probabilities mean greater value of the

alternative. As for Σ , it refers to the overall reinforcement probability of an alternative. In the standard task, for alternative 1, $\Sigma_1=.2$, and for alternative 2, $\Sigma_2=.5$; hence, $\Sigma_1 < \Sigma_2$. The hypothesis assumes that, all else equal, value increases directly with Σ ; greater overall reinforcement probability means greater value of the alternative. From these two assumptions it follows that the value of an alternative increases directly with the higher-order variables Δ and Σ : In symbols, if V denotes the value of an alternative, and f denotes a function of Δ and Σ , then $V=f(\Delta, \Sigma)$ with positive partial derivatives, $\partial f/\partial\Delta>0$ and $\partial f/\partial\Sigma>0$.

Nevertheless, what is measured during the task is preference for one alternative over the other. As a consequence, it is necessary to relate value to preference. We assume that preference for alternative 1, P_1 , equals $g(V_1, V_2)$, for some function g that is always between 0 and 1, increases with the ratio V_1/V_2 , and, bias aside, equals 0.5 when $V_1=V_2$. Putting the two functions together yields $P_1= g(V_1, V_2) = g(f(\Delta_1, \Sigma_1), f(\Delta_2, \Sigma_2))$. Suppose $\Sigma_1=\Sigma_2$; then preference (P_1) should increase with Δ_1 and decrease with Δ_2 . Suppose $\Delta_1=\Delta_2$; then preference for P_1 should increase with Σ_1 and decrease with Σ_2 .

To illustrate with a specific example, suppose that, for alternative i (with $i=1, 2$), we let

$$V_i = f(\Sigma_i, \Delta_i) \propto (\Sigma_i)^c * e^{\beta*\Delta_i}, \quad (1)$$

where \propto means ‘proportional to’, and $c>0$ and $\beta>0$ are scaling parameters. In addition, with no bias we can let

$$P_1 = g(V_1, V_2) \propto \frac{V_1}{V_1+V_2}, \quad (2)$$

perhaps the simplest choice rule. Substituting Equation 1 into Equation 2 and simplifying yields,

$$P_1 = \frac{1}{1 + \left(\frac{\Sigma_2}{\Sigma_1}\right)^c e^{-\beta(\Delta_1-\Delta_2)}}. \quad (3)$$

In this case, preference is a function that increases with the difference between the Δ 's, which we refer to as Δ_{diff} , and with the ratio between the Σ 's. Although we do not need specific forms for functions f and g in order to derive ordinal predictions from the Δ - Σ

hypothesis, we will use the preceding example to explain our reasoning and describe the experiments reported below.

The reinforcement probabilities in the terminal links (the p_i 's) play center stage in the Δ - Σ hypothesis because they determine the higher-order variables Δ_1 and Δ_2 and contribute to Σ_1 and Σ_2 . They also figure prominently in other accounts of suboptimal choice. Surprisingly, however, these probabilities have not been manipulated systematically to measure their effects on preference. There are two noteworthy exceptions, one that changed $p_{1,1}$, the reinforcement probability associated with S+, and the other that changed $p_{1,2}$, the reinforcement probability associated with S-. Zentall and Stagner (2011) reduced $p_{1,1}$ from 1 to .8 while maintaining constant the remaining probabilities ($p_{1,2}=0, p_{2,1}=p_{2,2}=.5$). Pigeons continued to prefer the suboptimal alternative. This result challenges Cunningham and Shahan's information model because the model predicts suboptimal preference only when the outcomes and the terminal link stimuli in Alternative 1 correlate perfectly (i.e. when $p_{1,1}=1$ and $p_{1,2}=0$). But the result is consistent with the Δ - Σ hypothesis because, in Equation 3, the difference $\Delta_1 - \Delta_2$ remains positive. To illustrate, with $c=1$ and $\beta=4$, Equation 1 yields $P_1=.96$ when $p_{1,1}=1$, and $P_1=.89$ when $p_{1,1}=0.8$, in both cases a strong preference for the suboptimal alternative.

Fortes et al. (2018) increased $p_{1,2}$ from 0 to .375 while maintaining constant the remaining probabilities ($p_{1,1}=1, p_{2,1}=p_{2,2}=.5$). Increasing $p_{1,2}$ retarded acquisition, but it did not change steady state preference for the suboptimal alternative. This result challenges most models mentioned above (but latest version of RRM; Fortes et al., 2018) because they do not explain how each terminal link is combined to determine the alternative's value. For instance, it is not clear the role of a signal with probability $p_{1,1}=1$ and a signal of probability $p_{1,2} \neq 0$ to determine the subjective value of Alternative 1 (e.g. Fortes et al., 2018). However, the result is consistent with the Δ - Σ hypothesis because, again, the difference $\Delta_1 - \Delta_2$ remained positive: With $c=1$ and $\beta=4$, Equation 1 yields $P_1=.96$ when $p_{1,2}=0$, and $P_1=.92$ when $p_{1,2}=0.375$, in both cases a preference for the suboptimal alternative.

On the other hand, studies that varied the overall probability of reinforcement Σ revealed either a small or no effect on preference. Some of those studies varied Σ in the optimal alternative (Smith, Bailey, Chow, Beckmann, & Zentall, 2016; Zentall, Laude, Stagner, & Smith, 2015), whereas others manipulated the ratio Σ_2/Σ_1 (Roper & Zentall, 1999;

Smith et al., 2016). To illustrate, in Roper and Zentall's (1999) study, it was always the case that, as in the standard task, $p_{1,1}=1$ and $p_{1,2}=0$, yielding $\Delta_1=1$ and $\Sigma_1=r_1$, and $p_{2,1}=p_{2,2}=p$ yielding $\Delta_2=0$ and $\Sigma_2=p$. By setting $p=r_1$ and varying r_1 across conditions (.125, .50, and .875), the authors ensured that $\Sigma_1=\Sigma_2$ while the overall reinforcement probability in the task varied. They found that, as the overall reinforcement probability increased from 0.125 to 0.50 to 0.875, preference for the suboptimal alternative remained significantly above chance and did not change systematically. These results agree with Equation 1: With $\Delta_1-\Delta_2=1$ and $\Sigma_2/\Sigma_1 = 1$, Equation 3 predicts $P_1=1/(1+e^{-\beta})$, a constant value (e.g., .98 with $\beta= 4$).

Stagner et al. (2012) studied the effect of Σ while maintaining constant the other task variables. Both alternatives had stimuli in the terminal link perfectly correlated with the presence and absence of food ($p_{1,1}=p_{2,1}= 1$ and $p_{1,2}=p_{2,2}= 0$), yielding $\Delta_1=\Delta_2= 1$ and $\Delta_{diff}= 0$, but one alternative had $\Sigma_1=r_1= .2$ whereas the other had $\Sigma_2=r_2= .5$. The authors found indifference between alternatives, a result at odds with Equation 3, which predicts a preference for the richer alternative (i.e., $P_1= .28$ with $c= 1$ and $\beta= 4$; indifference requires $c= 0$). Although Stagner et al.'s results suggest that pigeons may become insensitive to Σ when the terminal link stimuli are perfectly correlated with the outcomes (i.e., Δ 's= 1), the result needs to be replicated and extended to determine whether preference changes when larger Σ ratios accompany Δ s of 1, and when similar Σ ratios accompany less extreme Δ s.

In the present study, we tested the Δ - Σ hypothesis in two experiments. In Experiment 1, we examined the effect on preference of varying the probabilities of food in the terminal links (p 's), a set of variables that have not been manipulated systematically and yet figure prominently in models of suboptimal choice including the Δ - Σ hypothesis. We created different Δ values while maintaining all other parameters, in particular the Σ s, equal between and within alternatives. We predicted a preference for the alternative with the larger Δ . In Experiment 2, we examined the effect of the overall reinforcement probabilities, the Σ s, while maintaining constant all other parameters of the task, in particular the Δ s. We predicted a preference for the larger Σ , with a preference proportional to the ratio between them.

Experiment 1

According to the Δ - Σ hypothesis, the value of each alternative in the suboptimal task depends on two variables, the difference between the reinforcement probabilities, Δ , and the overall reinforcement probability, Σ . Experiment 1 examined the effect of the first variable, by manipulating the probabilities of reinforcement on each terminal link. Across conditions, we maintained $\Sigma_1 = \Sigma_2 = .5$ but varied the Δ s. We expected animals to show a preference for the greater Δ .

Method

Subjects.

Five pigeons (*Columba Livia*) with similar experience in studies of timing and reinforcement schedules, were maintained at about 80% of their free-feeding weights. They were housed individually in a temperature-controlled colony room with grit and water always available. The colony room had 13:11 h light/dark cycle (lights on at 8:00 a.m.). The pigeons were cared for in accordance with the animal care guidelines of the Directorate-General for Food and Veterinary, the Portuguese national authority for animal health, and the University of Minho. All experimental procedures were conducted in agreement with European (Directive 2010/63/EU) and Portuguese law (Ordinance 1005/92 of October 23), and were approved by the Directorate-General for Food and Veterinary (Authorization #024946).

Apparatus.

The experiment used three identical operant chambers from Med Associates®, 24 cm long, 30 cm wide, and 28.5 cm high. In each chamber, the response panel was equipped with three keys, 2.5 cm in diameter, centered on the panel 21 cm above the floor grid, and 6-cm apart center-to-center. In the same panel, a 5 x 6 cm hopper opening centered on the panel and 4 cm above the floor grid, provided access to grain when the hopper was raised and illuminated by a 1.1-W light. A 2.8-W houselight located 23 cm above the grid in the opposite panel provided general illumination. Each chamber was enclosed by an outer box equipped with a fan that circulated air and helped to mask external noises. The experiment was programmed, and the data were recorded, with the ABET II software (Lafayette Instruments).

Procedure.

Pre-Training. Pigeons were exposed to four sessions of pretraining, the first two with a Fixed Ratio (FR) 1, the third with a FR5, and the last one with a FR10 schedule. In each session, all six stimuli (two figures, circle and cross, and four colors, red, green, blue, and yellow) were presented on the side keys, eight times each, in random order, for a total of 48 trials. On each trial, the chosen stimulus remained displayed until the pigeon pecked at it and completed the schedule requirement at which time the key turned off and the feeder raised and illuminated for 3 s. A 10-s ITI with the houselight on followed. At the end of the pretraining all pigeons pecked all stimuli readily.

Experimental conditions. A session comprised 120 trials using a concurrent chain schedule. Forty trials were choice trials with both alternatives presented simultaneously; eighty trials were forced trials, with only one alternative presented. Choice and forced trials intermixed randomly.

On a choice trial, pigeons had to peck one of the two initial link keys (IL₁ or IL₂), each illuminated with a geometric figure. After a single peck (FR1), both keys turned off and the terminal link of the chosen alternative started. The chosen key illuminated with one of two possible colors (TL_{1,1} and TL_{1,2} for Alternative 1, and TL_{2,1} or TL_{2,2} for Alternative 2). The probability of occurrence of each terminal link remained the same ($r_1 = r_2 = .5$). Each terminal link lasted 10 s, after which the trial ended with or without food according to the scheduled probability (i.e., the p). A 10-s ITI with the houselight on separated trials. Forced trials followed the same structure, but only one initial link key was illuminated. After the pigeon pecked at it, the terminal link proceeded exactly as on free-choice trials. Figures and colors were counterbalanced across pigeons.

The experiment had six conditions, C1 to C6, each defined by the p 's associated with the terminal links. Table 1 shows the details. The p 's yielded three possible values for $\Delta 1$ and $\Delta 2$ (0, .5 and 1), and from these values three signed differences were formed ($\Delta \text{diff} = \Delta 1 - \Delta 2 = -0.5, 0, \text{ or } +0.5$). Each difference occurred in two conditions, one for each possible assignment of $\Delta 1$ and $\Delta 2$ to Alternatives 1 and 2. In all conditions, the overall probabilities of reinforcement, $\sum 1$ and $\sum 2$, equaled .5.

Table 1
Experimental conditions of Experiment 1.

Conditions	p _{1,1}	p _{1,2}	Δ_1	p _{2,1}	p _{2,2}	Δ_2	Δ_{diff}
C1	.5	.5	0	.5	.5	0	0
C2	.75	.25	.5	.5	.5	0	.5
C3	.5	.5	0	.75	.25	.5	-.5
C4	1	0	1	.75	.25	.5	.5
C5	.75	.25	.5	1	0	1	-.5
C6	1	0	1	1	0	1	0

Table 2
Order of experimental conditions for each pigeon in Experiment 1.

Pigeon/Order	1	2	3	4	5	6	7	8
161	0 vs 0	0 vs .5	.5 vs 0	.5 vs 1	1 vs .5	.5 vs 1	.5 vs 0	1 vs 1
	C1	C3	C2	C5	C4	C5	C2	C6
724	0 vs 0	1 vs .5	.5 vs 1	.5 vs 0	0 vs .5	.5 vs 0	.5 vs 1	1 vs 1
	C1	C4	C5	C2	C3	C2	C5	C6
918	0 vs 0	.5 vs 1	1 vs .5	0 vs .5	.5 vs 0	0 vs .5	1 vs .5	1 vs 1
	C1	C5	C4	C3	C2	C3	C4	C6
960	0 vs 0	.5 vs 0	0 vs .5	1 vs .5	.5 vs 1	1 vs .5	0 vs .5	1 vs 1
	C1	C2	C3	C4	C5	C4	C3	C6
G39	0 vs 0	1 vs .5	0 vs .5	.5 vs 0	.5 vs 1	1 vs .5	0 vs .5	1 vs 1
	C1	C4	C3	C2	C5	C4	C3	C6

Note. The experimental conditions are presented for each pigeon in the order they occurred. The first and last conditions were the same for all pigeons; the sixth and seventh conditions repeated either C3 and C4 or C2 and C5.

The Δ - Σ hypothesis predicts that preference should favor the alternative with the highest Δ . Thus, pigeons should favor Alternative 1 in conditions 2 and 4, Alternative 2 in

conditions 3 and 5, and be indifferent in conditions 1 and 6. Furthermore, because the hypothesis states that preference depends on Δ_{diff} (see Equation 1), we predict a similar preference in conditions 2 and 4 ($\Delta_1 - \Delta_2 = 0.5$), conditions 3 and 5 ($\Delta_1 - \Delta_2 = -0.5$), and conditions 1 and 6 ($\Delta_1 - \Delta_2 = 0.0$).

The experiment was divided into 8 phases that corresponded to the six conditions C1 to C6 plus the repetition of either conditions C2 and C3, or C4 and C5. Table 2 shows for each pigeon the experimental condition in each phase. The order of conditions was counterbalanced with two restrictions. First, all pigeons experienced C1 first and C6 last. And second, if Alternative 1 signaled a larger Δ in one condition, it had to signal the smaller Δ in the next condition. Each phase lasted for 16 sessions.

Results and Discussion

Figure 4 shows for each pigeon how preference for Alternative 1 varied with the difference $\Delta_1 - \Delta_2$. Each data point shows the preference averaged over the last four sessions of the corresponding condition. The two sets of symbols correspond to conditions in which at least one of the Δ 's equaled 1 (black circles) or 0 (gray circles); in some conditions with $\Delta_{\text{diff}} = -0.5$ or $\Delta_{\text{diff}} = +0.5$, a second black or gray circle shows a replication. The lines are the least-squares best-fitting regression lines. The bottom right panel shows the average data and linear fits.

The symbols show that preference for Alternative 1 increased with Δ_{diff} . They also show no systematic differences between the two sets of symbols. Statistical analyses corroborated these impressions. First, two paired t-tests comparing the results from the original and replicated conditions revealed no significant differences (for Δ 's of 0 vs .5, $t(4) = 0.101$, $p = .924$, 95% CI [-0.298, .265]; for Δ 's of 1 vs .5, $t(4) = 2.416$, $p = .074$, 95% CI [.020, .293]). Hence, for all subsequent analyses we averaged the results of the original and replicated conditions. Second, a 3x2 repeated-measures ANOVA, with Condition ($\Delta_{\text{diff}} = -0.5$, 0.0, and +0.5) and Δ value ($\Delta = 1$ and $\Delta = 0$) as factors, revealed a significant effect of Condition ($F[2, 8] = 11.832$, $p = .004$, $\eta^2 = 0.609$) but non-significant effects of Δ value ($F[1, 4] = 0.852$, $p = .408$, $\eta^2 = 0.033$) or of their interaction ($F[2, 8] = 2.497$, $p = .144$, $\eta^2 = 0.005$). And third, the slopes of the regression lines were all positive, with values ranging from 0.125 to 0.869.

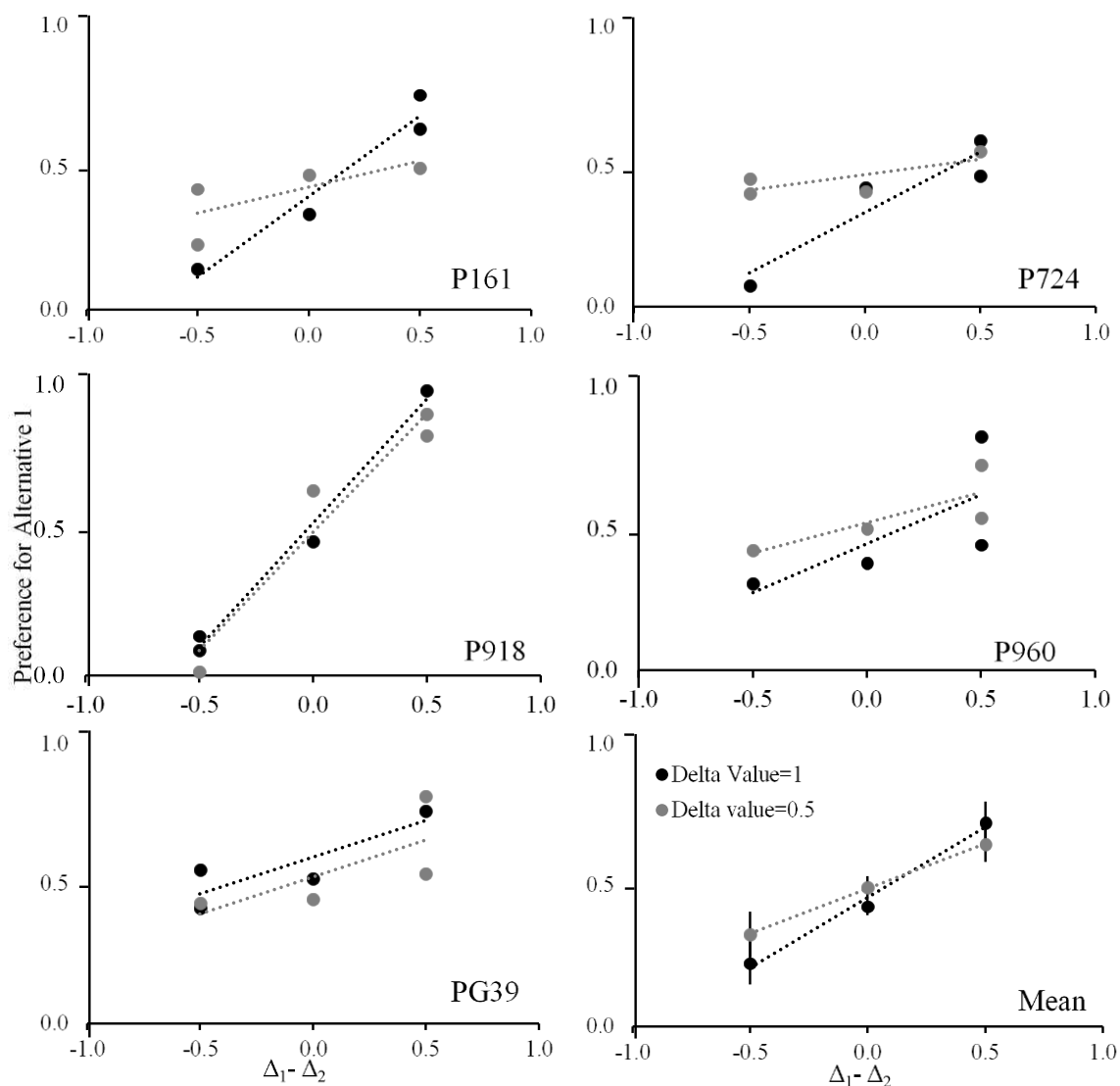


Figure 4. Average preference for Alternative 1 across Δ_{diffS} . The first five graphs show individual data and the last graph shows the average preference across pigeons. The black dots show preference when one (or two for the middle black dot) of the alternatives had stimuli perfectly correlated with the presence and absence of food. The grey dots show the other cases. Lines correspond to linear fittings of the data for both data sets. Error bars in the bottom right graph correspond to the standard error of the mean.

We also analyzed response latencies as an indirect measure of the value of each alternative. It has been suggested that latencies to peck the preferred alternative are shorter than the latencies to peck the non-preferred alternative (Kacelnik, Vasconcelos, Monteiro, & Aw, 2011). If higher values yield shorter latencies, and larger Δ s determine higher values, then we would expect latencies to vary negatively with Δ . Or, to put it differently, if pigeons prefer alternatives with higher Δ s, then they should be faster at pecking the preferred alternative. To check whether this was the case, we averaged the latencies on the forced trials of the last four sessions of each condition.

As for preference data, we averaged the data from the original and repeated conditions. Figure 5 shows the average latency to peck Alternatives 1 and 2 in each experimental condition. The left, middle, and right panels show the data for $\Delta_{\text{diff}} = -0.5$, 0.0, and +0.5, respectively. Within each panel, we further separate the conditions according to the specific Δ values used to obtain the same Δ_{diff} . In the four conditions with different Δ s (left and right panels), latencies were shorter for the higher Δ , more-preferred alternative in three cases (except Condition 1 vs .5, in which no difference between alternatives was observed).

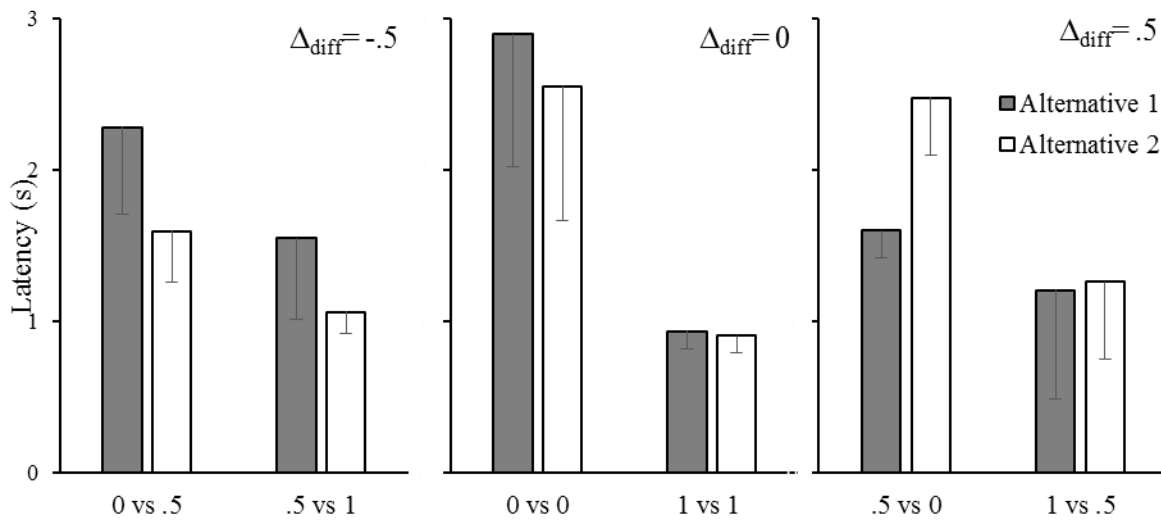


Figure 5. Average latencies to peck each alternative in Experiment 1. Latencies are shown by condition for each Δ_{diff} . Dark gray and white bars refer to latencies for Alternatives 1 and 2, respectively. The error bars correspond to standard error of the mean.

However, it is unclear whether latency depended on the Δ of the alternative, or on the difference between the two Δ s, a sort of context effect. To contrast these hypotheses, we performed multiple regressions on average latency to respond to a given alternative, using as predictors the Δ of that alternative and the Δ of the other alternative. The analysis revealed that (a) latency was not modulated by the context effect (the coefficient of the Δ of the other alternative never differed significantly from 0, all p 's > .05), but (b) latency correlated strongly with the Δ of the chosen alternative for three of the pigeons ($\beta_{P960} = -.783$, $p < .001$, $R^2 = .663$; $\beta_{P161} = -2.759$, $p < .001$, $R^2 = .52$; and $\beta_{PG39} = -1.645$, $p < .005$, $R^2 = .499$). The other two subjects did not yield coefficients significantly different from 0 ($\beta_{P724} = -.171$, $p = .495$, $R^2 = .034$; and $\beta_{P918} = 0.079$, $p = .77$, $R^2 = .006$). We conclude that when latencies varied, they tended to be shortest when $\Delta = 1$, longest when $\Delta = 0$, and intermediate when $\Delta = .5$.

To summarize, when we varied the difference between the reinforcement probabilities in the terminal links of each alternative, the two Δ s, while maintaining constant the overall reinforcement probabilities, the two Σ s, the pigeons preferred the alternative with a higher Δ . Importantly, the preference for the higher Δ was similar between conditions, suggesting that pigeons' preferences depended on the difference and not the specific probabilities that yielded the difference. However, the latency data suggested an effect of the individual Δ s – the subjects showed larger latencies for $\Delta=0$ when it was paired with $\Delta= .5$ than for $\Delta= .5$ when it was paired with $\Delta= 1$. That is, latencies differed even though in both cases $\Delta_{diff}= -0.5$. Similarly, when the Δ s were equal and $\Delta_{diff}= 0$, larger latencies were found for the condition with both Δ s= 0 than with both Δ s= 1. It seems that, in the suboptimal choice task, preference and latency may depend on different variables.

Overall, the results of Experiment 1 are consistent with the Δ - Σ hypothesis. According to the hypothesis, for constant Σ , value should increase with Δ ; hence, when $\Sigma_1= \Sigma_2$, but $\Delta_{diff}= \Delta_1-\Delta_2$ varies, pigeons should favor IL1 when $\Delta_{diff}>0$ and IL2 when $\Delta_{diff}<0$. More generally, preference should increase with Δ_{diff} , the observed finding. Our results also show that the sensitivity to the reinforcement probabilities is not limited to the extreme cases of $\Delta_1=1$ and $\Delta_2=0.0$ (i.e., $\Delta_{diff}=+1$); this sensitivity seems to extend to other values. Future studies should extend the range of probability pairs to determine whether, for example, preference remains constant regardless of whether a $\Delta= .5$ is achieved with $p_{1,1}= .75$ and $p_{1,2}= .25$ or with $p_{1,1}= .6$ and $p_{1,2}= .1$.

Experiment 2

According to the Δ - Σ hypothesis, the value of each alternative in the suboptimal task depends on two variables, the difference between the reinforcement probabilities, Δ , and the overall reinforcement probability, Σ . Experiment 1 examined the effect of the first variable and found results consistent with the hypothesis. Experiment 2 examined the effect of the second variable. Across conditions, we maintained $\Delta_1=\Delta_2 =1.0$ but varied the Σ s correlatively: As Σ_1 increased, Σ_2 decreased such that $\Sigma_1+\Sigma_2=1$. According to the hypothesis, pigeons should prefer the alternative with the higher Σ .

Method

Subjects.

Six pigeons (*Columba Livia*) with previous experience in timing and schedules of reinforcement procedures were used. The animals were kept in the same conditions as in Experiment 1.

Apparatus.

The experiment used the same chambers as in Experiment 1.

Procedure.

Pre-Training. The same pre-training of Experiment 1 was used.

Experimental Conditions. The general structure of the task was the same as in Experiment 1 (see Figure 3, left panel). Animals chose between alternatives 1 and 2, each leading to one of two 10-s terminal links. The probabilities of reinforcement in the terminal links of Alternative 1 were $p_{1,1} = 1$ and $p_{1,2} = 0$; the same was true for the Alternative 2, $p_{2,1} = 1$ and $p_{2,2} = 0$, yielding $\Delta_1 = \Delta_2 = 1$ and $\Delta_{diff} = 0$. To evaluate the effect of Σ , we manipulated the probabilities $\Sigma_1 = r_1$ and $\Sigma_2 = r_2$.

To maintain the richness of the task constant, we set $\Sigma_1 + \Sigma_2 = 1$, which is equivalent to $r_1 + r_2 = 1$. Hence, the probability of occurrence of TL_{1,1} (r_1) equaled the probability of occurrence of TL_{2,2} ($1 - r_2$). Similarly, the probability of occurrence of TL_{1,2} ($1 - r_1$) equaled the probability of occurrence of TL_{2,1} (r_2). Across the five conditions shown in Table 3, r_1 equaled .1, .3, .5, .7 and .9.

Table 3
Experimental conditions of Experiment 2.

Conditions	r_1	Σ_1	r_2	Σ_2
.5	.5	.5	.5	.5
.1	.1	.1	.9	.9
.3	.3	.3	.7	.7
.7	.7	.7	.3	.3
.9	.9	.9	.1	.1

Note. r_1 and r_2 are the probabilities of occurrence of each stimulus in the terminal link. Σ s correspond to the overall probability of reinforcement in each alternative.

As in Experiment 1, pigeons experienced every condition at least once, started and ended with Condition .5, and repeated Condition .1 or .9, and Condition .3 or .7 (See Table 4). Each condition lasted for 16 sessions. The alternative associated with the highest Σ changed between adjacent conditions. For instance, if a pigeon was in Condition .1, with $\Sigma_1=.1$ and $\Sigma_2=.9$, the next condition had to be Condition .7 or .9.

Table 4.

Order of experimental conditions for each pigeon in Experiment 2.

Pigeon/Order	1	2	3	4	5	6	7	8
458	.5 vs .5	.1 vs .9	.9 vs .1	.3 vs .7	.7 vs .3	.3 vs .7	.9 vs .1	.5 vs .5
967	.5 vs .5	.9 vs .1	.1 vs .9	.7 vs .3	.3 vs .7	.7 vs .3	.1 vs .9	.5 vs .5
157	.5 vs .5	.3 vs .7	.7 vs .3	.1 vs .9	.9 vs .1	.1 vs .9	.7 vs .3	.5 vs .5
449	.5 vs .5	.7 vs .3	.3 vs .7	.9 vs .1	.1 vs .9	.9 vs .1	.3 vs .7	.5 vs .5
286	.5 vs .5	.1 vs .9	.7 vs .3	.3 vs .7	.9 vs .1	.1 vs .9	.7 vs .3	.5 vs .5
444	.5 vs .5	.3 vs .7	.9 vs .1	.1 vs .9	.7 vs .3	.3 vs .7	.9 vs .1	.5 vs .5

Note. The experimental conditions (Σ_1 vs Σ_2) are presented in the order they occurred for each pigeon. Notice that the first and last conditions were fixed for all animals. Conditions 6 and 7 are repetitions of condition 2 or 3 and 4 or 5.

Results and Discussion

We analyzed the data from the last four sessions. Figure 6 shows how preference for Alternative 1 varied with Σ_1 for each pigeon. A second point in some conditions show the replications. The lines are the least-squares best-fitting regression lines. The bottom panel shows the average data and linear fit.

A preliminary analysis revealed that preference did not differ between a condition and its replication. Three paired-sample t-tests showed no significant effect when $\Sigma_1 = .5$ ($t(5) = .427$, $p = .687$, CI 95% [-0.151, 0.211]), when $\Sigma_1 = .9$ or $\Sigma_2 = .9$ ($t(5) = .460$, $p = .665$, CI 95% [-0.092, 0.132]), and when $\Sigma_1 = .7$ or $\Sigma_2 = .7$ ($t(5) = .719$, $p = .504$, CI 95% [-0.051, 0.091]). Hence, in subsequent analysis we averaged the results from the original and replicated conditions.

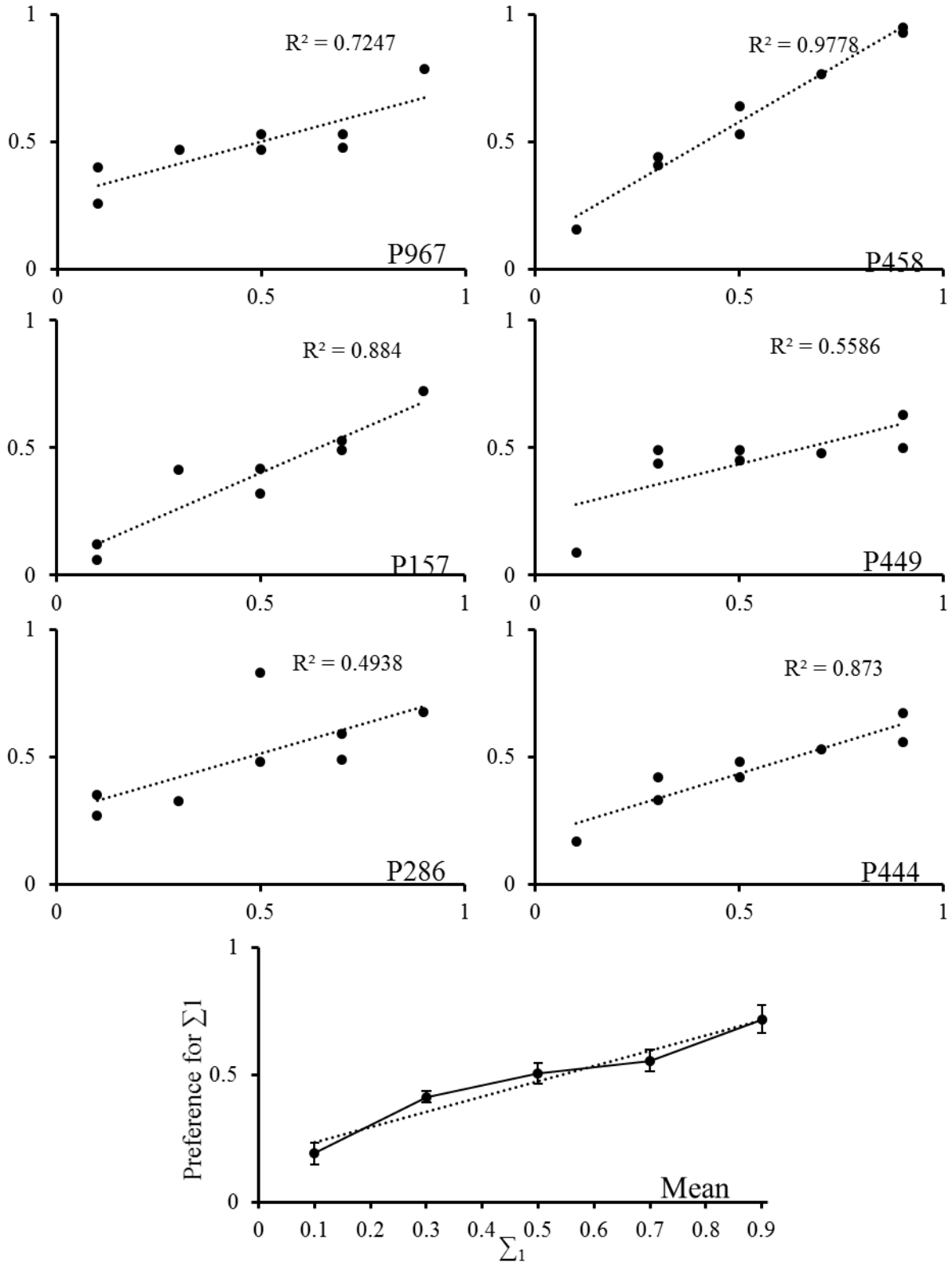


Figure 6. Average preference for Alternative 1 as a function of Σ_1 . The first six graphs show individual data and the bottom graph shows the average preference across pigeons. Lines correspond to the linear fitting of the data. Error bars in the bottom graph correspond to the standard error of the mean.

A repeated-measures ANOVA using Condition as the within-subject factor, revealed a significant effect of the factor ($F[4, 20] = 28.186, p < .001, \eta^2 = 0.772$), corroborating what was observed in Figure 6. Finally, the slopes of the regression lines for all pigeons were positive, with values ranging from 0.400 to 0.928. A t-test showed that the average slope was significantly different from 0 ($t(5) = 6.758, p < .001, 95\%CI [.352, .784], d = 2.759$), confirming the trend observed for all pigeons.

As in Experiment 1, we analyzed the latencies on the forced trials. If higher values yield shorter latencies, and a higher Σ corresponds to a higher value, then we would expect latencies to vary negatively with Σ . That is, if pigeons prefer the alternative with higher Σ , then they should be faster at choosing that alternative. As for the choice data, we used the average data of each condition. Figure 7 shows the averaged latencies to peck at Alternatives 1 and 2 on forced trials averaged over the last four sessions of each condition. The graph shows a small but consistent difference between the alternatives, with the alternative with the highest Σ yielding the shortest average latency. This result suggests that the latencies on the forced trials were in line with the preference on the choice trials: The alternative with the higher Σ yielded shorter latencies and more choices.

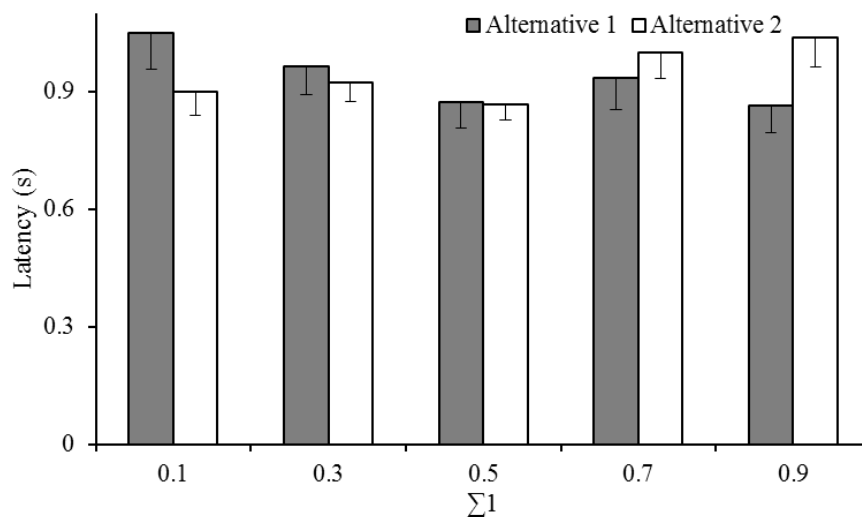


Figure 7. Average latencies to peck each alternative across conditions in Experiment 2. Gray and white bars refer to latencies for Alternative 1 and 2, respectively. The error bars correspond to the standard error of the mean.

The results in Figure 7 suggest that the difference between the two sets of latencies depended on the ratio between the two Σ s. To evaluate this hypothesis, we performed linear regressions for each pigeon taking the difference between latencies (i.e. latency of IL1–latency of IL2) as the dependent variable and the ratio Σ_1/Σ_2 as the independent variable. The

slope of the regression lines were all negative and significantly different from 0 ($\beta_{P458} = -.199$, $R^2 = .925$; $\beta_{P967} = -0.321$, $R^2 = .931$; $\beta_{P157} = -0.327$, $R^2 = .947$, $\beta_{P449} = 0.151$, $R^2 = .776$; $\beta_{P286} = 0.185$, $R^2 = .944$; $\beta_{P444} = -0.117$, $R^2 = .790$; in all cases, $p < .05$).

In conclusion, consistent with the Δ - Σ hypothesis, when Σ changed while Δ remained constant, preference increased with the ratio of the Σ 's. When the Σ 's differed, no pigeon showed exclusive preference for the richest alternative, the optimal choice. Whether this result holds also when the Δ 's are not 1 and 0 (and $\Delta_{diff} \neq 0$) remains to be seen. Moreover, preference correlated negatively with latency – longer latencies occurred on the less preferred alternative. The difference in latency between the preferred and non-preferred alternative was proportional to the Σ_1/Σ_2 ratio. To further understand the effect of the Σ s on preference, future experiments should evaluate a similar range of Σ 's with different Δ 's.

General Discussion

Suboptimal choice is hard to explain. Part of the difficulty stems from the complexity of the task, from its large number of features including various stimuli, delays, and probabilities. Previous research has examined the role of, among other variables, the terminal link delays (Cunningham & Shahan, 2019; Fortes et al., 2016; Kendall, 1974; Spetch, Belke, Barnet, Dunn, & Pierce, 1990; Zentall & Stagner, 2011), the degree of contiguity between choice and the terminal link stimulus onset (McDevitt, Spetch, & Dunn, 1997; Vasconcelos et al., 2015), and the response requirement during the initial links (Fantino et al., 1979; Kendall, 1974; Zentall, Andrews, & Case, 2017). However, most models of suboptimal choice rely on the predictability of the signals in the terminal link of the suboptimal alternative and the difference in the overall rate of reinforcement between alternatives (Cunningham & Shahan, 2018; Fortes et al., 2018; Stagner & Zentall, 2010). Some recent experiments have manipulated the overall rate of reinforcement (Roper & Zentall, 1999; Smith et al., 2016; Stagner et al., 2012), but none manipulated systematically the probabilities of reinforcement in the terminal links.

The aim of our study was to advance the Δ - Σ hypothesis and to test it in two experiments. The hypothesis states that the value of an alternative increases with the difference between the reinforcement probabilities associated with the two terminal links of the alternative, a within-alternative contrast or Delta (Δ); and it increases also with the overall

probability of reinforcement of the alternative, Sigma (Σ); see Equation 1. Moreover, preference between two alternatives depends on the relative value of the alternatives (Equation 2). In Experiment 1, we manipulated the difference between the probabilities of reinforcement of each terminal link. We expected that, with all other variables equal, the pigeons would prefer the alternative with the greater Δ . In Experiment 2, we manipulated the overall probability of reinforcement of each alternative. We expected that, with all other variables equal, the pigeons would prefer the alternative with the greater Σ .

Notice that, in Experiment 1, the Σ ratio always equaled 1 and therefore Equation 3 simplifies to

$$P_1 = \frac{1}{1 + e^{-\beta(\Delta_1 - \Delta_2)}}; \quad (4)$$

a one parameter equation. Similarly, in Experiment 2, because $\Delta_1 = \Delta_2 = 1$ across conditions, the model simplifies to

$$P_1 = \frac{1}{1 + \left(\frac{\Sigma_2}{\Sigma_1}\right)^c} \quad (5)$$

also a one parameter equation.

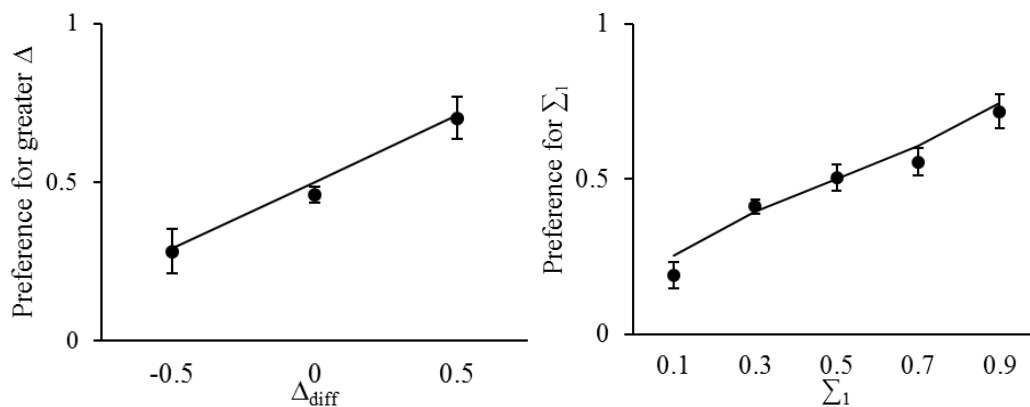


Figure 8. Average fit of the Delta-Sigma hypothesis (solid line) for Experiments 1 (left panel) and 2 (right panel). The predicted preference was estimated for each animal across conditions, then the fits were averaged. Error bars are the standard error of the mean.

Figure 8 shows how well the Δ - Σ hypothesis fits the data from Experiments 1 and 2. In Experiment 1, the fits of Equation 4 to the individual data yielded β parameter values ranging from 1.1 to 5.2 and the average of the individual fits was close to the average of the individual data (left panel). In Experiment 2, the fits of Equation 5 yielded c parameter values

ranging from .39 to .91. Again, the average of the fits was close to the average of the data (right panel). In summary, the Δ - Σ hypothesis captures well variations in preference engendered by manipulations of the within-alternative contrast (Δ) and by manipulations of the overall probability of reinforcement of each alternative (Σ).

Most, if not all, other models of suboptimal choice face difficulties in dealing with the findings of at least one of our experiments. Cunningham and Shahan's (2018) model, for instance, does predict the findings of our second experiment: provided that the terminal stimuli are perfectly correlated with the presence and absence of food, animals ought to prefer the alternative reinforced most often. Yet, when the perfect correlation between terminal stimuli and outcome is degraded (i.e., when both $p_{i,1}$ and $p_{i,2} > 0$) as in our first experiment, the model is unable to generate predictions. Without ancillary assumptions, the model cannot combine the temporal information conveyed by two terminal stimuli into a single metric.

The latest version of the RRM (see Fortes et al., 2018) faces similar, albeit opposite, difficulties. The recently proposed engagement function allows the model to deal with imperfect correlations of the sort implemented in our first experiment and consistently predict a preference for the alternative with the greater Δ . Note that is not Δ per se that drives preference, but the exact values of the terminal link probabilities. Still, the model is unable to cope with the findings of our second experiment. Because in this experiment the two terminal links of each alternative are perfectly correlated with food (S+) and its absence (S-) and because the model proposes that animals do not attend to the S-, animals ought to be functionally choosing between alternatives that always end with food reward. Consequently, but contrary to our findings, the model predicts indifference in all conditions of Experiment 2. Mazur's hyperbolic discounting model (Mazur, 1996, 1995), despite originating from a different research tradition, is mathematically very similar to the RRM and thus fails and succeeds in exactly the same conditions as the RRM.

Zentall's contrast-like account (e.g., Roper & Zentall, 1999), on the other hand, makes correct predictions for the first but not for the second experiment. Assuming that (1) only positive contrast ($p-\Sigma > 0$) affects choice and that (2) animals prefer options with greater contrast, this account correctly predicts the findings of our first experiment, but it also predicts a preference for the lower Σ in the second experiment. To see why, note that in the second experiment the terminal links of both alternatives are perfectly correlated with the

presence and absence of food. Thus, the S+ in each alternative yields positive contrast ($1.0-\Sigma$) and the S- yields negative contrast ($0.0-\Sigma$). Negative contrast aside, this means that positive contrast and Σ (with $\Sigma = r$, in this case) are inversely related and consequently any preference for higher contrast entails a preference for lower Σ s. Table 5 summarizes the failures and successes of the models aforementioned.

Table 5.
Failures and successes of sub-optimal choice models.

Model	Manipulation	
	Δ_{diff}	Σ_2/Σ_1
Cunningham and Shahan's model	n.a.	✓
RRM	✓	---
Mazur's hyperbolic model	✓	---
Zentall's contrast account	✓	---

Note. “✓” correct prediction; “---” incorrect prediction; “n.a.” manipulation outside model's domain.

To further test the scope of the Δ - Σ hypothesis, we generated predictions for 28 studies/conditions published in 13 different papers. Table 6 details 20 of these studies, all with two terminal links per alternative; Table 7 details the remaining eight wherein one of the alternatives had a single terminal link. For each study, we calculated Δ_{diff} , and Σ_2/Σ_1 , assuming $\Delta = 0$ whenever a single terminal link was used (cf. Table 7). The bulk of studies involved alternatives that differed in both Δ (i.e., $\Delta_{diff} \neq 0$) and Σ (i.e., $\Sigma_2/\Sigma_1 > 1$); four had different Δ s but equal Σ s and two had different Σ s but equal Δ s. Across studies, Δ_{diff} has not been manipulated systematically; as Table 6 and Table 7 show, Δ_{diff} was set to 1.0 in 22 of the 28 studies (range: 0.0 - 1.0). Σ_2/Σ_1 , on the other hand, was somewhat more variable across studies ranging from 1 to 20. Joint manipulations of Δ_{diff} and Σ_2/Σ_1 are absent.

To generate predictions when the two alternatives differed in only Δ or Σ , we searched for the best fitting parameter, β or c , respectively. When the two alternatives differed in both Δ and Σ , we set c equal to the average value estimated in our second experiment and searched for the best fitting β . We opted for this constrained search because, when animals are sensitive to Δ_{diff} , Σ_2/Σ_1 or both, the model can accurately predict the

observed preference via an infinite number of parameter pairs. In fact, the predictions matched the observed preferences in all cases.

Overall, when the alternatives involved two terminal links each (Table 6), β varied between 1.02 and 26.6. Compared with our first experiment, the range is much wider but this appears to be a case of overfitting. Indeed, we were able to predict the exact reported preference for each study (see last column). The outliers in β were necessary to predict Fortes et al.'s (2018) results only, but had we set its maximum value to 6.0, the maximum prediction error would have been of about 2%. In the single study listed in Table 6 where we searched for the best fitting c (cf. Stagner et al., 2012), it equaled .03. This is one order of magnitude below the minimum in our second experiment. Yet, Stagner et al.'s results suggest that animals are insensitive to Σ when the terminal-link stimuli are perfectly correlated with the outcomes, which contradicts the findings of our second experiment.

Despite early reports suggesting contradictory findings when one of the alternatives involves a single terminal link (e.g., Fantino et al., 1979; Kendall, 1974), recent research indicates that the procedure does yield consistent preference patterns (Smith et al., 2016; Smith & Zentall, 2016). As shown in Table 7, the Δ - Σ model captures the observed preferences findings well, both when Δ and Σ differed across alternatives (i.e., $\Delta_{\text{diff}} \neq 0$ and $\Sigma_2/\Sigma_1 > 1$) and when only one of them differed (Table 7). β varied between 0.36 and 2.27 and c equaled 1.26 in the only occasion we had to estimate this parameter.

More important than the ability of the Δ - Σ hypothesis to predict preference in a variety of manipulations, is the obvious conclusion from these 28 studies (cf. Tables 6 and 7) that the parametric exploration has been quite limited. Note, for instance, that typically $\Delta_1 = 1$ and $\Delta_2 = 0$, which limits not only our conclusions about the effect of Δ_{diff} itself but also confines our interpretations of the effects of Σ ratio to situations where $\Delta_{\text{diff}} = 1$. Perhaps this is simply due to the complexity of the task. The procedure involves six different stimuli (2 ILs, 4 TLs) and six probabilities (2 r 's and 4 p 's), making parametric studies extremely laborious. The Δ - Σ hypothesis clarifies the nomenclature and reduces the complexity to two higher-order variables driving preference: Δ , the difference between the two reinforcement probabilities associated with the terminal links of an alternative, and Σ , the overall reinforcement probability of an alternative. For the moment, our findings indicate that the value of an alternative increases with both Δ and Σ , as predicted.

Table 6.
Suboptimal choice studies using the 4-stimuli procedure.

Reference	Exp. reference	r_1	$p_{1,1}$	$1-r_1$	$p_{1,2}$	ΔI	r_2	$p_{2,1}$	$1-r_2$	$p_{2,2}$	$\Delta 2$	Δ_{diff}	β	Σ_1	Σ_2	Σ_2/Σ_1	c	Preference and Prediction
Roper & Zentall (1999)	Experiment 1	.5, .125 and .875	1	.5, .875 and .125	0	1	0.5	.5, .125 and .875	0.5	.5, .125 and .875	0	1	2.51	.5, .125 and .875	.5, .125 and .875	1	--	0.925
Gipson, Jerome, Miller & Zentall (2009)	Experiment 2	0.5	1	0.5	0	1	0.5	0.75	0.5	0.75	0	1	1.02	0.5	0.75	1.5	0.5	0.693
Stagner & Zentall (2010)	Training	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	3.91	0.2	0.5	2.5	0.5	0.969
Stagner, Laude & Zentall (2011)	Average groups	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	3.57	0.2	0.5	2.5	0.5	0.957
Zentall and Stagner (2011)	Exp.1 Acquisition	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	1.93	0.2	0.5	2.5	0.5	0.812
	Exp.2 Acquisition	0.2	0.8	0.8	0	0.8	0.2	0.5	0.8	0.5	0	0.8	3.65	0.16	0.4	2.5	0.5	0.912
Stagner, Laude & Zentall (2012)	Condition both 1	0.2	1	0.8	0	1	0.5	1	0.5	0	1	0	--	0.2	0.5	2.5	0.0	0.494
	Condition disc. One	0.2	1	0.8	0	1	0.5	0.5	0.5	0.5	0	1	2.06	0.2	0.5	2.5	0.5	0.831
Vasconcelos, Monteiro & Kacelnik (2015)	Exp 1. Standard group	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	7.38	0.2	0.5	2.5	0.5	0.999

Zentall, Laude, Stagner & Smith (2015)	G50/75	0.5	1	0.5	0	1	0.5	0.75	0.5	0.75	0	1	1.31	0.5	0.75	1.5	0.5	0.750*
	G25/75	0.25	1	0.75	0	1	0.5	0.75	0.5	0.75	0	1	1.66	0.25	0.75	3	0.5	0.750*
Fortes, Vasconcelos & Machado (2016)	Condition 1 (Exp1)	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	5.07	0.2	0.5	2.5	0.5	0.990
	Condition 2 (Exp1)	0.1	1	0.9	0	1	0.2	0.5	0.8	0.5	0	1	3.58	0.1	0.5	5	0.5	0.940
	Condition 3 (Exp1)	0.05	1	0.95	0	1	0.2	0.5	0.8	0.5	0	1	2.84	0.05	0.5	10	0.5	0.840
	Condition 4 (Exp1)	0.025	1	0.975	0	1	0.2	0.5	0.8	0.5	0	1	1.70	0.025	0.5	20	0.5	0.540
	Exp 2 (First condition)	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	3.41	0.2	0.5	2.5	0.5	0.950
Fortes, Machado & Vasconcelos (20017)	Exp. 1 and 2	0.2	1	0.8	0	1	0.2	0.5	0.8	0.5	0	1	25.60	0.2	0.5	2.5	0.5	1.00*
Fortes, Pinto, Machado & Vasconcelos (2018)	Condition 1 (Exp1)	0.2	1	0.8	0.0313	0.97	0.2	0.5	0.8	0.5	0	0.968	25.60	0.225	0.5	2.2	0.5	1.00*
	Condition 2 (Exp1)	0.2	1	0.8	0.1875	0.81	0.2	0.5	0.8	0.5	0	0.812	25.60	0.35	0.5	1.429	0.5	1.00*
	Condition 3 (Exp1)	0.2	1	0.8	0.375	0.63	0.2	0.5	0.8	0.5	0	0.62	26.60	0.5	0.5	1	--	1.00*

*Data preference estimated from graphs.

Note. All the parameters, the calculated Δ_{diff} , \sum_2/\sum_1 , and the used values of sensitivity (β and c) are presented for each study, experiment, condition or group as indicated in the second column. The last column presents only one value for predicted and observed preference for Alternative 1, because the model was fitted to the exact proportion of choice.

Table 7.

Suboptimal choice studies using the 3-stimuli procedure.

Reference	Exp. reference	r_1	$p_{1,1}$	$1-r_1$	$p_{1,2}$	$\Delta 1$	r_2	$p_{2,1}$	$\Delta 2$	Δ_{diff}	β	$\Sigma 1$	$\Sigma 2$	$\Sigma 2/\Sigma 1$	c	Preference /Prediction
Smith, Bailey, Chow, Beckmann & Zentall (2016)	Exp 1 cond 1	0.5	1	0.5	0	1	1	0.5	0	1	2.04	0.5	0.5	1	--	0.885
	cond 2	0.3	1	0.75	0	1	1	0.5	0	1	2.27	0.25	0.5	2	0.5	0.872
	cond 3	0.1	1	0.88	0	1	1	0.5	0	1	2.08	0.13	0.5	4	0.5	0.798
	Exp 2 cond 2	0.5	1	0.5	0	1	1	0.8	0	1	1.54	0.5	0.8	1.5	0.5	0.792
	cond 3	0.5	1	0.5	0	1	1	1	0	1	1.04	0.5	1	2	0.5	0.665
Smith & Zentall (2016)	Exp 1	0.5	1	0.5	0	1	1	1	0	1	0.36	0.5	1	2	0.5	0.5
	Exp 2	0.5	0.5	0.5	0.5	0	1	1	0	0	--	0.5	1	2	1.3	0.294
	Exp 3	0.5	1	0.5	0	1	1	0.5	0	1	0.61	0.5	0.5	1	--	0.647
Case & Zentall (2018)	Control	0.5	1	0.5	0	1	1	1	0	1	0.97	0.5	1	2	0.5	0.65

Note. The table shows for each study, experiment, and condition, the values of the procedural variables (r_1 , $p_{1,1}$, $p_{1,2}$ and $p_{2,1}$), higher-order variables (Δ_{diff} , Σ_2 and Σ_1), and model parameters (β and c). The last column presents only one value for predicted and observed preference for Alternative 1, because the model was fitted to the exact proportion of choice.

CHAPTER III²
TESTING THE Δ - Σ MODEL: SAME
DELTA WITH DIFFERENT
PROBABILITIES OF REINFORCEMENT

²I would like to thank Carlos Pinto for help collecting data for Experiment 4.

Abstract

In a concurrent-chain procedure, pigeons make a choice between two alternatives. One has an initial-link stimulus that is followed by reliable signals that food will be delivered or not after a delay, while the other alternative has an initial-link stimulus that is followed by unreliable signals of food. Pigeons prefer the former even when it yields a lower overall probability of food. Recently, a new model has been developed to explain the effect. The Delta-Sigma model proposes that the preference depends on two variables: the difference of reinforcement probabilities associated with the signals (Delta, Δ), a sort of contrast effect; and the overall probability of reinforcement (Sigma, Σ). An interesting prediction of the model is that animals should prefer greater delta values independently of the specific probabilities of reinforcement of the signals in each alternative. In two experiments, we tested this prediction comparing a $\Delta = .5$ against a $\Delta = 0$, with the former built with different pairs of probabilities of reinforcement; and a constant Sigma around .5 for both alternatives. The results are partially supported by the Delta-Sigma model.

Introduction

In the last few decades, researchers have made great efforts to understand a phenomenon in which, given a choice between two alternatives, animals strongly prefer the alternative that offers the least reinforcement. This apparently suboptimal choice occurs when the leaner alternative informs whether and when reinforcement will be delivered, but the richer alternative does not. The suboptimal choice effect has been observed with pigeons (Fortes et al., 2016; Gipson, Alessandri, Miller, & Zentall, 2009), starlings (Vasconcelos et al., 2015), humans (Lalli, Mauro, & Mace, 2000; Molet et al., 2012), and sometimes with rats (Cunningham & Shahan, 2019; Ojeda et al., 2018; but Trujano et al., 2016; Trujano & Orduña, 2015). With pigeons, the effect is reliable across a broad variety of conditions (for review, see McDevitt et al., 2016; Zentall, 2016a).

However, we still do not know the necessary and sufficient conditions to produce the phenomenon, partly because the procedure includes a relatively large number of parameters. The left panel of Figure 9 shows the typical procedure, a concurrent-chain schedule with two initial links, concurrently available during a choice phase. After choice, one of the four terminal links (two per alternatives) is available. The links are signaled by distinctive stimuli projected on the two keys, typically geometric figures such as a “plus” and a “circle” to signal the initial links, and light hues such as Red, Green, Blue and Yellow to signal the terminal links. To illustrate, suppose the pigeon chooses Alternative 1, the “plus”, during the initial-link phase. Then, the “circle” is turned off and the “plus” is replaced by either a Red light (with probability $r_1 = .2$) or a Green light (with probability $1-r_1 = .8$). After 10 s, the terminal-link delay, the pigeon always receives food when the light is Red (i.e., food occurs with probability $p_{1,1} = 1$), and it never receives food when the light is Green ($p_{1,2} = 0$). Conversely, if the pigeon chooses Alternative 2, the “circle” key, during the initial-link phase, then the “plus” key turns off and the “circle” is replaced by either a Yellow light (with probability $r_2 = .2$) or a Blue light (with probability $1-r_2 = .8$). After the 10-s delay, food is delivered in half of the trials (i.e., with probabilities $p_{2,1} = p_{2,2} = .5$). Given the six probabilities (two r 's and four p 's), Alternative 1 is 2.5 times poorer than Alternative 2 because it yields food on 20% of the trials, whereas Alternative 2 yields food on 50% of the trials. And yet, pigeons prefer Alternative 1, the suboptimal alternative.

The usual account of this instance of “suboptimal choice” points to the fact that the terminal-link stimuli correlate perfectly with the trial outcome in Alternative 1 (Red is an S+

and Green is an S-), but not in Alternative 2 (Blue and Yellow are both S+/-). That is, the terminal-link stimuli are informative in Alternative 1 but not in Alternative 2, and suboptimal choice may express the pigeons' strong preference for information concerning food (McDevitt et al., 2016; Zentall, 2016a).

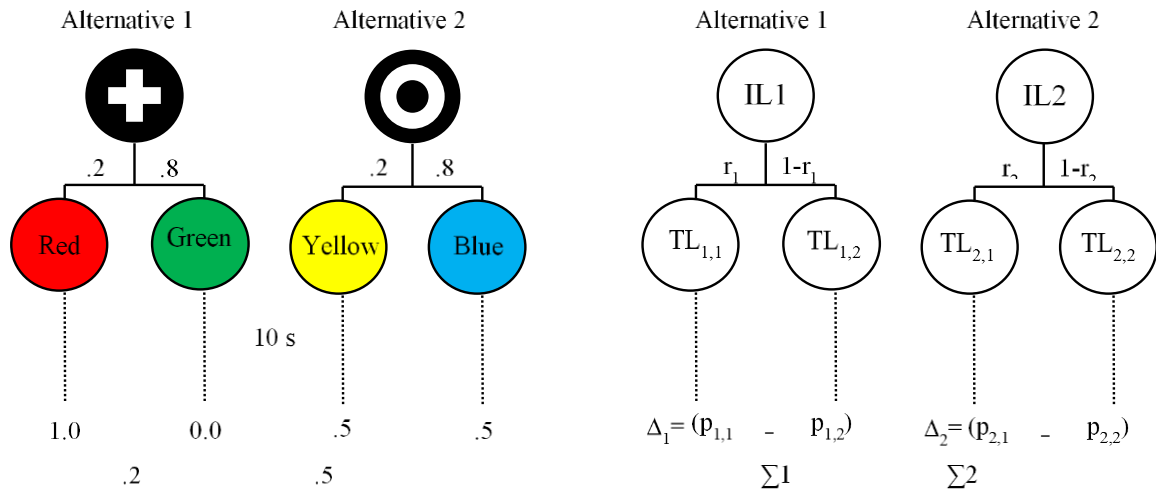


Figure 9. Suboptimal choice task. Left panel: typical procedure, in which animals choose with one peck in the initial links, a cross for the suboptimal (i.e. Alternative 1) or a circle for the optimal (i.e. Alternative 2) alternative. After the choice, animals see with probability .2 or .8, one out of two terminal links signaled by color, each lasting for 10s. The terminal link Red is always reinforced, Green is never reinforced; Yellow and Blue are reinforced half of the time. The overall probability of reinforcement was .2 and .5 for suboptimal and optimal alternative, respectively. Right panel: General task. Animals are confronted to a concurrent-chain schedule with two links. IL₁ and IL₂ correspond to the Initial links where the choice is made. r₁ and r₂ are the probabilities of occurrence of the terminal links in each alternative. p_{1,1} and p_{1,2} are the probabilities of reinforcement for Alternative 1, p_{2,1} and p_{2,2} for Alternative 2. The difference of those correspond to Δ_1 and Δ_2 . Finally, Σ_1 and Σ_2 refer to the overall probability of reinforcement of each alternative.

We have previously advanced an account of suboptimal choice with a simple quantitative model (Chapter II). According to the Delta-Sigma model, the value of an alternative, V , increases directly with two higher-order variables associated with that alternative, the difference between the reinforcement probabilities in the terminal links of the alternative (Delta, Δ), and the overall probability of reinforcement obtained by choosing the alternative (Sigma, Σ). The model suggested a specific way to integrate these two variables,

$$V = \Sigma^c \times e^{-\beta\Delta},$$

where c and β are two model parameters ($c > 0$ and $\beta > 0$). The Δ s for alternatives 1 and 2 equal, respectively, $\Delta_1 = (p_{1,1} - p_{1,2})$ and $\Delta_2 = (p_{2,1} - p_{2,2})$; and the Σ s equal, correspondingly,

$\Sigma_1=r_1 \times p_{1,1}+(1-r_1) \times p_{1,2}$ and $\Sigma_2=r_2 \times p_{2,1}+(1-r_2) \times p_{2,2}$. Hence, the values of alternatives 1 and 2 equal, respectively,

$$V_1 = (\Sigma_1)^c \times e^{-\beta \Delta_1}, \text{ and}$$

$$V_2 = (\Sigma_2)^c \times e^{-\beta \Delta_2}.$$

To predict the probability of choosing Alternative 1, P_1 , from the values of the two alternatives, the model suggested a simple ratio rule,

$$P_1 = \frac{V_1}{V_1+V_2},$$

which yields,

$$P_1 = \frac{(\Sigma_1)^c \times e^{-\beta \Delta_1}}{(\Sigma_1)^c \times e^{-\beta \Delta_1} + (\Sigma_2)^c \times e^{-\beta \Delta_2}}$$

$$= \frac{1}{1 + \left(\frac{\Sigma_2}{\Sigma_1}\right)^c e^{-\beta(\Delta_1 - \Delta_2)}}. \quad (1)$$

Notice that, if the choice is between two alternatives with different Σ s but equal Δ s (i.e., $\Sigma_1 \neq \Sigma_2, \Delta_1 = \Delta_2$), then Equation 1 simplifies to

$$P_1 = \frac{1}{1 + \left(\frac{\Sigma_2}{\Sigma_1}\right)^c} \quad (2)$$

with a single parameter. Similarly, if the animal chooses between two alternatives with different Δ s but equal Σ s, then Equation 1 simplifies to

$$P_1 = \frac{1}{1 + e^{-\beta(\Delta_1 - \Delta_2)}}. \quad (3)$$

also with a single parameter.

Results from two experiments showed that animals were sensitive to both higher-order variables, Delta and Sigma (see Chapter II). In Experiment 1, the Δ s varied across conditions while the Σ s remained equal. Consistent with the model (Equation 3), the pigeons preferred the option with the greater Δ . In Experiment 2, the Σ s varied across conditions while the Δ s remained equal. The pigeons preferred the alternative with the greater Σ , a result

also consistent with the model (Equation 2). The authors also showed that the 2-parameter Delta-Sigma hypothesis could explain most of the data in the suboptimal choice literature.

An interesting prediction of the Delta-Sigma model is that the value of an alternative depends on the difference between the two reinforcement probabilities associated with the terminal links (i.e., the Δ), and not on the two specific probabilities used to obtain the difference. In other words, provided Σ does not change, pigeons should be equally as sensitive to probabilities of .75 and .25 as to probabilities of .9 and .4 because both yield a Δ of .5. One could test this prediction in one of two ways. First, if the two pairs of probabilities (.75/.25) and (.9/.4) are pitted against each other as Alternatives 1 and 2, with r_1 and r_2 adjusted so that the Σ s remain equal, the model predicts that V_1 will equal V_2 and, therefore, that $P_1=.5$; i.e., the pigeon should be indifferent. Second, if the two pairs are used in two separate conditions as Alternative 1, and each pair pitted against the same Alternative 2, then, provided Σ_1 remains constant, the model predicts the same degree of preference for Alternative 1 in the two conditions.

Our previous results suggest consistency with the model (Chapter II), but no previous study was designed to test the foregoing prediction systematically. On the other hand, other models predict different results depending on the specific pairs of probabilities of reinforcement used in the terminal links. From previous example, if a Δ of .5 with probabilities of .75/.25 is compared with a Δ of 0 (.5/.5), some models predict a preference for the alternative with the pairs .75/.25 over .5/.5 (for instance, see Fortes et al., 2018) as the Delta Sigma hypothesis does; while others make no predictions (for instance, see Cunningham & Shahan, 2018). More interesting, most models assume that the pigeon ignores a stimulus never followed by food (Green, in Alternative 1; see left panel of Figure 9). In consequence, a probability of 0 does not contribute to the decision process. According to the models, when Green is never reinforced, the pigeon behaves as if Alternative 1 comprised a single terminal link, the Red key that is always followed by food. In consequence, when a $\Delta = .5$ (with the pairs .5/0) is compared with a $\Delta = 0$ (both probabilities equal to .5), the models predict indifference between the alternatives (see Cunningham & Shahan, 2018; Mazur, 1995; Vasconcelos et al., 2015); a prediction that differs from the Delta-Sigma hypothesis.

The objective of the present study is to continue to evaluate the Delta-Sigma model by testing the prediction that equal Deltas obtained from different pairs of probabilities of reinforcement have the same value. It followed the second strategy mentioned above: Across

experimental conditions (see right panel Figure 9), we varied the pair of probabilities ($p_{1,1}$; $p_{1,2}$) associated with Alternative 1 but in such a way that Δ_1 always equaled .5, while in Alternative 2 we maintained the probability pair used in the standard task (i.e., $p_{2,1}=p_{2,2}=.5$) such that Δ_2 always equaled 0. Moreover, by adjusting r_1 and r_2 , Σ_1 remained equal or approximately equal to Σ_2 in all conditions. Because in each condition $\Delta_1 > \Delta_2$ and $\Sigma_1 \approx \Sigma_2$, the Delta-Sigma model predicts a preference for Alternative 1 ($P_1 > .5$), but the same degree of preference regardless of the values of $p_{1,1}$ and $p_{1,2}$.

The study comprised two experiments. In Experiment 3, a Δ_1 of .5 was obtained by using the “extreme” probabilities of 1 and 0 included in the standard task. Specifically, across three conditions, the pairs of probabilities were 1/.5, .75/.25 and .5/0, the first using the extreme probability of 1, and the third the extreme probability of 0. In Experiment 4, the same Δ_1 value of .5 was obtained using two other pairs of probabilities close to the two extremes, .9/.4 and .6/.1. We contrast the results of both experiments against the Delta-Sigma model as well as other models of suboptimal choice: The Temporal-information model (Cunningham & Shahan, 2018), the RRM (Fortes et al., 2018; Vasconcelos et al., 2015), the HDM (Mazur, 1996; 1995) and the Zentall’s contrast-like account (Gipson et al., 2009; Roper & Zentall, 1999). Thus, the different predictions of the models give an interesting opportunity to evaluate their relevance in the suboptimal choice task.

Experiment 3

To evaluate the Delta-Sigma hypothesis, we varied across three conditions the pair of reward probabilities of the terminal links of Alternative 1 while maintaining constant the pair of probabilities of reinforcement in the terminal links of Alternative 2. The pair in Alternative 1 (i.e., $p_{1,1}$, $p_{1,2}$) was 1/.5, .75/.25, or .5/0, in each case yielding $\Delta_1 = .5$. In Alternative 2, the pair ($p_{2,1}$, $p_{2,2}$) was always .5/.5, yielding $\Delta_2 = 0$. Moreover, r_1 and r_2 were adjusted such that, in each condition and throughout the study, $\Sigma_1 \approx \Sigma_2$. According to the Delta-Sigma hypothesis, the pigeons should always prefer Alternative 1, and the degree of preference should remain constant across conditions.

Method

Subjects.

Nine pigeons (*Columba Livia*) with previous experience in procedures involving timing and different schedules of reinforcement, were used and maintained at 80% of their free-feeding weight. One of the pigeons become ill after a few sessions of training and was removed from the experiment. All animals were housed individually with permanent access to water and grit, in a 13:11 light:dark cycle in a temperature-controlled room (23° C). The pigeons were cared in accordance with the animal care guidelines of the Directorate-General for Food and Veterinary, the Portuguese national authority for animal health, and the University of Minho. All experimental procedures were conducted in agreement with European (Directive 2010/63/EU) and Portuguese law (Ordinance 1005/92 of October 23), and were approved by the Directorate-General for Food and Veterinary (Authorization #024946).

Apparatus.

We used three Med Associates® operant conditioning chambers each enclosed by an outer box equipped with a fan that circulated air and helped to mask extraneous noises. Each chamber measured 28.5 x 24 x 30 cm (height x length x width). The response panel contained three keys with LCD stimulus display, each 2.5 cm in diameter and 21 cm above the floor grid. The keys were centered on the panel, 6-cm apart center-to-center. The center key was not used during the experiment. The lateral keys displayed all 6 stimuli, the four solid colors (Red, Green, Blue and Yellow) and the two figures (Circle and Cross). Below the center key, a 5 x 6 cm (height x width) opening, 4 cm above the floor, gave the pigeon access to grain when the hopper was raised and illuminated by a 1.1-W light. The rear wall included a 2.8-W houselight, 23 cm above the floor. The experiment was programmed and the data recorded in ABET II software (Lafayette Instruments) with a temporal resolution of about 10 ms.

Procedure

Pretraining. In four sessions, pigeons were exposed to all the 6 stimuli used in the experiment, the two geometric figures and the four colors. In each session, the six stimuli were presented in the two side keys four times each, for a total of 48 trials. Trial order was random. The first two sessions used a FR 1 schedule – a single peck at the illuminated key turned it off and delivered the reinforcer – the feeder was raised and illuminated for 2 s. After the reinforcer, the houselight was turned on for a 10-s Inter-Trial Interval (ITI), and then a

new trial began, or the session ended. The FR parameter increased to 5 during the third session, and to 10 during the fourth session.

General procedure. Pigeons were exposed to the concurrent-chains procedure displayed in the right panel of Figure 9. During the initial link, the cross and the circle were randomly presented in the right and left keys. After a single peck at one of the keys (FR 1), both keys were turned off, and then one of the two possible terminal link stimuli appeared in the chosen key. The terminal-link stimulus always remained on for 10 s (FT 10 s). After this interval, the feeder was either raised for 2 s or not raised, depending on the scheduled reinforcement probabilities (see below). Then the houselight was turned on and the 10-s ITI began.

Sessions comprised 120 trials, from which 40 were choice trials (Alternative 1 and 2 presented simultaneously, as described above), and 80 forced trials (Alternative 1 or Alternative 2, but not both, presented in the initial link). The assignment of the two geometric figures as initial-links and the four hues as terminal-links were counterbalanced across pigeons with the restriction that R was always paired with G, and B always paired with Y. Table 8 shows the assignment of stimuli for each pigeon; only seven out of the eight possible permutations are presented, given that the last one was assigned to the pigeon that was removed from the experiment.

Table 8
Stimulus assignments for Experiment 3.

Pigeon	Alternative 1		Alternative 2	
	Plus		Circle	
	TL _{1,1}	TL _{1,2}	TL _{2,1}	TL _{2,2}
P730	R	G	B	Y
P053	Y	B	R	G
P435	R	G	B	Y
P795	Y	B	G	R
P917	R	G	Y	B
P588	G	R	Y	B
P816	B	Y	G	R
PG45	B	Y	R	G

Note. The letters R, G, Y and B refer to red, green, yellow and blue, respectively.

The pigeons experienced the three experimental conditions in pseudorandom order. Each condition lasted for 16 sessions. Table 9 shows the order of experimental conditions for each pigeon. The probabilities of reinforcement of each terminal link (p's) and the probabilities of occurrence (r's) are described.

Table 9
Order of experimental conditions in Experiment 3.

	Alternative	Probabilities	P730	P053	P435	P795	P917	P588	P816	PG45
Condition 1	1	r ₁	.5	.5	.9	.1	.1	.5	.9	.1
		p _{1,1}	.75	.75	.5	.5	.5	.75	.5	1
		p _{1,2}	.25	.25	.5	.5	.5	.25	0	.5
	2	r ₂	.5	.5	.9	.1	.1	.5	.9	.1
		p _{2,1}	.5	.5	.5	1	1	.5	.5	.5
		p _{2,2}	.5	.5	0	.5	.5	.5	.5	.5
Condition 2	1	r ₁	.9	.1	.1	.5	.5	.9	.1	.9
		p _{1,1}	.5	.5	1	.75	.75	.5	.5	.5
		p _{1,2}	.5	.5	.5	.25	.25	.5	.5	.5
	2	r ₂	.9	.1	.1	.5	.5	.9	.1	.9
		p _{2,1}	.5	1	.5	.5	.5	.5	1	.5
		p _{2,2}	0	.5	.5	.5	.5	0	.5	0
Condition 3	1	r ₁	.1	.9	.5	.9	.9	.1	.5	.5
		p _{1,1}	1	.5	.5	.5	.5	1	.75	.75
		p _{1,2}	.5	0	.5	.5	.5	.5	.25	.25
	2	r ₂	.1	.9	.5	.9	.9	.1	.5	.5
		p _{2,1}	.5	.5	.75	.5	.5	.5	.5	.5
		p _{2,2}	.5	.5	.25	0	0	.5	.5	.5

Note. r₁, r₂ correspond to the probability of occurrence of p_{1,1} and p_{2,1}; the probability of occurrence of p_{1,2} and p_{2,2} correspond to 1-r₁ and 1-r₂, respectively. p's are the probabilities of reinforcement of each terminal link.

To simplify the explanation of the conditions, we will take as an example pigeon P730. In condition 1, P730 experienced a Δ = .5 from the pair .75/.25 against the constant Δ = 0 (pair .5/.5). From the 40 forced trials of Alternative 1, 20 of them were followed by the

terminal link $TL_{1,1}$ ($r_1 = .5$) and the other 20 by $TL_{1,2}$ ($1-r_1 = .5$); 15 of the $TL_{1,1}$ trials were programmed to be baited ($p_{1,1} = .75$), while 5 of the $TL_{1,2}$ presentations were programmed to be baited ($p_{1,2} = .25$). From the 40 forced trials of Alternative 2, 20 were followed by $TL_{2,1}$ and 20 by $TL_{2,2}$ ($r_2 = .5$, $1-r_2 = .5$). Ten trials of each terminal link in this alternative were programmed to be baited ($p_{2,1} = p_{2,2} = .5$). Notice that r_1 was always equal to r_2 , to assure pigeons faced the terminal links with the same frequency between alternatives in all conditions. In condition 2 of P730, $\Delta = .5$ with the pair $.5/0$ was assigned to Alternative 2. As $r_1 = r_2 = .9$, P730 entered the $TL_{1,1}$ and $TL_{2,1}$ on 36 trials and $TL_{1,2}$ and $TL_{2,2}$ on 4 trials. From those, $TL_{1,1}$ and $TL_{2,1}$ were programmed to be reinforced on 18 of them ($p_{1,1} = p_{2,1} = .5$). $TL_{1,2}$ and $TL_{2,2}$ were programmed to be reinforced on 2 ($p_{1,2} = .5$) and 0 ($p_{2,2} = 0$) of the 4 trials, respectively. Finally, in condition 3, the $\Delta = .5$ with the pair $1/.5$ was assigned again to Alternative 1; P730 experienced $TL_{1,2}$ and $TL_{2,2}$ on 36 trials and $TL_{1,1}$ and $TL_{2,1}$ on four trials ($r_1 = r_2 = .1$). $TL_{1,1}$ was programmed to always be baited ($p_{1,1} = 1$), while $TL_{2,1}$, $TL_{1,2}$, and $TL_{2,2}$ were programmed to be baited on half of them (two out of four trials for the first terminal link, 18 out of 36 for the other two).

As described in the previous paragraph, the predicted preference for Pigeon P730 was Alternative 1 in condition 1, then Alternative 2 in condition 2, and finally Alternative 1 again in condition 3. The Δ of $.5$ was allocated to different alternatives between conditions for all subjects. The overall rate of reinforcement (Σs) was $.5$ for the alternative with $\Delta = 0$, but varied across conditions in the $\Delta = .5$; Σs was equal to $.55$ for the pair $1/.5$, $.5$ for the pair $.75/.25$ and $.45$ for the pair $.5/0$.

Results and discussion

Preference was calculated for the alternative containing the greater Δ value. Figure 10 shows the average preference for the alternative with $\Delta = .5$ of the last four sessions for all conditions. In the condition with the pair $1/.5$, most pigeons showed a clear preference for $\Delta = .5$ ($M = .73$, $SD = .14$). With the pair $.75/.25$, most pigeons had a preference for the alternative containing these probabilities of reinforcement ($M = .71$, $SD = .14$). Finally, when pigeons experienced the pair $.5/0$, preference for the $\Delta = .5$ was slightly below $.5$ ($M = .40$, $SD = .12$).

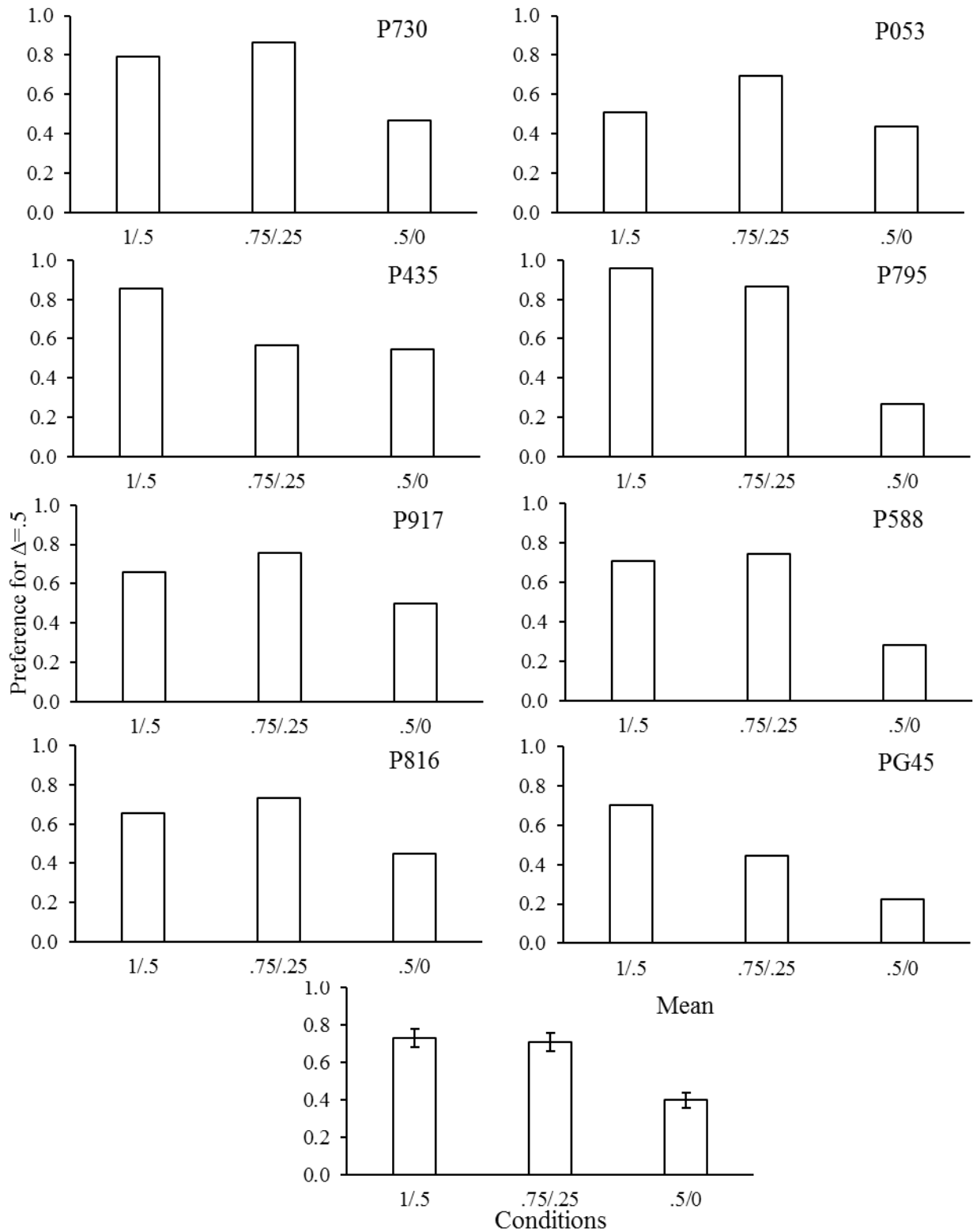


Figure 10. Choice data Experiment 3. The average preference for a $\Delta = .5$ on the last four sessions for each condition is presented by pigeon. The bottom panel shows the average preference for each condition. Error bars show SEM.

A repeated-measures ANOVA using Condition as a within-subject factor, showed a significant effect of Condition ($F[2, 14] = 16.724, p < .001, \eta^2 = .59$). Pairwise-comparisons using Bonferroni corrections showed no difference between the pairs 1/.5 and .75/.25 ($p = 1$), but a significant difference between the pairs 1/.5 and .5/0 ($p = .006$), and between .75/.25 and .5/0 ($p = .004$). The question whether preference with the pair .5/0 was evidence of indifference or a preference for the $\Delta = 0$ is theoretically relevant because some explanations of the suboptimal choice effect assume animals ignore signals for the absence of food, in which case pigeons would see two identical alternatives, thus indifference is expected. A one sample t-test found that the preference for the pair .5/0 was significantly below .5 ($t(7) = -2.426, p = .046, CI\ 95\% [-0.204, -0.300], d = -.858$). In sum, as shown in Figure 10, all pigeons preferred less the $\Delta = .5$ when the pair was .5/0 compared with the other two pairs. There was no systematic difference in preference between the pair 1/.5 and the pair .75/.25.

Latencies have shown to correlate with preference, being a sensitive metric of the value assigned to each alternative (Kacelnik et al., 2011; Shapiro, Siller, & Kacelnik, 2008). In general, shorter latencies are found in the preferred alternative. Latencies in this procedure were defined as the time from the start of a trial until the first peck on the initial link was made. Since the amount of data in the choice trial depended on the preference (e.g. preferences closer to 1 would give almost no data for the non-preferred alternative), we compared the median latency on the forced trials in the last four sessions for each condition.

Figure 11 shows latency for the three experimental conditions, averaged over the last four sessions for the $\Delta = .5$ (black dots) and $\Delta = 0$ (white dots). Visual inspection suggests that the shorter latency corresponds to the preferred alternative. Latencies for the pairs 1/.5 and .75/.25 were shorter in the alternative with $\Delta = .5$ than with $\Delta = 0$; however, the opposite was observed with the pair .5/0. A repeated-measures ANOVA, using Condition and Alternative as within-subject factors, confirmed no significant effect of Condition ($F[2, 14] = 0.995, p = .394, \eta p^2 = 0.026$), but a significant effect of Alternative ($F[1, 7] = 15.864, p = .005, \eta p^2 = 0.083$). The interaction Condition x Alternative was marginally significant ($F[2, 14] = 2.960, p = .085, \eta p^2 = 0.103$). To explore this interaction, a paired-sample t-test was performed for each condition. The analysis showed that the preferred alternative had significantly shorter latencies for the condition with the pairs 1/.5 ($t(7) = -2.614, p = .035, 95\% CI [-2.751, -0.138], d = -0.924$) and .75/.25 ($t(7) = -2.632, p = .034, 95\% CI [-1.711, -0.092], d = -0.931$), but no difference was found between alternatives for condition with the pair .5/0 ($t(7) = 0.723, p = .493, 95\% CI [-0.792, 1.489], d = 0.256$). The results of this

analysis suggest that, when comparing alternatives, their differences in latency are related to their levels of preference. Indeed, the $\Delta = .5$ with the pairs 1/.5 and .75/.25 was strongly preferred over the $\Delta = 0$ and a clear difference in latency between alternatives was found, while the preference for Condition .5/0 is approximately 10% below indifference with a non-significant difference in latency.

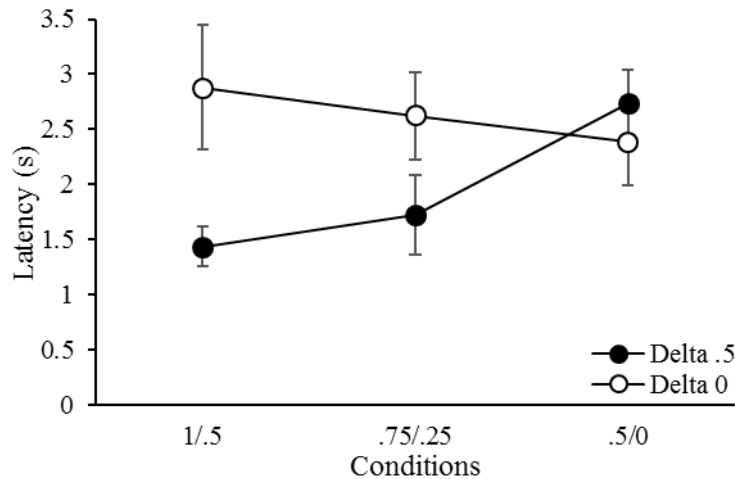


Figure 11. Averaged median latency data on forced trials of Experiment 3. Black dots show latency for the $\Delta = .5$, white dots show latencies for alternatives with $\Delta = 0$. Error bars correspond to the standard error of the mean.

Overall, Experiment 3 showed that the specific probabilities of reinforcement in the terminal links play a role in the value of the alternatives. Even though we found a preference for the greater Delta with the pairs .75/.25 and 1/.5, the condition with the probabilities of .5/0 revealed that the absence of reinforcement could be a special case. The result challenges our proposal, the Delta-Sigma hypothesis. It could indicate that the difference in probabilities of reinforcement in the terminal links is not the main variable responsible for the preference, or it may suggest that the difference in probability of reinforcement follows a dichotomy: There is a preference for the greater difference when that alternative contains a signal with high probability of reinforcement, but it is rejected when the alternative contains a signal for the absence of reinforcement. This idea supports other models of suboptimal choice which assume that a probability of reinforcement of 0 is ignored. However, those models predict indifference between the alternatives, a result that was not found in this experiment.

Some experiments showed that extreme values, such as probabilities of 1 or 0, are special cases, which animals value more and show more sensitivity towards them (Ludvig, Madan, McMillan, Xu, & Spetch, 2018). The results add evidence that in the pair .5/0 the

probability of 0 is a special case. Indeed, Pisklak, Madan, Ludvig and Spetch (2019) stated that pigeons choose an alternative to avoid the signal for no reward. However, the probabilities of 1/.5 did not yield different preference from the pair .75/.25, suggesting that 1, another extreme value, is not a special condition. Perhaps .5, the point that maximizes the absence of information, is aversive for the animals, and overshadows the effect of a probability of reinforcement of 1. In accordance with this idea, a series of experiments with humans found that, when given a choice between getting information about a negative outcome or no information at all, participants preferred the former (Lieberman et al., 1997).

Experiment 4

Experiment 4 aimed to extend the results found in Experiment 3. Using the same design of Experiment 3 we evaluated the effect of the extreme values (0 and 1) by using values close to but different (.6/.1 and .9/.4) than the extremes. When compared with an Alternative with $\Delta = 0$, the Delta-Sigma hypothesis predicts a preference for both $\Delta = .5$.

Method

Subjects and apparatus.

The same eight pigeons (*Columba livia*) from Experiment 3 were used in Experiment 4. Pigeons were maintained in the same conditions. The experimental chambers were the same as Experiment 3.

Procedure.

Pretraining. The same pretraining from Experiment 3 was used.

General procedure. We used the same task from Experiment 3. Stimuli were pseudorandomly assigned avoiding the same mapping from the previous experiment to reduce carry-over effects (See Table 10). As in the previous experiment, all conditions lasted for 16 sessions, each session had the same number of trials of each type, $r_1 = r_2$ across conditions, and the alternative allocated with the greater delta changed between conditions.

Table 10
Stimulus assignments for Experiment 4.

Pigeon	Alternative 1		Alternative 2	
	Plus		Circle	
	TL _{1,1}	TL _{1,2}	TL _{2,1}	TL _{2,2}
P730	Y	B	G	R
P053	G	R	B	Y
P435	Y	B	G	R
P795	R	G	B	Y
P917	B	Y	G	R
P588	B	Y	R	G
P816	R	G	Y	B
PG45	G	R	Y	B

Note. The letters R, G, Y and B refer to red, green, yellow and blue, respectively.

All pigeons started with Condition 1, in which they experienced the traditional suboptimal choice task (See Figure 9, left panel). Five pigeons acquired a suboptimal preference within 16 sessions. Nevertheless, the other three pigeons did not acquire a suboptimal preference after the training was extended for another eight sessions (total of 24). As those pigeons showed a strong bias for one key location, they were excluded from the next two conditions.

Table 11 shows the order of the experimental conditions for the pigeons that completed the experiment. Conditions 2 and 3 were counterbalanced across pigeons. As in the previous experiment, the number of times the pigeons experienced each terminal link was fixed (i.e. on 32 trials when r 's equaled .8; on 8 trials when r 's equaled .2). However, the number of reinforced trials after each terminal link for the pairs .9/.4 and .6/.1 was not an integer; therefore, the programmed reinforced trials were sampled with replacement. In consequence, the obtained probability of reinforcement at the end of each session was not always the same.

Table 11
Order of experimental conditions in Experiment 4.

Alternatives		Probabilities	P730	P053	P435	P795	P917
Condition 1	1	r_1	.2	.2	.2	.2	.2
		$p_{1,1}$.5	.5	1	1	1
		$p_{1,2}$.5	.5	0	0	0
	2	r_2	.2	.2	.2	.2	.2
		$p_{2,1}$	1	1	.5	.5	.5
		$p_{2,2}$	0	0	.5	.5	.5
Condition 2	1	r_1	.2	.8	.8	.2	.8
		$p_{1,1}$.9	.6	.5	.5	.5
		$p_{1,2}$.4	.1	.5	.5	.5
	2	r_2	.2	.8	.8	.2	.8
		$p_{2,1}$.5	.5	.6	.9	.6
		$p_{2,2}$.5	.5	.1	.4	.1
Condition 3	1	r_1	.8	.2	.2	.8	.2
		$p_{1,1}$.5	.5	.9	.6	.9
		$p_{1,2}$.5	.5	.4	.1	.4
	2	r_2	.8	.2	.2	.8	.2
		$p_{2,1}$.6	.9	.5	.5	.5
		$p_{2,2}$.1	.4	.5	.5	.5

Note. r_1 , r_2 correspond to the probability of occurrence of $p_{1,1}$ and $p_{2,1}$; the probability of occurrence of $p_{1,2}$ and $p_{2,2}$ correspond to $1-r_1$ and $1-r_2$, respectively. p 's are the probabilities of reinforcement of each terminal link.

Results and discussion

First, all the subjects were exposed again to the typical suboptimal task, with the pair of probabilities of 1/0 in one alternative ($\Delta=1$), and probabilities of .5 for the other alternative ($\Delta=0$). The pigeons who developed a suboptimal preference within the 16 sessions of training got an average preference in the last 4 sessions of .78 (SD = .09). The three pigeons that did not show a suboptimal preference had an average preference of .51 by the end of training ($M_{P816} = .55$, $M_{P588} = .58$, $M_{PG45} = .42$). This performance was unexpected, considering that these pigeons were showing sensitivity to the changes in probabilities in Experiment 3. An analysis of their choice data revealed a strong preference for one side.

As mentioned before, the schedule and obtained probabilities differed. Table 12 shows the average, minimum and maximum obtained probabilities for the terminal links of the 16 sessions by condition. Overall, pigeons' experienced probabilities approached the schedule probabilities across condition.

Table 12
Averaged, minimum and maximum obtained probability of reinforcement of each terminal link stimuli.

Pigeon	P730				P053				P435				P795				P917			
Pair .9/.4																				
Scheduled	.90	.40	.50	.50	.90	.40	.50	.50	.90	.40	.50	.50	.90	.40	.50	.50	.90	.40	.50	.50
Avg. obtained	.91	.39	.48	.50	.88	.41	.50	.50	.90	.40	.50	.50	.91	.40	.49	.51	.90	.40	.50	.50
Min obtained	.75	.34	.25	.47	.63	.38	.38	.47	.75	.31	.38	.47	.75	.31	.38	.44	.75	.34	.38	.47
Max obtained	1.0	.47	.63	.53	1.0	.47	.63	.56	1.0	.59	.63	.56	1.0	.59	.63	.56	1.0	.47	.63	.53
Pair .6/.1																				
Scheduled	.10	.60	.50	.50	.10	.60	.50	.50	.10	.60	.50	.50	.10	.60	.50	.50	.10	.60	.50	.50
Avg. obtained	.11	.61	.51	.51	.12	.60	.50	.49	.11	.60	.52	.50	.11	.60	.50	.50	.11	.60	.51	.50
Min obtained	0.0	.56	.38	.47	0.0	.41	.25	.38	0.0	.59	.38	.47	0.0	.53	.38	.38	0.0	.56	.38	.44
Max obtained	.25	.66	.63	.56	.25	.69	.75	.56	.25	.66	.75	.56	.25	.66	.63	.56	.25	.66	.75	.53

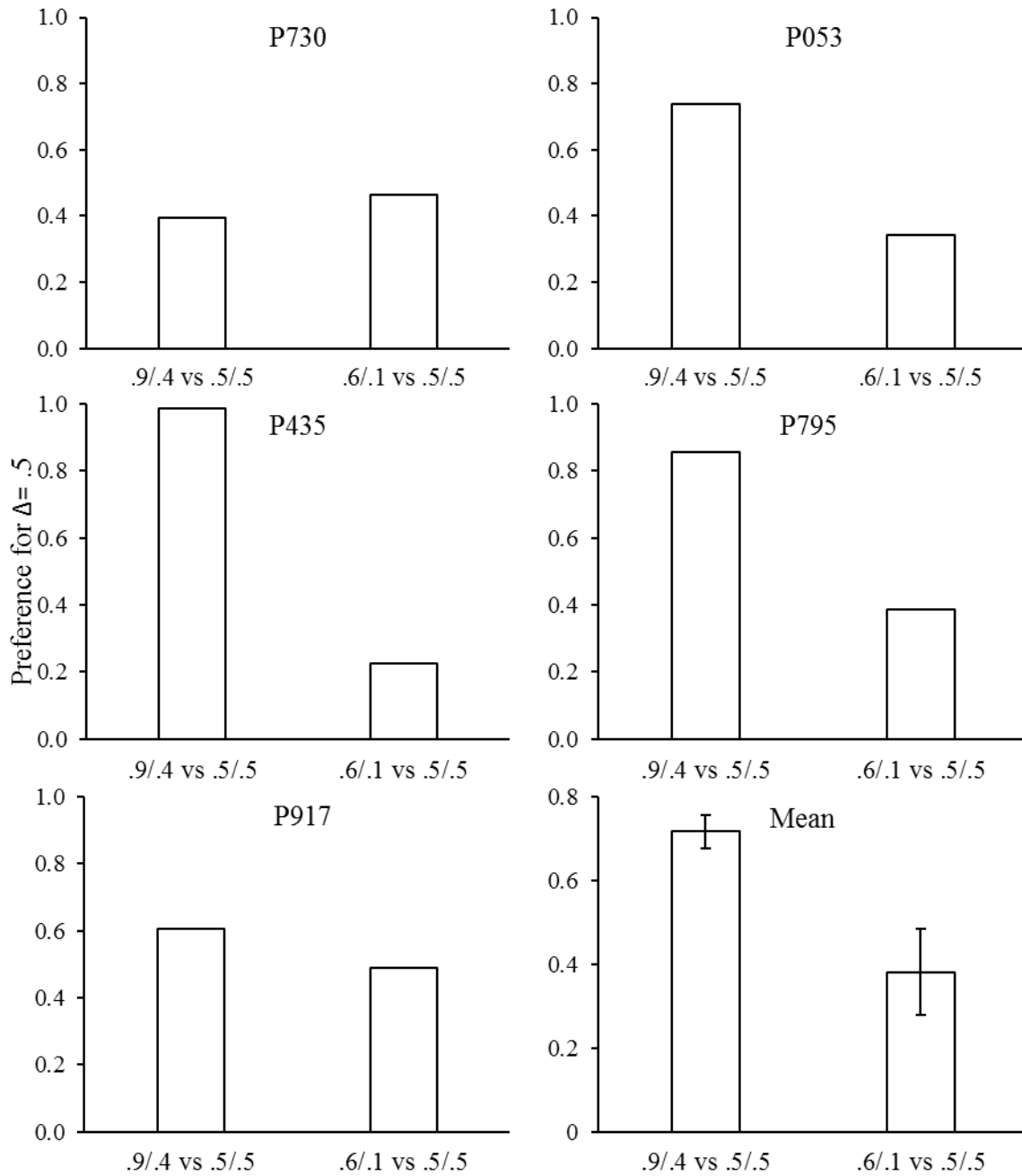


Figure 12. Choice data Experiment 4. The average preference for a $\Delta = .5$ of the last four sessions for the two conditions is presented by pigeon. The bottom right panel shows the average preference. Error bars show SEM.

Figure 12 shows the preference for the alternative with $\Delta = .5$ in the last four sessions. In condition with the pair .9/.4, four out of five pigeons developed a preference for the alternative with the greater Δ . In the condition with the pair .6/.1, all pigeons showed a preference near or below indifference (.5). A paired sample t-test found a marginal difference between the two conditions ($t(4) = 2.329, p = .080, 95\% \text{ CI } [-0.064, 0.734], d = 1.041$). A one

sample t-test for Condition .6/.1 found that preference for the alternative with $\Delta = .5$ was marginally below .5, $t(4) = -2.539$, $p = .064$, 95% CI [-0.249, 0.011], $d = -1.135$.

The average of the median latencies of the last four sessions in Figure 13 shows that latencies for the preferred alternative with the pair .9/.4 were shorter than on the other alternative. Latencies for alternative with the pair .6/.1 were slightly longer than on the other alternative. A repeated-measures ANOVA with Condition and Alternative as within-subject factors found no significant effect of Condition ($F[1, 4] = 1.160$, $p = .342$, $\eta p^2 = 0.225$) or Alternative ($F[1, 4] = 0.019$, $p = .896$, $\eta p^2 = .05$), but found a significant Condition x Alternative interaction ($F[1, 4] = 10.294$, $p = .033$, $\eta p^2 = 0.720$). This suggests there is a difference in latency between alternatives in .9/.4 but not in .6/.1.

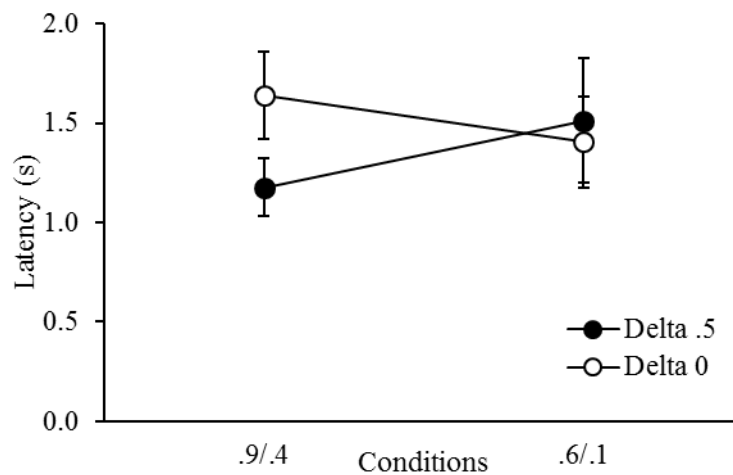


Figure 13. Averaged median latency data for forced trials in Experiment 4. Black dots show latencies for alternatives with $\Delta = .5$, white dots show latencies for alternatives with $\Delta = 0$. Error bars correspond to the standard error of the mean.

Overall, the results of Experiment 4 were similar to Experiment 3: Pigeons prefer the greater Delta with the pair .9/.4, as they preferred the pair 1/.5; and do not prefer the pair .6/.1, similar to the pair .5/0 in the previous experiment. Even though the reinforcers were programmed differently between experiments, the results were alike. This was also corroborated in the analysis of latency, in which shorter latencies were found for the preferred alternative, but no significant difference between alternatives appeared when the preference was marginally below indifference.

Furthermore, to evaluate whether $p = 1$ or $p = 0$ were special conditions, we compared preference for these conditions between experiments for the pigeons that participated in both

experiments ($n = 5$). A paired sample t-test was performed between conditions with the pair 1/.5 and .9/.4, finding no significant difference ($t(4) = 0.345$, $p = .747$, 95% CI [-0.264, 0.339]). A comparison between the pairs .5/.0 and .6/.1 also found no difference ($t(4) = 0.862$, $p = .437$, 95% CI [-0.139, 0.264]). Figure 14 shows the preference for the conditions of both experiments by pigeon. We observed a similar preference for the first two conditions in the graph (pairs .5/.0 and .6/.1); and a general preference above chance for the other three conditions with the exception of one pigeon (Pigeon P730 with the pair .9/.4), suggesting some type of threshold between the first two pairs and the other pairs of $\Delta = .5$.

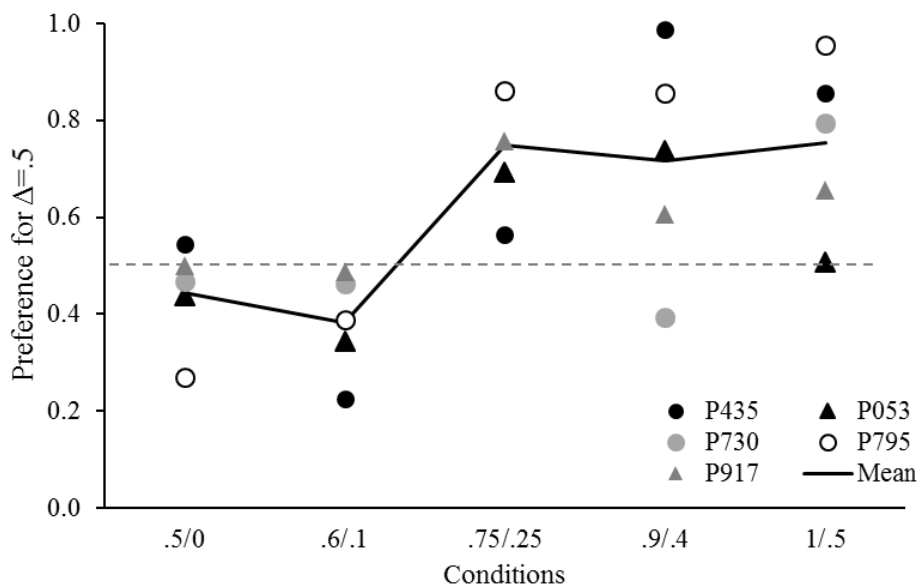


Figure 14. Preference for the Delta of .5 in experiments 3 and 4 by pigeon.

The results did not fully support the prediction of the Delta-Sigma hypothesis: First, preference depends on the specific probabilities of reinforcement used in the terminal links. Second, the results suggest that neither 1 nor 0 are special conditions. Finally, it is unclear how the two probabilities in the terminal link contribute to the final value of the alternative.

General discussion

The present study further tests the Delta-Sigma hypothesis. It evaluated the prediction that when all parameters are constant but Delta, preference depends only on the difference in reinforcement probability independently of the specific probabilities associated with each signal. We tested a constant Delta ($\Delta = .5$) combined with different probabilities of

reinforcement in each terminal link against a constant Delta ($\Delta = 0$) with constant probabilities of .5 in both terminal links. In Experiment 3, we replicated the results from previous research (see Chapter II), finding a preference for $\Delta = .5$ relative to $\Delta = 0$ when the pair in the former was .75/.25. A preference for $\Delta = .5$ was also obtained with the pair 1/.5. However, a preference for $\Delta = 0$ over $\Delta = .5$ was found with the pair .5/0. Experiment 4 extended these results with another two sets of values that avoided the extremes in the distribution of reinforcement probabilities, i.e. pairs .9/.4 and .6/.1 (both conditions compared to $\Delta = 0$). Similar to Experiment 3, a preference for the $\Delta = .5$ was found with the first pair, but not with the pair of probabilities of .6 and .1. Latency was consistent with the preference data in both experiments, where we found that the preferred alternatives had shorter latencies (Kacelnik et al., 2011).

The results from both experiments are partially supported by the Delta-Sigma hypothesis. Because the Σ ratio was approximately constant across conditions, it does not play a role in the calculation of the preference (see Equation 3). According to the model, animals should always prefer the greater Delta, independently of the specific probabilities of reinforcement. Even with a minimum β (i.e. value near zero), a parameter for individual differences in sensitivity to the Delta value, the model predicts near indifference but never a preference against the greater Delta. The study found that pigeons preferred the greater Delta in three out of five conditions between both experiments; these three $\Delta = .5$ were preferred in a similar degree, as predicted for the model. The two conditions in which pigeons did not prefer the expected alternative were those in which one of the terminal links was associated with very low or no reinforcement (condition with pairs .5/0 and .6/.1).

The Delta-Sigma hypothesis was proposed as another case in which contrast is influencing behavior. The idea is very similar to the within-trial contrast proposed by Zentall (2005), which presumes that the hedonic state of a subject changes between the end of a less appetitive event and a stimulus signaling reward or the reward itself (see, Singer, Berry, & Zentall, 2007). In the typical suboptimal choice task, the preferred alternative has a $\Delta_1 = 1$ ($p_{1,1} = 1$ and $p_{1,2} = 0$), the maximum possible difference in probability of reinforcement; against a $\Delta_2 = 0$ ($p_{2,1} = p_{2,2} = .5$), the minimum difference in probability of reinforcement within an alternative. When the former alternative is chosen, the subject is often confronted with a signal for no reinforcement ($p_{1,2} = 0$). This negative event sets a context in which the appearance of the signal for reinforcement changes the hedonic states by contrast, and in

consequence, the alternative gains more value. A similar situation happens in the present experiments when a signal for unlikely reinforcement appears ($p_{1,2} = .25$). This negative event sets the context in which the appearance of a signal associated with a probability of reinforcement of .75 gives a positive contrast, and in consequence, the alternative gains more value compared with an alternative with no contrast. In comparison, in the two conditions in which the animals prefer the Delta of 0 ($p_{2,1} = p_{2,2} = .5$), the alternative with the Delta of .5 had a signal that gives information about low or no reinforcement ($p = 0$ or $p = .1$ in different conditions) together with a low or uninformative signal ($p = .6$ or $p = .5$, respectively). In these two conditions, the alternative with the Delta of .5 had a terminal link that gave a context of uncertainty. When the other terminal link was experienced, indicating certain absence of reinforcement, the contrast was negative, ergo the value of the alternative decreased.

There are, however, some problems with the idea of one of the two terminal links within an alternative working as a context, because we need to assume a priori which terminal link has that role. One solution is to assume the context is determined by the most frequent stimulus (i.e., r_1 for Alternative 1 or r_2 for Alternative 2). In the condition with the pair 1/.5 for the Delta .5, the most frequent signal was the one reinforced with probability of .5 ($r_1 = .1$), then when the signal associated with a probability of 1 appeared, the contrast was positive. The same occurs with the pair .9/.4 ($r_1 = .2$). The opposite is true for the pairs .5/0 ($r_1 = .9$) and .6/.1 ($r_1 = .8$), in which the contrast was negative. However, the pair .75/.25 ($r_1 = .5$) is problematic because both signals were presented equally often. As a whole, the results of the present study suggest that the two variables proposed by the Delta-Sigma model are not sufficient to explain the preference in the suboptimal choice task.

The Temporal-information approach (Cunningham & Shahan, 2018) is based on the idea that temporal information conveyed by a signal predicting food works as a Pavlovian conditioned stimulus and as an instrumental conditioned reinforcer. The model explains that the suboptimal choice preference is developed because the suboptimal alternative has more information about when food will be delivered rather than the optimal alternative. Moreover, the model includes a role for competition between the temporal information associated with the signals and the relative primary reinforcement of each alternative. Thus, the model highlights the same variables that the Delta-Sigma model does. However, as the authors recognized, the Temporal-Information approach does not specify how to quantify the temporal information offered by signals with different probability of reinforcement.

Specifically they can only deal with probabilities of 1 and 0 in the suboptimal alternative, in which they also assumed the signal reinforced with probability 0 has no role in choice, and a probability of .5 in the optimal alternative. Therefore, the model has no predictions but in the condition with the pair of probabilities of .5/0, where indifference is anticipated.

In general terms, RRM is based on optimal foraging principles and suggests that animals behave following an information-seeking strategy (Fortes et al., 2016; Fortes et al., 2017; Vasconcelos et al., 2015). In the original proposal (see Figure 9, left panel), Red is a valuable stimulus because it provides information about reinforcement; thus, animals attend to it. However, Green is a signal for no reinforcement and since in the natural environment that would trigger another foraging bout, they never learn to attend to it. However, as in Cunningham and Shahan's model, the RRM could not explain situations in which the probabilities of reinforcement were different from 1 and 0. The latest version of the RRM (see Fortes et al., 2018) added an engagement function that allows the model to deal with these situations. It is assumed that animals engage or pay attention to a signal depending on its associated probability of reinforcement. The latest version of RRM also faces some difficulties to explain the results of this study. The model predicts a preference for the greater Delta in all conditions, but the pair .5/0, in which indifference is expected. Similarly, Mazur's hyperbolic discounting model (Mazur, 1996; 1995), even though it assumes different processes, is mathematically similar to the RRM and in consequence fails and succeeds in the same conditions as the RRM.

Finally, Zentall's contrast (e.g., Roper & Zentall, 1999) proposed that when a pigeon chooses the suboptimal alternative, its food expectation equals the overall probability of reinforcement associated with that alternative, $\Sigma_1 = .2$. However, when Red appears, its food expectation increases to 1, a positive contrast of +0.8 expectation. When the pigeon chooses the optimal alternative, it expects food with probability $\Sigma_2 = .5$. That does not change because the probabilities of reinforcement are equal to .5; a contrast of 0. It would follow that a preference for the greater contrast is always expected. However, this account does not explicitly suggest what happens in situations where there is a negative contrast. We previously took a step forward assuming that only positive contrast ($p-\Sigma > 0$) influences preference (see Chapter II). With this assumption, Zentall's account encounters the same difficulties as previous models. It predicts a preference for the greater delta ($\Delta = .5 > \Delta = 0$)

in the conditions with a positive contrast (i.e., when one of the signals has a probability of reinforcement greater than .5), but it also predicts indifference in condition with the pair .5/0.

Overall, the presented models do not fully address all or some conditions of the study. The Temporal-information approach, Mazur's model, RRM and Zentall's contrast account predict indifference when the pair .5/0 was presented against the pair of probabilities .5/.5. The Delta-Sigma hypothesis predicts preference for the pair .5/0. In the other conditions of Experiment 3 and 4, the models are either silent (i.e. Temporal-information approach) or predict, as hypothesized for the Delta-Sigma model, a preference for the greater Delta. Notably, all the models that have predictions for both experiments encounter difficulties in the same conditions, suggesting perhaps the influence of another variable that was not considered before.

Some research has found that information about reinforcement or absence of reinforcement are not equivalent (Dinsmoor, 1983). In effect, this has been extensively discussed in the 'observing response' literature. Observing response refers to situations in which subjects can respond to access a stimulus that gives information about the contingencies; however, this behavior does not alter the schedule of reinforcement. Hirota (1972) found that pigeons spent more time pressing a pedal that shows a signal associated with reinforcement than they did for a signal for no reinforcement. Similarly, Browne and Dinsmoor (1974) designed an experiment where the behaviors to access information for food or absence of food were mutually exclusive, and their results indicated that pigeons prefer information about food over absence of food.

Previous evidence suggests that preference for information depends on the type or *valence* of the information, i.e., the attractiveness/goodness or averseness/badness of an event or stimulus (Beyeler et al., 2016; Sharot & Garrett, 2016). Nevertheless, it may be problematic to extrapolate these results to the suboptimal choice task, because in this procedure each alternative offers two stimuli associated with different probabilities of reinforcement that can have different degrees (i.e. amount) of information and valence (i.e. attractiveness). The problem of how organisms integrate different sources of information is not new (Massaro & Friedman, 1990; Shaw, 1982) and does not have a clear solution yet.

Some authors have proposed that information about something appealing is treated and learned in a different way than information about something repelling or aversive

(Beyeler et al., 2016; Olsen & Pracejus, 2004; Tversky & Kahneman, 1992). However, how we learn and combine information about different aspects of our environment is still an open question. Ultimately, evaluating a broad range of parameters is necessary and fundamental to having a better understanding of what is tracking animals' preference in the suboptimal choice task.

CHAPTER IV³

THE ROLE OF INHIBITION IN THE SUBOPTIMAL CHOICE TASK

³This Chapter reproduces the submitted version of the paper:

González, V. V., & Blaisdell, A. (2019). The Role of Inhibition in The Suboptimal Choice Task. *Manuscript submitted for publication.*

Abstract

Given a choice, pigeons prefer an initial-link stimulus that is followed by reliable signals that food will be delivered (S+) or not (S-) after a delay, over an alternative initial-link stimulus that is followed by unreliable signals of food; even when the former yields a lower overall probability of food. This suboptimal preference has been attributed to the combination of a biased attraction to the S+ and ignoring the S-. We evaluated the inhibitory properties of the S- in two experiments to investigate its role in suboptimal choice. In Experiment 5, pigeons were trained in an autoshaping procedure with the four terminal link stimuli of the suboptimal choice task; S+ was continuously reinforced, S3 and S4 were each partially reinforced on a 50% schedule, and S- was never reinforced. Summation tests showed that S- acquired inhibitory properties during training. Experiment 6 replicated the results of the summation tests after training on the full suboptimal choice procedure. Furthermore, the inhibitory properties of the S- positively correlated with the strength of suboptimal preference. Future models explaining performance in the suboptimal choice task should take into account inhibition to the S- as key player in suboptimal choice.

When hungry pigeons are confronted with a choice between two alternatives, one predicting a low probability of food but that is followed by predictable signals of food, and another predicting a higher probability of food but that is followed by unpredictable signals of food; they systematically choose the former. This preference has been called suboptimal because they fail to maximize food intake (Vasconcelos et al., 2018). For instance, Stagner and Zentall (2010) gave pigeons a choice between two initial-link options in a concurrent chain. If pigeons chose the suboptimal alternative, on 20% of the trials the choice led to a terminal-link stimulus (S+) that was always followed by food after 10 s, while on the other 80% of the trials another terminal-link stimulus (S-) appeared for 10 s ending always without food. If the optimal alternative was chosen, one of two terminal-link stimuli (S3 and S4) would appear, and after 10 s was followed by food on half of the trials regardless of which terminal-link stimulus had been presented. Pigeons showed a strong preference for the suboptimal option, even though the overall probability of food for the suboptimal alternative was 2.5 times lower than for the optimal alternative (20% vs 50%, respectively).

Most models, although differing in the explanation or mechanism, assumed that the preference is due to the combination of a biased attraction to the S+ and ignoring the S- (Zentall, 2016a). For instance, the RRM (Vasconcelos et al., 2015), a model based on optimal foraging theory, proposed that animals follow an information-seeking strategy, in which the suboptimal preference is due to the information embedded in the S+. The model also assumed that, since in nature animals naturally are not attracted to signals that reliably predict the absence of reward, such as food, the S- plays no role in the decision process. Similarly, Cunningham and Shahan (2018) advanced an explanation based on the idea that animals learn the temporal relation between events, and given there is no event that follows the S-, it is ignored when making response decisions (for other proposals that disregard the role of the S- see McDevitt et al., 2016; Zentall, 2016a).

The general assumption that the S- plays no role in the decision process is based on evidence suggesting animals are insensitive to changes to the S-. For instance, Fortes et al. (2016) manipulated the probability of occurrence and the duration of the S-, finding that animals continued showing a suboptimal preference even when the S- was presented on 95% of the trials, or the delay of this stimulus was increased to 200 s. Stronger evidence comes from Laude, Stagner and Zentall (2014) in a study suggesting that the S- develops inhibitory

properties at the beginning of the training, but that inhibition diminishes with more training. They reported this effect using a variant of the typical suboptimal choice task in which the magnitude of the reinforcer following each stimulus, rather than the probability of reinforcement of each stimulus differed between alternatives (See Figure 15). Specifically, the S+ and S- in the suboptimal alternative led to 10 and 0 pellets, respectively; while S3 and S4 in the optimal alternative always resulted in 3 pellets. Thus, the overall probability of reinforcement was 2 pellets for the suboptimal choice and 3 pellets for the optimal choice. Pigeons were trained in this task using color keys for the S+, S3, and S4 terminal links, and a vertical line for the S-. To assess the inhibitory properties of the vertical line S-, the rate of response to the compound S+S- was compared to S+ alone early and late in training using a within-subject design (Experiment 1) and between-groups design (Experiment 2). In both experiments, they found strong inhibition when tested early during training, but a significant reduction in inhibition when tested much later in training.

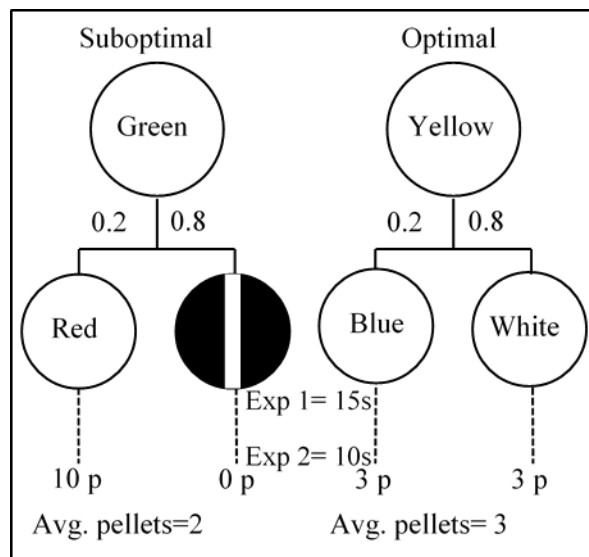


Figure 15. Procedure used by Laude, Stagner and Zentall (2014) to evaluate the inhibitory properties of the stimulus never reinforced (S-, black circle with a white vertical line). After the choice, each signal was followed by a different magnitude of reinforcement.

Despite the large effects in both experiments, the research designs have some particularities that warrant caution in the interpretation of the results of Laude et al. (2014). First, they used an atypical procedure in which reinforcement was always 100% and therefore each signal was equally and perfectly informative about the outcome. This is relevant especially for explanations of suboptimal choice in terms of the informativeness of the terminal link stimuli. In the procedure used by Laude et al. (2014), all terminal link stimuli

were informative. Until now, it is uncertain whether the same mechanisms for the development of suboptimal choice underlie both types of procedure (Daniels & Sanabria, 2018). Second, the S- stimulus was a vertical line, while the remaining stimuli used for the initial and terminal links were colored keys. Color has been shown to be a more salient dimension to pigeons than stimulus shape (e.g., Blaisdell & Cook, 2005). Thus, failure to counterbalance the stimulus roles as S+ and S-, and choice of the less salient visual dimension for the S- may have contributed to the loss of stimulus control by the line orientation stimulus in their study. Third, the only assessment of inhibition to the S- was by measuring the difference in peck rate to the S+ versus the S+S- compound. Additional tests involving other stimulus compounds of the S+ and another stimulus (e.g., a novel stimulus or a different excitatory stimulus) are required to rule out external inhibition as an explanation for the reduction in response on compound test trials (Rescorla, 1969). Finally, an experiment using rats as subjects and the typical suboptimal choice task (i.e. the signals after choice have different probabilities instead of magnitude of reinforcement) found that inhibition to the S- increased with training. Despite showing development of strong inhibition to the S-, the rats did not show a suboptimal preference (Trujano et al., 2016). The authors concluded that the strong difference found between pigeons and rats in the suboptimal choice task is related to differences in the impact of conditioned inhibitors.

The purpose of the current experiments was to provide an evaluation of the inhibitory properties of the S- using the typical suboptimal choice task in pigeons. In Experiment 5, pigeons were trained in a Pavlovian autoshaping procedure involving only the terminal link stimuli (S+, S-, S3, and S4). Pigeons received 30-s presentations of each stimulus on separate trials. Each stimulus signaled a specific probability of reinforcement: $p(\text{food}|S+) = 1$; $p(\text{food}|S-) = 0$; $p(\text{food}|S3) = .5$; $p(\text{food}|S4) = .5$. Occasional non-reinforced probe trials with elements and compounds S+S-, S3S4, S3S- and S+S4 were delivered to assess the inhibitory properties of S-. If the S- had acquired inhibitory properties, we would expect a reduction in the responses to S+S- compared to S+ alone or to the S3S4 compound. The compound S3S- and S+S4 compounds were used to compare the excitatory properties of a stimulus partially reinforced with a stimulus continuously reinforced. If the Pavlovian contingencies signaled by the terminal link stimuli are learned, the S- should show the development of conditioned inhibition as a function of amount of training.

In Experiment 6, pigeons were trained on the typical suboptimal choice task involving both initial and terminal link stimuli. We measured the development of both conditioned inhibition to S- using summation tests with compound stimuli, and of preference for the suboptimal choice stimulus on choice trials during training. If suboptimal choice correlates with the development of inhibition to the S-, as we predict, then development of suboptimal choice should track the emergence of conditioned inhibition to the S- as shown on summation tests. Alternatively, if suboptimal choice depends on ignoring the S-, as shown by Laude et al. (2014), then development of suboptimal choices should track the loss of stimulus control by the S- on compound tests. If inhibition does not decay during training, it will challenge most models of suboptimal choice that assumed the S- plays no role in the decision process. Although it is possible that learning inhibition to S- is independent from the development of a suboptimal preference, current and new models would need to integrate an explanation of how learning about a stimulus is not considered at the moment of choice.

Experiment 5

Before assessing the role of inhibition in suboptimal choice, we wished to determine whether terminal link stimuli, when trained on their own, acquire excitatory and inhibitory properties. Thus, pigeons received training with each of the terminal link stimulus elements (S+, S-, S3 and S4) on separate trials in a Pavlovian autoshaping procedure. S+ was always followed by food, S3 and S4 were followed by food on 50% of the trials in each session, and the S- was never followed by food. We measured the peck response rate to each stimulus. The inhibitory properties of the S- was evaluated using a summation test, in which the S- was presented in compound with a stimulus with excitatory properties (S+S- and S3S-). The response rate to the compounds was compared to the response rate to the elements and to other compounds such as S3S4 and S+S4. If S- acquired inhibitory properties as a function of training, we predicted a reduction in response rate to the compounds S+S- and S3S- compared to S+, S3S4, and S+S4.

Subjects

Five adult homing pigeons (*Columba livia*) from Double T farms, three males and two females, served as subjects. These pigeons had previously participated in a wide variety of behavioral experiments, including spatial overshadowing, object location encoding, response variability, and pattern learning, but were naïve with respect to the current procedures and

stimuli which were selected to minimize transfer from prior experience. Subjects were individually housed in steel home cages with metal wire mesh floors in a vivarium. They were maintained at 80% of their free-feeding weight, but were given free access to water and grit while in their home cages. Testing occurred at approximately the midpoint of the light portion of the 12-hour light-dark cycle.

Materials and Methods

Apparatus.

The experiment was conducted in a flat-black Plexiglas chamber (38 cm wide x 36 cm deep x 38 cm high). All stimuli were presented by computer on a color LCD monitor (NEC MultiSync LCD1550M). Stimuli were presented using the coding language Python (Python Software Foundation, <https://www.python.org/>) and the extension PsychoPy (Peirce, 2007). The bottom edge of the viewing window was 13 cm above the chamber floor. Pecks to the monitor were detected by an infrared touchscreen (Carroll Touch, Elotouch Systems, Fremont, CA) mounted on the front panel. A custom-built food hopper (Pololu, Robotics and Electronics, Las Vegas, NV) was located in the center of the front panel, its access hole flush with the floor. The hopper could deliver 3-s access to mixed grain as a food reward. All experimental events were controlled and recorded with a Pentium III-class computer (Dell, Austin, TX). A video card controlled the monitor in the SVGA graphics mode (800 x 600 pixels).

Stimuli.

Each of the four stimuli, S+, S-, S3, and S4, were composed of two circles (either red, green, yellow, or blue) vertically or horizontally aligned (Figure 16), with a size of 100 x 100 pixels. All the stimuli were presented in the center of the screen against a grey background.

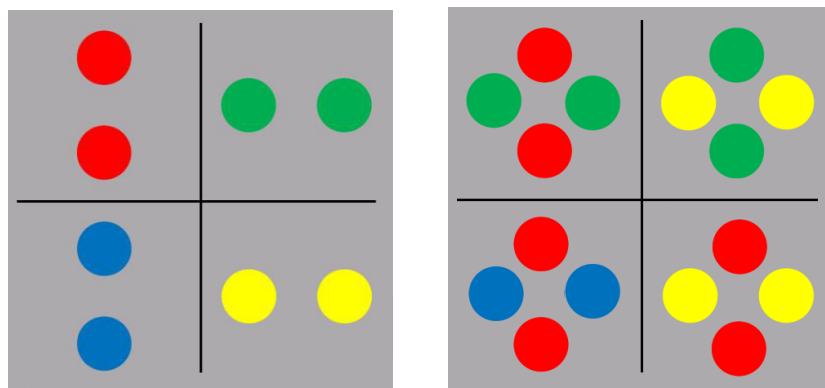


Figure 16. Left panel: Examples of stimuli presented during the Pre-training and Training phases of Experiment 5. Right panel: Examples of compound stimuli presented on probe trials during the Testing phase of Experiment 5. Color and orientation were counterbalanced across pigeons with the restriction that S+ and S3, and S- and S4 should have the same orientation to be able to create the compound stimuli.

Procedure

Pre-training.

Pigeons were initially trained to peck each stimulus. A session consisted of 40 trials with an average duration of 10 minutes. Each stimulus was presented 10 times in random order within each session. On each trial, one pseudorandomly selected stimulus from the set of four was presented at the center of the grey screen. The stimulus remained on the screen until the required number of pecks was completed. Each trial ended with the delivery of the food reward followed by a 10-s Intertrial-Interval (ITI) during which the screen was black. Pigeons received two sessions of pre-training, the first under a continuous schedule of reinforcement (CRF), the second under a fixed-ratio (FR) 10 schedule of reinforcement.

Training.

Pigeons received a total of 40 sessions of training on an autoshaping procedure, six days per week. Each session consisted of 80 trials. In each session, the S+ was presented on eight trials each ending in food reward, the S- was presented 32 times and was never followed by food, S3 was presented eight times, four trials followed by food reward and four non-rewarded, and S4 was presented on 32 trials, in which 16 were followed by food reward and the remainder were non-rewarded. On each trial, a stimulus appeared in the center of the screen for either 10 s (during the first 10 sessions), or 30 s (during the remainder of training). The order of the trials was randomized. An ITI of 10 s with a black screen separated trials. Stimulus role was pseudorandomly assigned across subjects. During the first 10 sessions of training we observed low rates of pecking by some pigeons, and thus increased presentation duration to 30 s to allow more time to accumulate pecks. Pigeons received 15 sessions of training with stimuli of 30-s duration.

Test.

Following the 15 training sessions with the longer duration stimuli, pigeons received test sessions in blocks of five sessions interspersed with blocks of 5 training-only sessions (e.g., Sessions 16-20 for testing, Sessions 21-25 for training, Sessions 26-30 for testing, etc.). Pigeons received four blocks of five test sessions. Each test session of 56 trials started with ten presentations of the training stimuli as described for the training procedure. After the 10th

trial of each test session, 9 non-reinforced test trials were interspersed among training trials for the remainder of the session. Test trials consisted of 2 presentations each of the following compounds: S+S-, S3S4, S3S-, and S+S4 (Figure 16, right panel), and one non-reinforced presentation of the S+. Each test trial was 30 s in duration and, as with training trials, the number of pecks during each test trial were recorded.

Results

The peck rate for each stimulus was calculated for each session. Session peck rates were calculated for each stimulus and then averaged across blocks of five sessions, obtaining eight blocks for the training phase and four blocks for the test phase. Given that the rate of response (RR) was highly variable between pigeons, the data were normalized. For each block, the S+ RR was defined as the reference value in percentage (i.e. $(\text{stimulus RR}_b / \text{S+ RR}_b) * 100$).

The upper panel of Figure 17 shows the normalized mean peck rate for each stimulus as a function of training blocks. S+, S3, and S4 maintained similar RRs, but the RR for S- decreased from Block 1 to Block 2 and remained low for the remainder of training.

Supporting these observations, a Repeated-Measures Analysis of Variance (ANOVA) conducted on normalized peck rates with Block and Stimulus as factors revealed a main effect of Stimulus, $F(3, 12) = 38.194, p < .001, \eta^2 = .905$, but no effect of Block, $F(7, 28) < 1.0$, nor interaction, $F(21, 84) = 1.481, p = .107, \eta^2 = .270$. Post-hoc analyses using Bonferroni correction indicated that S- RR was lower than S+, S3, and S4 RR ($p_{\text{bonf}} < .05$), while RRs across the remaining stimuli did not differ. A paired-sample t-test comparing the RR to S- in the first two blocks, showed a significant reduction of pecks from Block 1 ($M = 61.64, SD = 60.16$) to Block 2 ($M = 15.54, SD = 18.59$), $t(4) = 2.290, p < .05, d = 1.024$.

The bottom panel of Figure 17 shows the normalized response rate across test blocks for each stimulus. To evaluate any possible influence of testing on performance during training, we compared the performance of Block 5 which occurred prior to the first block of testing to training performance in Block 6. A repeated-measures ANOVA conducted on normalized response rates using Block and Stimulus as factors, showed a main effect of Stimulus, $F(3, 12) = 12.466, p < .001, \eta^2 = .757$, but no effect of Block, $F(1, 4) < 1.0$, nor interaction, $F(3, 12) < 1.0$. Suggesting that the training performance did not change due to the introduction of testing sessions.

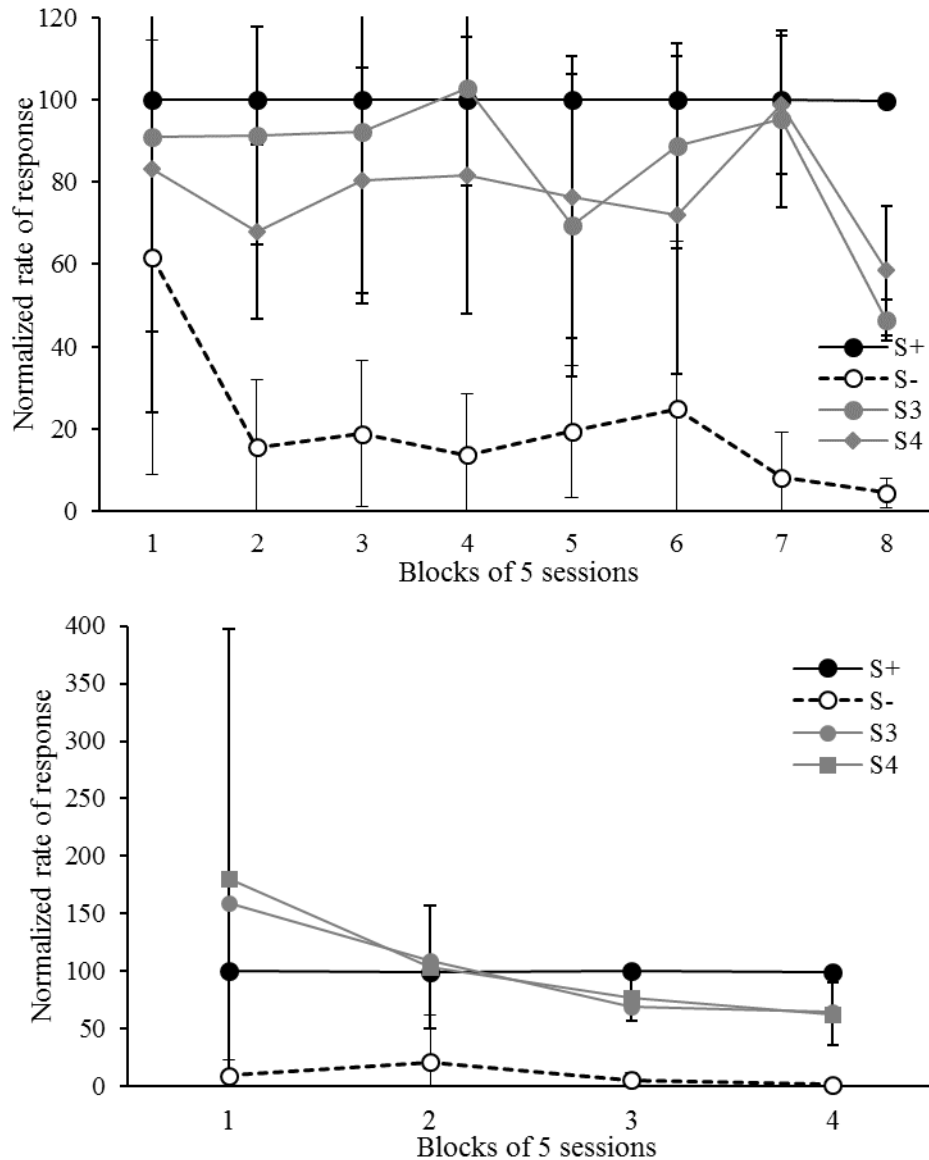


Figure 17. Mean normalized RR for each block of five sessions of training (Upper panel) and testing (Lower panel) for each stimulus in Experiment 5. Error bars represent 95% confidence interval.

A repeated-measures ANOVA conducted on normalized response rates to the probe trial stimuli in the test phase, with Block and Stimulus as factors, found an effect of Stimulus, $F(7, 28) = 4.987, p < .001, \eta^2 = .555$, but no effect of Block, $F(3, 12) = 1.302, p = .319, \eta^2 = .246$, nor interaction, $F(21, 84) = 1.397, p = .143, \eta^2 = .259$. Since the main objective of the experiment was to assess inhibitory strength of the S- using a summation test, we performed a paired-sample t-test comparing responding to S+ with responding to S+S- trials in every block. There were no differences in Block 1, $t(4) < 1.0$ or Block 2, $t(4) < 1.0$, but there were differences in Block 3, $t(4) = 2.875, p < .05, 95\% \text{ CI}[0.765, 4.983]$, and 4, $t(4) = 12.614, p < .05, 95\% \text{ CI}[4.333, 20.611]$.

Since evidence for inhibition to S- did not develop until the third block of testing, peck rate data were pooled across the last two blocks of testing to assess our predictions (Figure 18). A repeated-measures ANOVA of peck rates with Stimulus has a factor found a main effect of Stimulus, $F(2.41, 21.67) = 30.70, p < .001, \eta^2 = .773$. Post-hoc analysis using Bonferroni corrections revealed a lower RR for S- compared to S+, S3, and S4 ($ps < .001$), evidencing that pigeons did not peck the stimulus associated with absence of food. Likewise, RR to S+ was higher than to S3 ($p = .006$) and S4 ($p = .002$). RR did not differ between S3 and S4 ($p = .689$). These results suggest that S+, a continuously reinforced stimulus, elicited higher peck rates than did partially reinforced stimuli S3 and S4. The lowest peck rates were observed to the S- which never signaled reinforcement. Post-hoc analyses using Bonferroni corrections comparing elements with compounds found higher RR to S+ than to the S+S- compound ($p < .001$), giving evidence of the inhibitory properties of the S-. Similarly, the RR to S3 was higher than to the S3S- compound ($p < .001$), also indicating inhibitory properties of the S- when tested in conjunction with a partially reinforced stimulus. RRs to S+ and S+S4 did not differ ($p = 1.00$). This is an important comparison as it shows that merely presenting two stimuli together in a novel compound did not produce external inhibition. Moreover, the similar peck rates between S+ and S+S4 suggests that excitatory properties were averaged rather than summed. Similar results were found when comparing S4 against S+S4 or S3S4 ($ps = 1.00$), and S3 with S3S4 ($p = 1.00$).

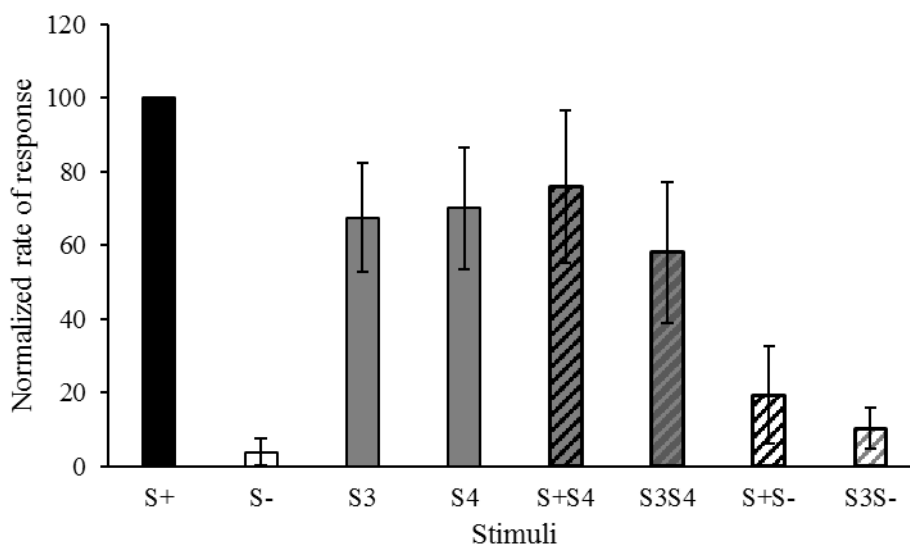


Figure 18. Mean normalized RR across the last two blocks of testing for each probe stimulus in Experiment 5. Error bars correspond to the 95% confidence interval.

We predicted RR to compound test stimuli to show a positive monotonic relation to the average probability of reinforcement signaled by the elements, with highest RR to the S+S4 compound signaling an average probability of reinforcement of .75 (i.e. $p(\text{food}|S+) = 1.0$ and $(\text{food}|S4) = .5$), followed by S3S4 with an overall probability of reinforcement of .5 (i.e. $p(\text{food}|S3) = .5$ and $(\text{food}|S4) = .5$), then S+S- (also overall probability of reinforcement of .5, but with S- serving as a conditioned inhibitor), and finally S3S- (overall probability of .25; from $(\text{food}|S3) = .5$ and $p(\text{food}|S-) = 0$). As anticipated, a within-subjects contrast for the compound stimuli revealed a significant linear trend, $F(1, 4) = 32.286, p < .001$. If RR to each compound reflected only the average probability of reinforcement signaled by the elements, we would predict equivalent RRs to the S3S4 compound and the S1S- compound, given that they both signal an average probability of reinforcement of .5. Given that the RR to the S+S- compound was significantly lower than to the S3S4 compound, this suggests the additional operation of conditioned inhibition to S- as a major factor in determining RR to the compound. That is, S- signals a 0 probability of reward that translates to a negative, rather than a neutral, value for that stimulus, or increased probability of reward omission. It is in this way that conditioned inhibitors exert their modulating effect over conditioned excitors, such as to reduce the response elicited by the CS and withdrawing from the inhibitory CS (Hearst, Bottjer, & Walker, 1980; Wasserman, Franklin, & Hearst, 1974).

Experiment 6

In the previous experiment, we used a Pavlovian autoshaping procedure, in which only the terminal links of the suboptimal choice task were presented, maintaining the frequency with which each stimulus is presented in the typical task. The objective was to assess inhibition to the S- in a summation test. The results confirmed that the S- passed a summation test of conditioned inhibition, thereby establishing that the S- acquired inhibitory properties during training. Having established inhibition to the S-, we can now test the hypothesis that development of preference for the suboptimal initial link stimulus is related to the acquisition of inhibition to the S- signaled by the suboptimal initial link stimulus. To test this, we trained pigeons on the typical suboptimal choice task. As Experiment 5, test sessions were periodically introduced later in training, in which we presented the nonreinforced compounds S+S4, S3S4, S+S- and S3S-. The same results from previous experiment were expected: A reduction in the number of responses (RR) with compound S+S- and S3S- compared to S+, S+S4 or S3S4. Furthermore, we expected suboptimal preference to develop

at the same time that the response rate to the S- decreased. Thus, we hypothesized that the strength of the suboptimal preference would correlate with the strength of inhibition.

Subjects

The same pigeons were used and maintained as in Experiment 5.

Materials and Methods

Apparatus.

The apparatus was the same as that used in Experiment 5.

Stimuli.

A new set of stimuli were created for purpose of Experiment 6. Two circles with a colorful pattern served as initial link stimuli, and four pairs of geometric shapes with a black and white pattern served as terminal link stimuli (Figure 19). All stimuli were 100 x 100 pixels. Stimuli were presented against a grey background, and appeared on the left and right side of the screen counterbalanced throughout each session. Assignment of stimuli to function was pseudorandomly assigned across subjects.

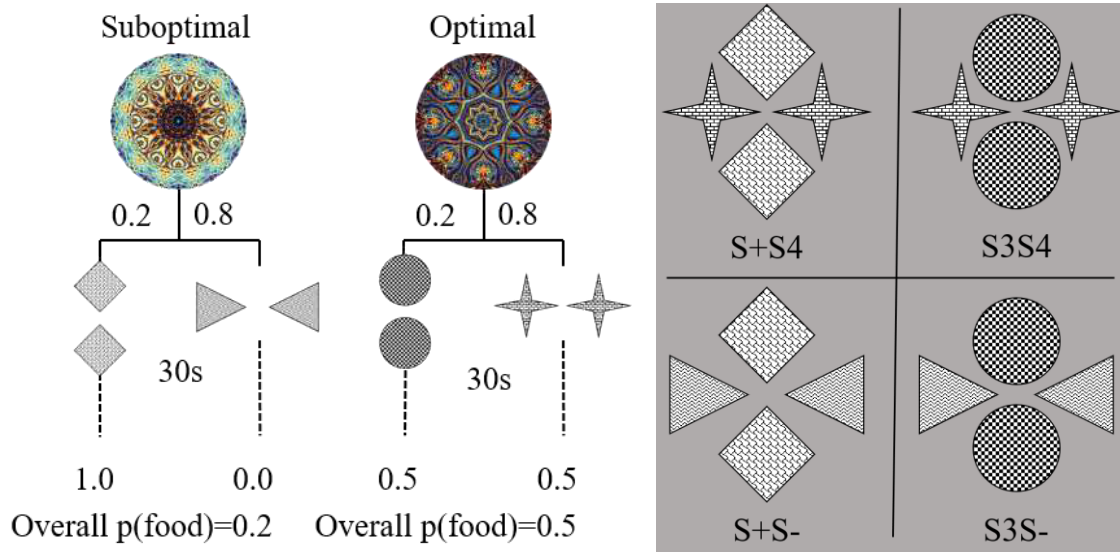


Figure 19. Left panel: Structure of the suboptimal choice task used on choice trials during training in Experiment 6. The initial link stimuli were counterbalanced across pigeons. Terminal link stimuli could be two pairs of figures (diamonds, triangles, circles, or 4-point stars) each one filled with a different pattern and presented vertically (S+ and S3) or horizontally (S- and S4). In this example, the vertical pair of diamonds correspond to the S+, the horizontal pair of triangles correspond to the S-, the vertical pair of circles correspond to the S3, and the horizontal pair of stars correspond to the S4. *Right panel:* Examples of four possible compound stimuli used during testing in Experiment 6. All stimuli used black patterns over a white background. Stimuli were presented half of the time in each side of the screen within each test session.

Procedure

Pre-training.

All stimuli (two initial links and four terminal links) were presented individually in a similar fashion as in Experiment 5.

Training.

There were two types of trials in each training session of 80 trials, free-choice (20 trials) and forced-choice (60 trials). In a free-choice trial, pigeons were presented with a choice between two circles presented on the left and right sides of the screen (Figure 19, left panel). When the pigeon pecked one of the initial link stimuli, both stimuli disappeared and was replaced by a terminal link stimulus presented for 30 s on the same side of the screen as the selected initial link stimulus. Choice of the suboptimal alternative was followed on 20% of the trials by the S+ stimulus (e.g. two diamonds arranged vertically), and on 80% of the trials by the S- stimulus (e.g. two triangles arranged horizontally). The S+ was always followed by food upon its termination, while the S- was never followed by food. Choice of the optimal alternative was followed on 20% of the trials by the S3 stimulus (e.g. two circles arranged vertically), and on 80% of the trials by the S4 stimulus (e.g. two horizontal stars). Both S3 and S4 were followed by food upon stimulus termination on 50% of trials in each session. A black screen presented for 10 s served as the ITI. On forced-choice trials, pigeons were presented with only one alternative initial link stimulus, with each type (optimal and suboptimal) appearing on 50% of the forced-choice trials in each session. The suboptimal and optimal alternatives for both forced-choice and free-choice trials appeared half of the time in each side of the screen (i.e., left/right counterbalanced). The duration of the session was set to a maximum of 90 minutes. The training phase consisted in 30 successful sessions (i.e. pigeon completing at least half of the trials on a session). Pigeons initially received 15 sessions of training, after which blocks of 5 training sessions were alternated with blocks of 5 test sessions.

Test.

Pigeons received four blocks of five test sessions. A test session had 96 trials, 80 training trials as described above, and 16 non-reinforced compound test trials. The first 10 trials in each test session always consisted of forced-choice training trials. After the 10th trial of the session, 70 forced-choice, 20 free-choice, and 16 test trials were randomly interspersed

throughout the remainder of the session. Test trials entailed the presentation of one of four compound stimuli (Figure 19, right panel) presented for 30 s on half of the trials on the left side and the other half on the right side of the screen.

Sessions were included in data analysis only if at least half of the training trials and all of the test trials had been presented. In consequence, two pigeons, Cousteau and Darwin, repeated some training sessions (10 and 5, respectively) and some test sessions (11 and 3, respectively). Sessions were averaged in blocks of five sessions, obtaining six blocks (seven for Herriot) for the training phase and four (five for Herriot) blocks for the test phase. Preference for the suboptimal alternative was defined as the number of choices to the suboptimal alternative divided by the total number of choice trials completed. RR to each terminal link stimulus or compound test stimulus was also recorded on each trial. Given that the RRs were highly variable across pigeons, the data were normalized to allow comparisons. For each block, the S+ RR was defined as the reference value in percentage, (stimulus RR/S+ RR) *100). As with choice preference, the preference was collapsed in blocks of five sessions.

Results

The left panel of Figure 20 shows choice preference across training blocks for each pigeon. The rate at which suboptimal preference developed varied across pigeons. Hawthorne acquire an almost exclusive preference for the suboptimal alternative by Block 3 (before the first test block was introduced). Goodall and Cousteau reached a similarly high preference for the suboptimal alternative by Block 4. Although not as strong as in the three birds discussed above, Darwin reached an asymptotic level of preference for the suboptimal alternative of around 0.6 by the second block of training. Finally, Herriot began with a strong preference for the optimal alternative, and only by Block 5 had reached 50% preference for the suboptimal alternative, and reaching a suboptimal preference close to .7 by the 6th block of training. A Repeated-Measures ANOVA conducted on percentage of suboptimal preference with Block as a factor showed a main effect of Block, $F(5, 20) = 15.72, p < .001, \eta^2 = .797$. Post-hoc analyses using Bonferroni correction showed a significant difference between Block 1 and Blocks 5 and 6 ($p < .05$), and Block 2 against Blocks 4 to 6 ($p < .05$), suggesting that most subjects developed a preference for the suboptimal alternative through training.

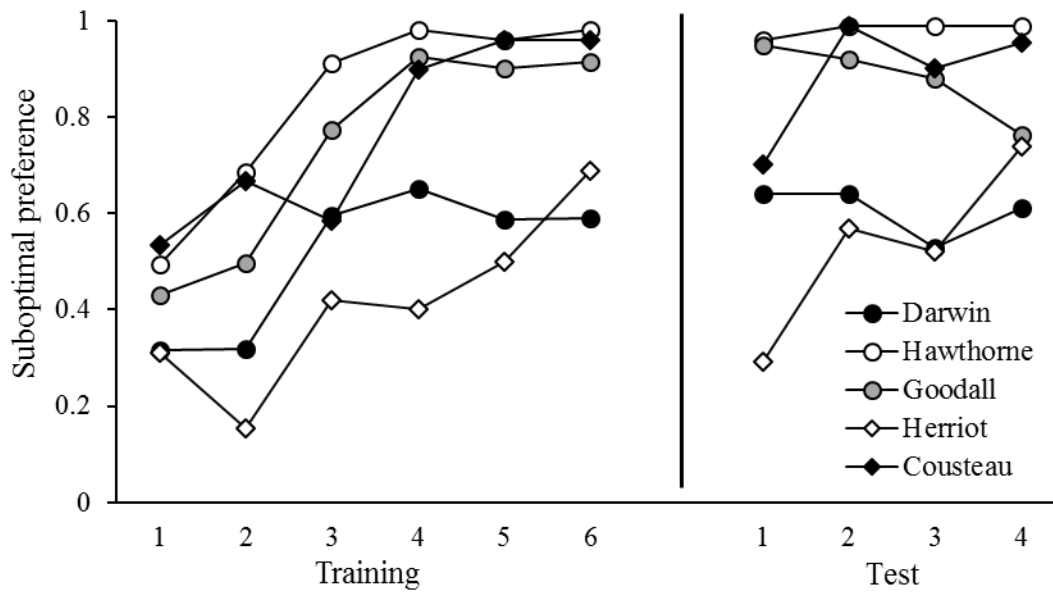


Figure 20. Proportion of suboptimal preference during training (left panel) and testing (right panel) in blocks of five sessions by pigeon in Experiment 6. As with the data analysis, for Herriot blocks 2 and 3 of training, and 1 and 2 for test were averaged for a better comparison.

Figure 20 right panel shows choice preference for the suboptimal choice during test. Given that one of the pigeons (Herriot) had an extra block of testing, the first two blocks ($M_1 = .64$ and $M_2 = .64$) were averaged to facilitate comparison with the other four pigeons. The graph suggests that preference during test blocks did not differ from the preference showed in later blocks of training. A repeated-measures ANOVA conducted on preference for suboptimal choice during test blocks using Block as a factor found a no main effect of Block, $F(1.539, 6.158) = 1.117, p = .366, \eta^2 = .218$, confirming that the preference did not change through test blocks. A Bayesian Repeated Measures ANOVA found a similar result ($BF_{01} = 2.13$); this suggests that the data provide more support for the null hypothesis.

Table 13 Table 13 shows the average RR by stimulus per block for each pigeon. There is a large difference in RR between pigeons, especially between Goodall and the other four subjects. However, we see a similar peck rate between stimuli; S+, S3 and S4 are peck similarly in each block with no systematic variation. In contrast, the RR of S- decreases for all pigeons across training. As explained in previous experiment, the RR was normalized to facilitate the comparison across pigeons. Figure 21 shows the normalized average RR to each stimulus across all blocks of training. An overall similar RR to S+, S3, and S4 was maintained across all blocks of training, while the RR to S- dropped considerably from Block

1 to Block 3 after which it remained close to 0. A Repeated-Measures ANOVA conducted on normalized RR during training with Block and Stimulus as factors found a main effect of Stimulus, $F(3, 12) = 34.933, p < .001, \eta^2 = .897$, Block, $F(5, 20) = 6.207, p < .001, \eta^2 = .608$, and their interaction, $F(15, 60) = 3.340, p < .001, \eta^2 = .455$. Post-hoc analyses using Bonferroni correction found RR to be lower to S- than to S+, S3, and S4 ($p_{bonf} < .05$), but no differences were found between the remaining three terminal link stimuli. The post-hoc analyses of Block, only showed a significant difference between Block 1 and Block 6 ($p_{bonf} = .024$). These results suggest that RR changed differently over time for the different stimuli, with RR to S- dropping across training. A paired sample t-test comparing the normalized S- RR for the first two blocks, showed a significant reduction in pecks, $t(4) = 2.963, p = .04, d = 1.325$.

Table 13

Averaged rate of response for stimuli by block for each pigeon for Experiment 6.

		Blocks of training						
		Stim	B1	B2	B3	B4	B5	B6
Darwin	S+		7.14	7.79	6.27	4.79	5.23	5.78
	S-		5.14	0.57	0.55	0.61	0.48	0.69
	S3		9.30	3.99	4.75	4.30	3.37	4.28
	S4		2.78	2.42	3.19	1.64	1.86	2.89
Goodall	S+		18.40	34.28	130.01	116.21	131.43	137.90
	S-		20.62	12.19	3.68	2.76	2.49	1.77
	S3		17.70	55.03	115.64	115.49	150.85	179.85
	S4		16.85	44.59	145.34	95.95	142.76	122.84
Hawthorne	S+		16.39	10.88	8.94	12.86	9.09	12.98
	S-		6.16	2.34	0.92	0.66	0.95	0.94
	S3		13.91	14.07	10.86	13.87	7.39	4.71
	S4		14.62	6.22	11.44	8.55	5.37	3.61
Herriot	S+		39.13	35.45	37.24	30.94	33.01	29.03
	S-		27.57	11.54	1.62	2.77	1.08	1.30
	S3		34.80	37.62	21.04	32.07	21.81	23.66
	S4		45.12	46.20	26.65	30.50	17.95	22.68
Cousteau	S+		13.47	7.76	8.06	6.54	11.32	6.89
	S-		12.28	5.74	0.54	0.34	0.13	0.17
	S3		13.74	10.29	8.21	4.67	6.23	2.79
	S4		13.72	7.77	5.82	5.57	6.76	5.16

A repeated-measures ANOVA conducted on normalized RR during test sessions with Block and Stimulus as factors found a main effect of Stimulus, $F(7, 28) = 24.394, p < .001, \eta^2 = .859$, but no effect of Block, $F(3, 12) = 2.535, p = .106, \eta^2 = .388$, nor interaction, $F(21, 84) = 0.976, p = .500, \eta^2 = .196$. A repeated-measures ANOVA focusing on the development of inhibition was conducted using S+ and S+S- as the Stimulus factor and Block, found a main effect of Block, $F(3, 12) = 6.330, p < .05, \eta^2 = .613$, Stimulus, $F(1, 4) = 31.538, p < .05, \eta^2 = .887$, and their interaction, $F(3, 12) = 6.307, p < .05, \eta^2 = .612$. Post-hoc analysis using Bonferroni corrections found that Block 1 differed from Blocks 3 and 4 ($p_{bonf} < .05$). A repeated-measures ANOVA conducted on normalized RR with S3 vs S3S1 as the Stimulus factor and Block found an effect of Stimulus, $F(1, 4) = 63.552, p < .001, \eta^2 = .941$, but not Block, $F(3, 12) < 1.0$, nor their interaction, $F(3, 12) < 1.0$.

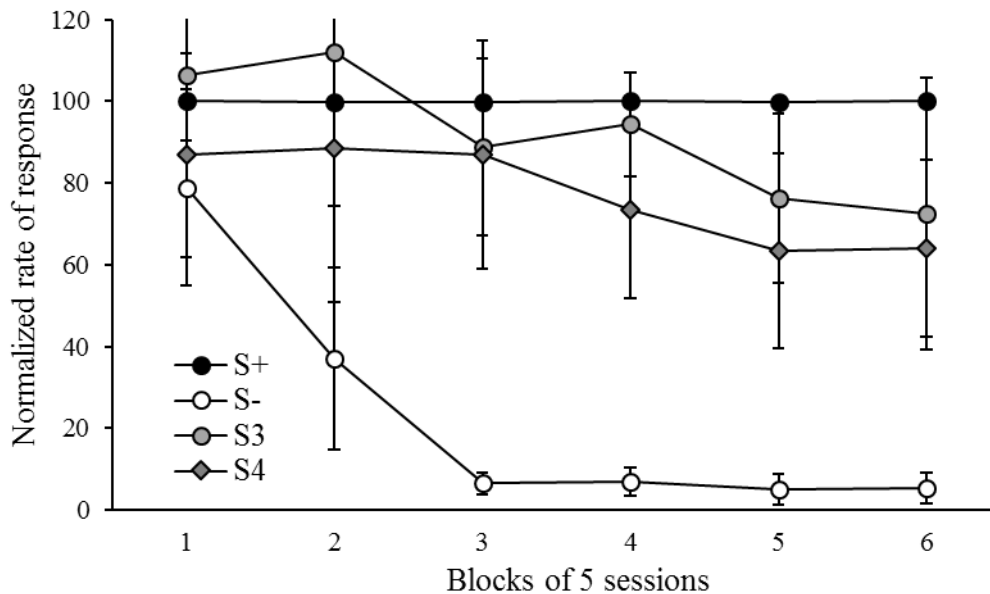


Figure 21. Mean normalized RR to each stimulus by 5-session block of training in Experiment 6. For Herriot, blocks 2 and 3 of training were averaged. Error bars correspond to the 95% confidence interval.

Figure 22 shows the average normalized response rate across all blocks of testing per stimulus. The left four bars correspond to the training stimuli while the right four bars correspond to the compound stimuli. S+ had the highest RR, followed by S3 and S4 which had a similar RR, and finally S-, the stimulus that had never been reinforced, had an RR close to zero. Visual inspection of normalized RR to the four compound stimuli suggests an order effect following the reinforcement value associated with them as well as the inhibitory properties of the S-, in which the compound S+S4 (i.e. $p = .75$, from $p = 1$ and $p = .5$,

respectively) showing a RR higher than S3S4 ($p = .5$), and S+S- ($p = .5$). Finally, S3S- ($p = .25$) showed the lowest RR.

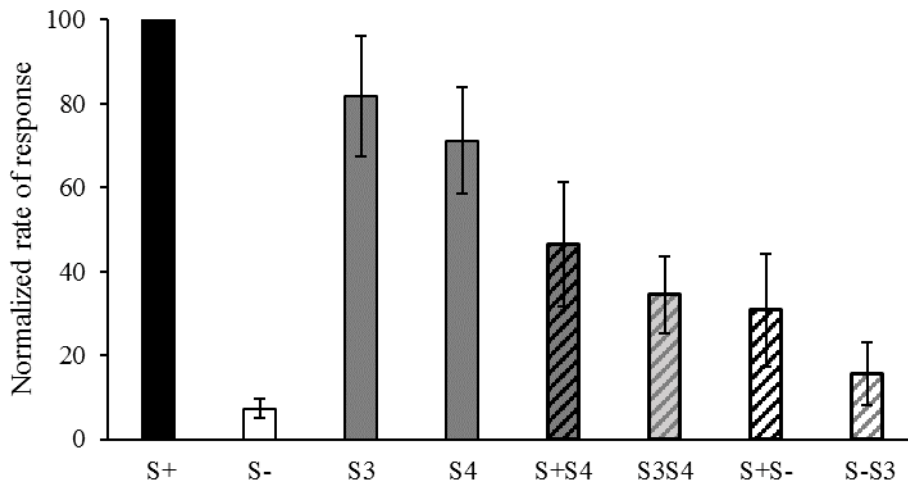


Figure 22. Mean normalized RR across all blocks of testing for each stimulus in Experiment 6. Error bars correspond to the 95% confidence interval.

Test data for normalized RR were pooled across all blocks to test specific predictions. A repeated-measures ANOVA conducted on normalized response rates with Stimulus as a factor found an effect of Stimulus, $F(2.41, 21.67) = 30.70, p < .001, \eta^2 = .773$. Planned-comparisons revealed lower normalized RR to S- than to S+, S3, and S4 ($ps < .001$), showing that, as in training, pigeons did not peck the nonreinforced stimulus. Normalized RR to S+ was higher than to S3 ($p = .022$) and S4 ($p < .001$), while normalized RR to S3 and S4 did not differ ($p = .079$). Thus, the greater the probability of food signaled by the terminal link stimulus, the greater the normalized RR. Post-hoc analysis with Bonferroni corrections comparing elements with compounds revealed that normalized RR was higher to S+ than to S+S- ($p < .001$), evidencing the inhibitory properties of the S-. Likewise, normalized RR to S3 was significantly higher than to S3S- ($p < .001$). Moreover, normalized RR to S+ was higher than to S+S4 ($p < .001$). Similarly, normalized RR to S3 and to S4 was higher than to the S3S4 test compound ($ps < .001$). Overall, RRs to compounds was lower than to elements, suggesting either generalization decrement or that animals were learning that the compounds were never reinforced.

We expected to replicate the results reported in Experiment 5 of a linear relationship between rate of reinforcement of the stimulus elements during training and peck rate to the probe trial stimulus compounds at test. Thus, we predicted the highest RR to S+S4, followed

by S3S4, S+S-, and finally S3S-. A polynomial within-subjects contrast of normalized response rates to the compound stimuli confirmed a significant linear trend, $F(1, 19) = 32.286, p < .001$, replicating the results found in Experiment 5.

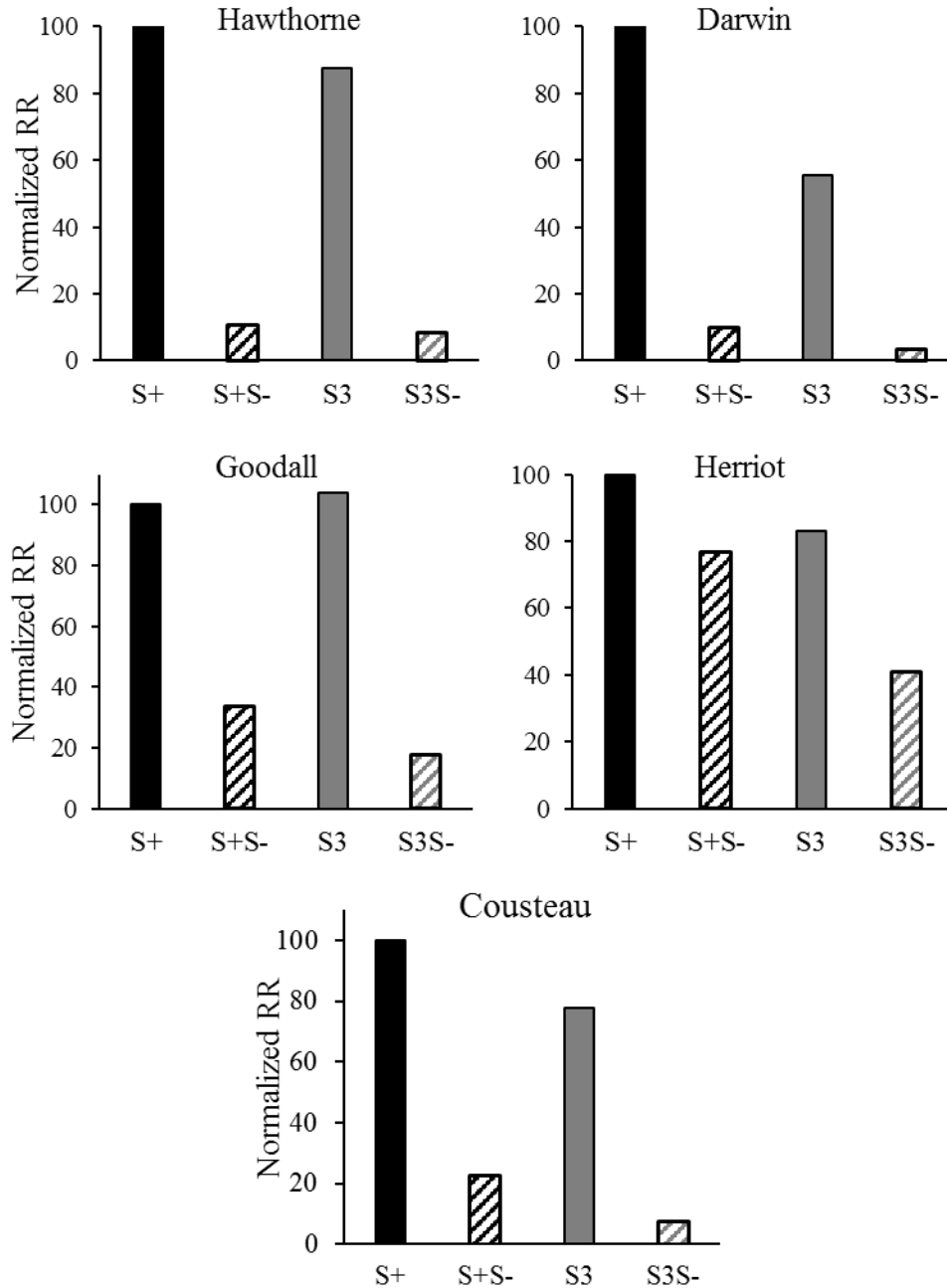


Figure 23. Normalized RR for each element and compound on probe trials for each pigeon in Experiment 6. Solid colors depict elements, black for S+ and grey for S3. Stripes depict compounds, black and white for S+S-, and grey and white for S3S-.

To further test the relationship between inhibition to S- and suboptimal preference, we performed an analysis of the summation test by bird. Figure 23 shows the average normalized

RR for S+ vs S+S- and S3 vs S3S- for each bird. As was found when data were averaged across pigeons, the individual data showed that the S+ was the most pecked stimulus, followed by S3. We also observe that RR was lower to the compounds than to the elements that was probably due to the fact that compound stimuli were never reinforced. Nevertheless, we did observe differences between compounds, the two compounds with continuous and partially reinforced stimuli (i.e. S+S4 and S3S4) elicited more pecks than the compounds containing the S- as one of the elements. Furthermore, the compound S+S- elicited a higher RR than did S3S-. A paired-sample t-test with pooled data across the four blocks of testing, comparing S+ with S+S-, and S3 with S3S- by pigeon found a significant reduction in RR to S+S- compared with S+ in five pigeons (Hawthorne: $t(3) = 33.231, p < .001, d = 16.616$; Darwin: $t(3) = 17.299, p < .001, d = 8.649$; Goodall: $t(3) = 5.293, p < .05, d = 2.647$; Herriot: $t(3) = 2.375, p < .05, d = 1.187$; and Cousteau: $t(3) = 5.648, p < .001, d = 2.864$). Likewise, Lower RR was shown to S3S- compared to S3 for four pigeons (Hawthorne: $t(3) = 5.370, p < .05, d = 2.685$; Darwin: $t(3) = 3.420, p < .05, d = 1.710$; Goodall: $t(3) = 8.533, p < .05, d = 4.266$; and Cousteau: $t(3) = 4.664, p < .05, d = 2.332$; but not Herriot: $t(3) = 1.908, p = .076, d = 0.954$).

We hypothesized that the development of inhibitory properties of S- should correlate with the level of suboptimal preference observed by each pigeon. To analyze this relationship, and avoid the RR values to be all cluttered close to 0, the S- RR were log transformed. Figure 24 24 shows the correlation between suboptimal preference and S- log RR by pigeon and average across pigeons (bottom right panel). Given that pigeons developed a suboptimal preference at different rates, we performed a Pearson's r correlation between S- RR and suboptimal preference by bird. Hawthorne, Goodall, and Cousteau showed a strong negative correlation ($r = -.95, r = -.99, \text{ and } r = -.81$, respectively; $ps < .05$). Herriot and Darwin also showed a negative but nonsignificant correlation ($r = -.77, p = .075; r = -.48, p = .337$, respectively). The three pigeons with a strong negative correlation between suboptimal preference and S- RR were also the subjects who showed inhibition in the summation test. Darwin did show inhibition in the summation test, however, its preference, although, suboptimal, was not as strong as Hawthorne, Goodall, and Cousteau. Herriot showed the weakest inhibition of the birds during testing (significant for the S+S1 compound, but not for the S3S- compound), which could explain the nonsignificant correlation between S- RR and suboptimal preference. A one-sample t-test using the Pearson's r values collapsed across all birds suggests that the mean Pearson's r was significantly more negative than zero, $t(4) = -$

8.85, $p < .001$, CI [-1.048, -.547], confirming a general negative relationship between S- RR and the level of preference for the suboptimal alternative across subjects.

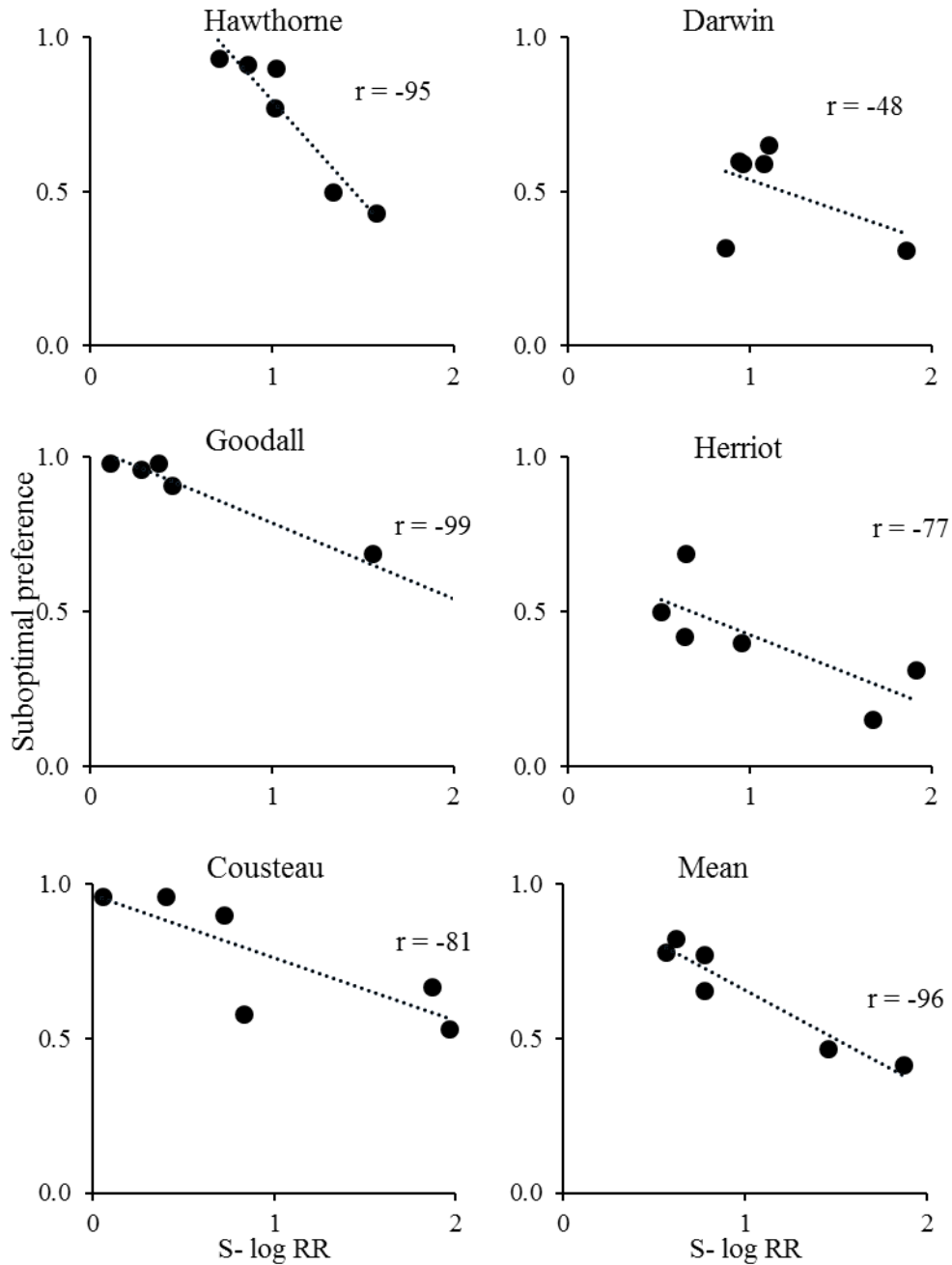


Figure 24. The correlation between suboptimal preference and log RR to S- during the 6 blocks of training for each pigeon in Experiment 6. The bottom right panel shows the mean correlation across all pigeons. Dashed lines represent the best linear fitting. r = Pearson's correlation.

General discussion

The preference observed in the suboptimal choice task is an intriguing phenomenon. The mechanisms that underlie this preference are still open to debate. The present studies tested the hypothesis that the S- has inhibitory properties, and the development of these is related to the degree of suboptimal preference. Previously, Laude et al. (2014) failed to find evidence supporting any relationship between S- inhibition and suboptimal choice, in fact showed that inhibition to S- waned with training compared with the strong inhibition shown early in training. Their result could indicate a de-correlation between inhibition and preference; however, we can only speculate because they did not report the changes in response to S- over training. Also, we identified a number of shortcomings with their study that precludes strong support of their conclusions.

In Experiment 5 of our current study, pigeons were trained only with the terminal link stimuli. Using this procedure, we found conditioned inhibition developed to S- with training, as assessed in a summation test. Moreover, we observed that the relative rate of response to each test compound correlated with the combined excitatory and inhibitory values of the elements. The compounds that contained only continuously or partially reinforced stimuli elicited higher peck rates than compounds that contained the non-reinforced S-. These results suggest that S- had acquired inhibitory properties such that it passed a summation test of conditioned inhibition. Inhibition on summation tests had a stronger effect on partially reinforced (S3S-) compared to continuously reinforced (S+S-) stimuli.

In Experiment 6, pigeons were trained on the typical suboptimal choice task and then received summation tests on compound stimuli as in Experiment 5. This allowed us to evaluate the relationship between inhibition to the S- and development of suboptimal preference. Experiment 6 replicated the linear pattern of responding to the compound test stimuli as found in Experiment 6. More interestingly, we found a negative correlation between the strength of suboptimal preference and the response rate to the S-, suggesting that as each bird acquired a suboptimal preference, the S- was becoming more inhibitory. Although these data are correlational, they could suggest a causal relation between them, perhaps with the development of preference for the suboptimal alternative initial link stimulus dependent on development of inhibition to the S- terminal stimulus. Even though none of the pigeons chose optimally by the end of training, there were individual differences

in the strength of suboptimality, and the time it took to reach asymptotic levels of suboptimal preference differed as well.

The results obtained here contradict some previous evidence that demonstrated inhibition to the S- at the outset of training, before animals had developed a suboptimal preference, and that inhibition to the S- waned with training and as animals developed suboptimal choice behavior (Laude et al., 2014). Nevertheless, the procedural differences between their study and the current one are considerable. Our study addresses some of these potentially problematic issues. First, we used the traditional probabilistic version of the suboptimal task, in which the choice is between an informative and a non-informative alternative; instead of the magnitude version of the task, in which the choice is between two informative alternatives that signal different magnitudes of reward. Second, the stimulus that served as the S- in our procedure was from the same dimension as the other terminal link stimuli, thereby avoiding potential confounds due to generalization decrement or differences in associability of the S- relative to the other stimuli. Third, the stimuli that served as the terminal links, including the S-, were counterbalanced across pigeons. Fourth, we assessed inhibition via summation tests with compounds of terminal link stimuli at various time points during training so that we could track the development of conditioned inhibition. Finally, we analyzed individual subject data to assess individual differences in learning and performance as a more sensitive test of the relationship between inhibition to the S- and suboptimal choice.

As mentioned before, Trujano et al. (2016) report an experiment evaluating inhibition in the suboptimal choice procedure with rats. They found evidence of inhibition to the S-terminal link stimulus, but in their study, rats never developed suboptimal preference. They suggested that there is a difference between how rats and pigeons learn the suboptimal choice task; and that the lack of inhibition in previous experiments with pigeons indicates that pigeons did not encode the task in the same way. Instead, rats that showed inhibition to the S- chose the optimal alternative, suggesting they were sensitive to the signal values of the terminal link stimuli. Our data contradict this conclusion, however. Moreover, a strong suboptimal preference in rats has recently been demonstrated in a task where the difference in rate of reinforcement between alternative initial link stimuli is reduced (Ojeda et al., 2018), the delay to reinforcement is increased (Cunningham & Shahan, 2019), or in which levers are used as initial link stimuli rather than lights or other visual cues (Chow et al., 2017). It would be interesting to evaluate the role of inhibition to the S- in a task in which rats develop a

preference for the suboptimal alternative. It is possible that the results of Trujano et al. (2016) could be explained by a difference in parameters, rather than an intrinsic species difference between rats and pigeons.

The results of our experiments challenge models of suboptimal choice that assumed the S- is ignored by the animal. For instance, the Temporal Information Approach (Cunningham & Shahan, 2018), developed from an associative learning perspective, claimed that the S- is not a conditioned reinforcer (Dinsmoor, 1983). However, the authors did not propose an alternative. The present study propose a mechanism. If an animal learns about a stimulus to the extent that it exerts behavioral control, such as in the negative summation test, then this stimulus should somehow contribute to choice behavior. Indeed, it is possible that the inhibitory properties of the S- indicates that the subject has learned that it signals the omission of an otherwise expected reinforcer (cf., the Rescorla-Wagner (1972) model). Differently, this idea does not contradict the RRM (Vasconcelos et al., 2015), which was developed from an ecological perspective. Although the authors also assumes the S- is ignored, the explanation is functional. The model suggests that in nature, the information about the S- is used to search food in a different patch. Hence, learning about the inhibitory properties of the S- could be necessary for the decision of foraging in a different area.

Another possibility is what the Delta-Sigma hypothesis proposed. The authors suggest that suboptimal preference is due to the interaction of two variables: the difference in probability of reinforcement within an alternative, and the ratio of the overall probability of reinforcement between alternatives. The former, also called Delta, suggests that animals prefer the alternative with a greater contrast between the signals (for instance, between S+ and S-. In terms of probability, between 1 and 0). Perhaps our data suggest that the contrast between the S+ and S- become maximum when both excitation to the S+ and inhibition to the S- develop.

Here, we propose that by acquiring properties of conditioned inhibition, the S- predicts the explicit absence of food. We further propose that the S- becoming a conditioned inhibitor is necessary for the development of suboptimal preference. Future assessments of the relationship between conditioned inhibition to the S- and development of suboptimal preference to validate our hypothesis could inform development of alternative models of suboptimal choice that include a role for processes of inhibition.

CHAPTER V

CONCLUSIONS

In the present dissertation, we used a concurrent-chain procedure in which animals have to choose between two alternatives. After the choice in the initial links, each alternative leads to one out of two possible terminal-link stimuli that remain on for a delay. When the delay elapses, the trial ends with or without reinforcement according to the programmed probabilities of reinforcement. One alternative gives food on 20% of the trials with the terminal-link stimuli associated with a probability of reinforcement equals to 1 (S+) and 0 (S-), while the other alternative gives food on half of the trials and both terminal links are associated with a probability of reinforcement of .5. Some authors called this preference *suboptimal* because the preferred alternative leads to less overall reinforcement, thus animals fail to maximize food intake.

The major goal of this dissertation was to evaluate the role of some of the key variables that have been proposed as responsible for the suboptimal choice effect, namely, the probabilities of reinforcement of the terminal links in each alternative, and the overall rate of reinforcement between alternatives. A subordinate goal was to test a newly developed hypothesis to fit the empirical data and compare it against other existing models of suboptimal choice.

Two high-order variables are responsible for the preference

Most models assume that preference for the suboptimal alternative is due to the information given by the terminal-link stimuli of that alternative: the presentation of S+ or S- indicates immediately the outcome of the trial. Indeed, most evidence suggests that, when the information is removed (i.e. the probabilities of reinforcement associated with each terminal link are set to .5), animals choose the alternative that gives more reinforcement. This suggests that animals are sensitive to the different rates of reinforcement on each alternative and that they are willing to exchange some reinforcement for information.

Although the role of information has been central in the explanation, there is no prior systematic manipulation of the information given by the terminal-link stimuli. Most literature in the suboptimal choice task used probabilities of reinforcement of 1 and 0 for the suboptimal alternative, and probability of reinforcement of .5 for the optimal alternative. Prior to the research reported here, we found two experiments that manipulated this variable.

Zentall and Stagner (2011) reduced the probability of reinforcement of S+ from 1 to .8 and continued to observe a strong suboptimal preference. Similarly, Fortes et al. (2017) increased the probability of reinforcement of the 'S-' from 0 to .375, finding a retardation but not a reduction in the final preference for the suboptimal alternative. On the other hand, some experiments that manipulated the overall reinforcement probability between alternatives have revealed either a small or no effect on preference. Some studies varied the overall probability of reinforcement in one alternative (Smith et al., 2016; Zentall et al., 2015), whereas others manipulated the ratio (Roper & Zentall, 1999; Smith et al., 2016), but none of them isolate the impact of this variable by itself in the observed preference.

In an effort towards a simpler explanation of the phenomena, in Chapter II we described a new model, the Delta-Sigma hypothesis, which proposes that the difference or contrast in the probabilities of reinforcement within the terminal links of each alternative (Delta, Δ) and the ratio of the overall probabilities of reinforcement between alternatives (Sigma, Σ), are responsible for the suboptimal choice effect. In Experiment 1, we manipulated Delta. We compared three different Deltas: 1, .5, and 0; the first and the last Delta against a Delta of .5. The results went along with the Delta-Sigma hypothesis: animals preferred the greater Delta. In Experiment 2 Sigma was manipulated. We used three different ratios (Σ/Σ): .9/.1, .7/.3 and .5/.5. The results were also consistent with the Delta-Sigma hypothesis, that is, animals always preferred the alternative given more food; but the preference was stronger when the ratio between alternatives was greater.

In light of these results, we found that the Delta-Sigma hypothesis was the only model that accounted for the results of both experiments. Other models such as Zentall's contrast-like account (Stagner & Zentall, 2010), the RRM (Fortes et al, 2018), Mazur's model (Mazur, 1995), and the Temporal Information approach (Cunningham & Shahan, 2018) failed to explain one experiment or the other. However, this does not mean that the mechanisms proposed by these models are wrong; in fact, a systematic test of these two variables (Delta and Sigma) has never been made before. We believe these results will encourage other researchers to update their models to deal with the new data and follow-up predictions.

The value of information

Chapter III reported two experiments that extended the results of Chapter II, evaluating a specific prediction that rises from the Delta-Sigma model. The model suggests that equal Deltas built with different probabilities of reinforcement are equivalent. For instance, a Delta of .5 constructed from the difference of probabilities of 1 and .5 or from .5 and 0 should have the same value.

In the two experiments reported here, Deltas of .5 built from different probabilities were all compared with a Delta of 0 (from the pair of probabilities of .5/.5). In the first experiment, the Delta of .5 was built from the following pair of probabilities: 1/.5, .75/.25, and .5/0. A similar preference for Delta of .5 was obtained in the first two conditions but not for the probabilities of .5/0, in which case the preference favored the alternative with the Delta of 0. Experiment 4 aimed to extend these results and test the hypothesis that probabilities of 1 and 0 are special conditions. To do so, two more deltas of .5 were added using the pairs of probabilities .9/.4 and .6/.1. Similarly, we obtained a strong preference for the first condition but a preference below indifference for the latter. These results suggest that the probabilities of 1 and 0 are not necessarily special conditions. The Delta-Sigma hypothesis can partially explained the results.

Overall these results are challenging. On one hand, most models of suboptimal choice have assumed that in the typical task, the S- (i.e. a probability of 0) is ignored, and in consequence, the S+ (i.e. a probability of 1) has an uneven impact on preference. The results reported here suggest that neither a probability of 1 nor 0 are different from, for instance, a probability of .9 or .1, respectively. Models such as Mazur's and the RRM had problems explaining the same conditions in which the Delta-Sigma model has difficulties. Other models, such as Zentall's contrast and the Temporal-information approach are silent and need further assumptions to deal with these results. On the other hand, the observing response literature found evidence that information about reinforcement and absence of reinforcement are not equally valuable for the animals (Dinsmoor, 1983). The results reported here also suggest that not all information is equally valuable. How this feature of information is integrated needs further examination.

The mechanism of a never reinforced stimulus

The role of S- in the suboptimal choice procedure was assessed based on the principle that most models assume that animals ignore a stimulus never reinforced. However, it is not completely clear if 'ignore' means that the value of the stimulus does not take part on the decision process, or if the animals do not learn about the stimulus. Previous research with pigeons (Laude et al., 2014), evaluated whether the S- acquires inhibitory properties. The results showed evidence of Pavlovian inhibition early in training, but this inhibition decreases by the end of the experiment. However, there were some particularities in the design that prevent us from generalizing the results.

Chapter IV, carried out two experiments in which conditioned inhibition and its relationship with the development of suboptimal choice was explored. In Experiment 5, pigeons were trained in an autoshaping procedure presenting the four terminal-link stimuli of the suboptimal choice task. Summation tests showed that S- acquired inhibitory properties during training. Experiment 6 replicated the results of the summation tests after training on the full suboptimal choice procedure. Furthermore, there was found a negative correlation between the development of suboptimal preference and the inhibitory properties of the S-, that is, suboptimal preference increased as the responses towards the s- decreased. Future experiments will need to test if there is a causal relationship between the inhibitory properties and the suboptimal preference, in other words, if the development of inhibition is necessary to show suboptimal preference.

Most models of suboptimal choice need additional assumptions to explain how learning about the S- does not contribute to the decision. For instance, the Temporal-Information approach comes from an associative learning tradition and perhaps would need to integrate these results on their assumptions. On the other hand, Zentall's like-contrast account and the Delta-Sigma model do not assume the S- is ignored, consequently, the results reported here do not contradict their explanations. Lastly, the results presented here do not contradict the explanation offered by ecological or functional models such as the RRM. Instead, they can represent an opportunity to integrate different perspectives.

Overall, the results presented in this dissertation found that animals are sensitive to information. Furthermore, this information is about food, thus the overall probability of food given by each option is also relevant and contributes to the decision. However, how

information from different sources is integrated is still an unsolved question. Finally, animals learn about the S-, and the strength of this learning correlates with the acquisition of the suboptimal preference. These results indicate that animals are sensitive to the proposed variables. It is well-known that the suboptimal choice preference is a reliable phenomenon that occurs in a broad variety of conditions. However, more research is needed focusing on the mechanisms behind the effect. The work presented here is an effort in that direction.

REFERENCES

- Beyeler, A., Namburi, P., Glober, G. F., Simonnet, C., Calhoun, G. G., Conyers, G. F., ... Tye, K. M. (2016). Divergent Routing of Positive and Negative Information from the Amygdala during Memory Retrieval. *Neuron*, *90*(2), 348–361. <https://doi.org/10.1016/j.neuron.2016.03.004>
- Blaisdell, A. P., & Cook, R. G. (2005). Two-item same-different concept learning in pigeons. *Learning and Behavior*, *33*(1), 67–77. <https://doi.org/10.3758/bf03196051>
- Blanchard, R. (1975). The effect of S- on observing behavior. *Learning and Motivation*, *6*(1), 1–10. [https://doi.org/10.1016/0023-9690\(75\)90031-4](https://doi.org/10.1016/0023-9690(75)90031-4)
- Bower, G., Mclean, J., & Meacham, J. (1966). Value of Knowing When Reinforcement Is Due. *Journal of Comparative and Physiological Psychology*, *62*(2), 184–192. <https://doi.org/10.1037/h0023682>
- Browne, M. P., & Dinsmoor, J. A. (1974). Wyckoff's observing response: Pigeons learn to observe stimuli for free food but not stimuli for extinction. *Learning and Motivation*, *5*(2), 165–173. [https://doi.org/10.1016/0023-9690\(74\)90023-X](https://doi.org/10.1016/0023-9690(74)90023-X)
- Case, J. P., & Zentall, T. R. (2018). Suboptimal choice in pigeons: Does the predictive value of the conditioned reinforcer alone determine choice? *Behavioural Processes*, *157*(July), 320–326. <https://doi.org/10.1016/j.beproc.2018.07.018>
- Chow, J. J., Smith, A. P., Wilson, A. G., Zentall, T. R., & Beckmann, J. S. (2017). Suboptimal choice in rats: Incentive salience attribution promotes maladaptive decision-making. *Behavioural Brain Research*, *320*, 244–254. <https://doi.org/10.1016/j.bbr.2016.12.013>
- Cunningham, P. J., & Shahan, T. A. (2018). Suboptimal choice, reward-predictive signals, and temporal information. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(1), 1–22. <https://doi.org/10.1037/xan0000160>
- Cunningham, P. J., & Shahan, T. A. (2019). Rats Engage in Suboptimal Choice When the Delay to Food Is Sufficiently Long. *Journal of Experimental Psychology: Animal Learning and Cognition*. <https://doi.org/10.1037/xan0000211>
- Daniels, C. W., & Sanabria, F. (2018). An associability decay model of paradoxical choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(3), 258–271. <https://doi.org/10.1037/xan0000179>
- Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, *6*(4), 693–704. <https://doi.org/10.1017/S0140525X00017969>
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior*, *53*(2), 201–218. <https://doi.org/10.1901/jeab.1990.53-201>
- Fantino, E., Dunn, R., & Meck, W. (1979). Percentage reinforcement and choice. *Journal of the Experimental Analysis of Behavior*, *32*(3), 335–340. <https://doi.org/10.1901/jeab.1979.32-335>
- Fantino, E., & Silberberg, A. (2010). Revisiting the role of bad news in maintaining human

- observing behavior. *Journal of the Experimental Analysis of Behavior*, 93(2), 157–170. <https://doi.org/10.1901/jeab.2010.93-157>
- Fortes, I., Machado, A., & 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(4), 277–289. <https://doi.org/10.1037/com0000079>
- Fortes, I., Pinto, C., Machado, A., & Vasconcelos, M. (2018). The paradoxical effect of low reward probabilities in suboptimal choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44(2), 180–193. <https://doi.org/10.1037/xan0000165>
- Fortes, I., Vasconcelos, M., & Machado, A. (2016). Testing the Boundaries of “Paradoxical” Predictions: Pigeons Do Disregard Bad News. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42(4), 336–346. <https://doi.org/10.1037/xan0000114>
- Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, 107(2), 289–344. <https://doi.org/10.1037/0033-295X.107.2.289>
- Gipson, C. D., Alessandri, J. J. D., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning & Behavior*, 37(4), 289–298. <https://doi.org/10.3758/LB.37.4.289>
- Hearst, E., Bottjer, S. W., & Walker, E. (1980). Conditioned approach-withdrawal behavior and some signal-food relations in pigeons: Performance and positive vs. negative “associative strength”. *Bulletin of the Psychonomic Society*, 16(3), 183–186. <https://doi.org/10.3758/BF03329516>
- Hirota, T. T. (1972). The Wyckoff Observing Response- A Reappraisal. *Journal of the Experimental Analysis of Behavior*, 18(2), 263–276.
- 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(3), 547–558. <https://doi.org/10.1007/s00265-010-1101-2>
- Kendall, S. B. (1974). Preference for Intermittent Reinforcement. *Journal of the Experimental Analysis of Behavior*, 21(3), 463–473. <https://doi.org/10.1901/jeab.1974.21-463>
- Lalli, J. S., Mauro, B. C., & Mace, F. C. (2000). Preference for unreliable reinforcement in children with mental retardation: the role of conditioned reinforcement. *Journal of Applied Behavior Analysis*, 33(4), 533–544. <https://doi.org/10.1901/jaba.2000.33-533>
- Laude, J. R., Beckmann, J. S., Daniels, C. W., & Zentall, T. R. (2014). Impulsivity affects suboptimal gambling-like choice by pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(1). <https://doi.org/10.1037/xan0000001>
- Laude, J. R., Stagner, J. P., & Zentall, T. R. (2014). Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(1), 12–21. <https://doi.org/10.1037/xan0000010>

- Lieberman, D. A., Cathro, J. S., Nichol, K., & Watson, E. (1997). The role of S-in human observing behavior: Bad news is sometimes better than no news. *Learning and Motivation*, 28(1), 20–42. <https://doi.org/10.1006/lmot.1997.0951>
- Ludvig, E. A., Madan, C. R., McMillan, N., Xu, Y., & Spetch, M. L. (2018). Living near the edge: How extreme outcomes and their neighbors drive risky choice. *Journal of Experimental Psychology: General*, 147(12), 1905–1918. <https://doi.org/10.1037/xge0000414>
- Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review*, 97(2), 225–252. <https://doi.org/10.1037/0033-295X.97.2.225>
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal Behavior Processes*, 10(4), 426–436. <https://doi.org/10.1037/0097-7403.10.4.426>
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In: *Commons, M.L., Mazur, J.E., Nevin, J.A., Rachlin, H., pp, 55–73*. Retrieved from https://books.google.com/books?hl=es&lr=&id=1q5mAgAAQBAJ&oi=fnd&pg=PA55&ots=eNqEXNxb1t&sig=LZ-72k6AwVwY_StT_zTS_wSoxXM
- Mazur, J. E. (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior*, 51(1), 87–99. <https://doi.org/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(2), 139–150. <https://doi.org/10.1901/jeab.1995.63-139>
- Mazur, J. E. (1996). Choice with certain and uncertain reinforcers in an adjusting-delay procedure. *Journal of the Experimental Analysis of Behavior*, 66(1), 63–73. <https://doi.org/10.1901/jeab.1996.66-63>
- McDevitt, M. A., Spetch, M. L., & Dunn, R. (1997). Contiguity and conditioned reinforcement in probabilistic choice. *Journal of the Experimental Analysis of Behavior*, 68(3), 317–327. <https://doi.org/10.1901/jeab.1997.68-317>
- McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*. <https://doi.org/10.1002/jeab.192>
- Molet, M., Miller, H. C., Laude, J. R., Kirk, C., Manning, B., & Zentall, T. R. (2012). Decision making by humans in a behavioral task: do humans, like pigeons, show suboptimal choice? *Learning & Behavior*, 40(4), 439–447. <https://doi.org/10.3758/s13420-012-0065-7>
- Ojeda, A., Murphy, R. A., & Kacelnik, A. (2018). Paradoxical choice in rats: Subjective valuation and mechanism of choice. *Behavioural Processes*, 152(March), 73–80. <https://doi.org/10.1016/j.beproc.2018.03.024>
- Olsen, D. G., & Pracejus, J. W. (2004). Integration of positive and negative affective stimuli. *Journal of Consumer Psychology*, 14(4), 374–384.

- https://doi.org/10.1207/s15327663jcp1404_7
- Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Pisklak, J. M., Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2019). The power of nothing: Risk preference in pigeons, but not people, is driven primarily by avoidance of zero outcomes, 1689–1699. <https://doi.org/10.1017/CBO9781107415324.004>
- Prokasy, W. F. (1956). The acquisition of observing responses in the absence of differential external reinforcement. *Journal of Comparative and Physiological Psychology*, 49(2), 131–134. <https://doi.org/10.1037/h0046740>
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*. <https://doi.org/10.1037/h0027760>
- Rescorla, R., & Wagner, A. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64–99). New York: Appleton-Century-Crofts. Retrieved from papers3://publication/uuid/2DC3F0FC-E35D-4C8F-8F24-72997FB92D83
- Roper, K. L., & 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(3), 201–220. <https://doi.org/10.1006/lmot.1999.1030>
- Shahan, T. A., & Cunningham, P. (2015). Conditioned reinforcement and information theory reconsidered. *Journal of the Experimental Analysis of Behavior*, 103(2), 405–418. <https://doi.org/10.1002/jeab.142>
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. The University of Illinois Press (Vol. 25). Chicago, IL. <https://doi.org/10.4992/jjpsy.25.110>
- Shapiro, M. S., Siller, S., & Kacelnik, A. (2008). Simultaneous and sequential choice as a function of reward delay and magnitude: Normative, descriptive and process-based models tested in the European starling (*Sturnus vulgaris*). *Journal of Experimental Psychology: Animal Behavior Processes*, 34(1), 75–93. <https://doi.org/10.1037/0097-7403.34.1.75>
- Sharot, T., & Garrett, N. (2016). Forming Beliefs: Why Valence Matters. *Trends in Cognitive Sciences*, 20(1), 25–33. <https://doi.org/10.1016/j.tics.2015.11.002>
- Shaw, M. L. (1982). Attending to multiple sources of information: I. The integration of information in decision making. *Cognitive Psychology*, 14(3), 353–409. [https://doi.org/10.1016/0010-0285\(82\)90014-7](https://doi.org/10.1016/0010-0285(82)90014-7)
- Singer, R. A., Berry, L. M., & Zentall, T. R. (2007). Preference for a stimulus that follows a relatively aversive event: contrast or delay reduction? *Journal of the Experimental Analysis of Behavior*, 87(2), 275–285. <https://doi.org/10.1901/jeab.2007.39-06>
- Smith, A. P., Bailey, A. R., Chow, J. J., Beckmann, J. S., & Zentall, T. R. (2016). Suboptimal

- choice in pigeons: Stimulus value predicts choice over frequencies. *PLoS ONE*, *11*(7), 1–18. <https://doi.org/10.1371/journal.pone.0159336>
- Smith, A. P., & 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*(2), 212–220. <https://doi.org/10.1037/xan0000092>
- Spetch, M. L., Belke, T. W., Barnet, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior*, *53*(2), 219–234. <https://doi.org/10.1901/jeab.1990.53-219>
- Stagner, J. P., Laude, J. R., & 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*(4), 282–287. <https://doi.org/10.1016/j.lmot.2011.09.001>
- Stagner, J. P., Laude, J. R., & 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*(4), 446–452. <https://doi.org/10.1037/a0030321>
- Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review*, *17*(3), 412–416. <https://doi.org/10.3758/PBR.17.3.412>
- Trujano, R. E., López, P., Rojas-Leguizamón, M., & Orduña, V. (2016). Optimal behavior by rats in a choice task is associated to a persistent conditioned inhibition effect. *Behavioural Processes*, *130*, 65–70. <https://doi.org/10.1016/j.beproc.2016.07.005>
- Trujano, R. E., & Orduña, V. (2015). Rats are optimal in a choice task in which pigeons are not. *Behavioural Processes*, *119*, 22–27. <https://doi.org/10.1016/j.beproc.2015.07.010>
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. <https://doi.org/10.1007/BF00122574>
- Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific Reports*, *5*, 13874. <https://doi.org/10.1038/srep13874>
- Vasconcelos, M., Machado, A., & Pandeirada, J. N. S. (2018). Ultimate explanations and suboptimal choice. *Behavioural Processes*. <https://doi.org/10.1016/j.beproc.2018.03.023>
- Wasserman, E. A., Franklin, S. R., & Hearst, E. (1974). Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of Comparative and Physiological Psychology*, *86*(4), 616–627. <https://doi.org/10.1037/h0036171>
- Wyckoff, L. B. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review*, *59*(6), 431–442. <https://doi.org/10.1037/h0053932>
- Wyckoff, L. B. (1969). The role of observing responses in discrimination learning. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 237–260). Homewood, IL: Dorsey Press.

- Zentall, T. R. (2005). A Within-trial Contrast Effect and its Implications for Several Social Psychological Phenomena. *International Journal of Comparative Psychology*, *18*(4), 273–297. Retrieved from <https://escholarship.org/uc/item/1476b0fg>
- Zentall, T. R. (2015). When animals misbehave: Analogs of human biases and suboptimal choice. *Behavioural Processes*, *112*, 3–13. <https://doi.org/10.1016/j.beproc.2014.08.001>
- Zentall, T. R. (2016a). Resolving the paradox of suboptimal choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*(1), 1–14. <https://doi.org/10.1037/xan0000085>
- Zentall, T. R. (2016b). When Humans and Other Animals Behave Irrationally. *Comparative Cognition & Behavior Reviews*, *11*(January), 25–48. <https://doi.org/10.3819/ccbr.2016.110002>
- Zentall, T. R., Andrews, D. M., & Case, J. P. (2017). Prior commitment: Its effect on suboptimal choice in a gambling-like task. *Behavioural Processes*, *145*(September), 1–9. <https://doi.org/10.1016/j.beproc.2017.09.008>
- Zentall, T. R., Laude, J. R., Stagner, J. P., & 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*(2), 223–229. <https://doi.org/10.1007/s40732-015-0119-2>
- Zentall, T. R., & Stagner, J. P. (2011). Sub-optimal choice by pigeons: Failure to support the Allais paradox. *Learning and Motivation*, *42*(3), 245–254. <https://doi.org/10.1016/j.lmot.2011.03.002>