



**Universidade do Minho**  
Escola de Psicologia

**Maria Inês Abreu Fortes**

**O efeito do esforço no valor de uma recompensa**

**Dissertação de Mestrado**

**Mestrado Integrado em Psicologia**

**Área de especialização em Psicologia Experimental e Suas  
Aplicações**

**Trabalho efectuado sob a orientação do  
Professora Armando Machado**

**Outubro de 2011**

## DECLARAÇÃO

Nome:

Maria Inês Abreu Fortes

Endereço electrónico: ines.fortes@gmail.com

Título da dissertação:

O efeito do esforço no valor de uma recompensa.

Orientador:

Professor Doutor Armando Machado

Ano de conclusão: 2011

Designação do Mestrado:

Mestrado Integrado em Psicologia – Área de Especialização de Psicologia Experimental e Suas Aplicações

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, 19/10/2011

Assinatura: \_\_\_\_\_

## **O efeito do esforço no valor de uma recompensa**

Este estudo teve como objectivo compreender como é que o esforço para obter uma recompensa afecta o seu valor. Na Experiência 1a, seis pombos escolheram entre duas opções, 2s de reforço entregue após um atraso de 20s ou 6s de reforço entregue após um atraso ajustável. Este atraso ajustável aumentava ou diminuía dependendo das escolhas do pombo. Pretendíamos assim ajustar este intervalo para encontrar o ponto de indiferença - duração do atraso com a qual as duas alternativas são igualmente escolhidas. De seguida variou-se o número de bicadas que o pombo tinha que dar para obter os 2s de reforço e verificou-se qual o efeito destas bicadas sobre os pontos de indiferença. Assim, podiam ser exigidas entre 1 e 10 bicadas durante o atraso (fase 'Taxa Baixa'), entre 10 e 30 bicadas (fase 'Taxa Intermédia') ou pelo menos 30 bicadas (fase 'Taxa Elevada'). Desde modo, com base nos pontos de indiferença, pretendemos verificar se a exigência de respostas durante o atraso leva a escolhas mais impulsivas (da pequena recompensa) ou autocontroladas (da grande recompensa). Os pombos mostraram dificuldade em aprender a tarefa, e por isso o procedimento foi simplificado. Na Experiência 1b o atraso para o reforço de 2s foi diminuído para 10s e houve apenas duas fases. Na fase 'Taxa Baixa', durante o atraso de 10s, o pombo não podia dar mais de 5 respostas e na fase 'Taxa Elevada' tinha de dar pelo menos 10 respostas. Os resultados mostram que, para a maior parte dos pombos, exigir respostas durante o atraso para uma recompensa, desvaloriza essa recompensa.

## **The effect of effort on the value of a reward**

This study aimed to understand how the effort to obtain a reward affects its value. In Experiment 1a, six pigeons chose between two options, 2s of reinforcer delayed 20s or 6s of reinforcer delivered after an adjustable delay. This adjustable delay increased or decreased depending on the pigeon's choices. We wanted to adjust this interval to find the indifference point – the delay duration at which the two alternatives are equally chosen. We then varied the number of pecks required for the 2-s food and assessed how these pecks affected the indifference points. Thus, it could be required between 1 and 10 pecks during the delay (phase 'Low Rate'), between 10 and 30 pecks (phase 'Intermediate Rate') or at least 30 pecks (phase 'High Rate'). Based on the indifference points, we assessed whether requiring responses during the delay led to more impulsive (small reward) or self-controlled (large reward) choices. The pigeons showed difficulties learning the task, so the procedure was simplified. In Experiment 1b the 2-s reinforcer delay was reduced to 10s and there were only two phases. In the phase 'Low Rate' during the 10-s delay the pigeon could not give more than 5 phase responses and in phase 'High Rate' it had to give at least 10 responses. The results showed that, for most of the pigeons, requiring pecks during the delay devalues the reward.

## ÍNDICE

Introdução	6
Experiência 1a	
Método	15
Resultados e Discussão	21
Experiência 1b	
Método	27
Resultados e Discussão	31
Conclusão	46
Referências	48

## INTRODUÇÃO

Todos os animais, quer humanos, quer não humanos, têm frequentemente que tomar decisões: para onde ir, o que comer, que actividade desempenhar em cada momento. Por estar tão frequentemente presente na vida de qualquer ser, o estudo do processo da tomada de decisão e do seu resultado, a escolha, é importante. Um dos paradigmas mais usados para estudar a escolha é o do autocontrolo. Numa situação de autocontrolo, os sujeitos são confrontados com uma escolha entre uma recompensa (ou reforço) pequena disponível imediatamente (ou após um atraso curto), e uma recompensa maior disponível após um atraso mais longo. Tipicamente, após várias escolhas, o ganho total de reforços é maximizado se for escolhido sempre o reforço maior e mais atrasado. Esta escolha é definida como autocontrolo (Rachlin & Green, 1972). Pelo contrário, a escolha do reforço menor e mais imediato, é nomeado de impulsividade.

Nos estudos de autocontrolo parte-se do pressuposto de que a escolha do sujeito depende da relação entre o atraso para a obtenção um reforço e o valor (subjectivo) desse mesmo reforço. Para estudar esta relação entre atraso e valor de um reforço, Mazur (1984; 1987) desenvolveu um procedimento com pombos, o procedimento de ajustamento do atraso.

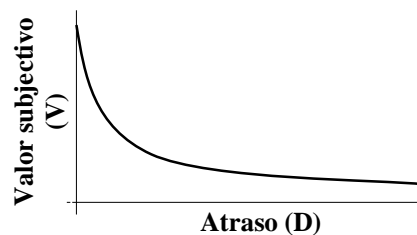
Neste procedimento o sujeito tem de escolher entre duas alternativas: a *alternativa padrão* e a *alternativa ajustável*. Geralmente a alternativa padrão corresponde ao reforço menor e menos atrasado e a alternativa ajustável corresponde ao reforço maior mas mais atrasado. Para a *alternativa padrão*, as contingências (atraso e quantidade/duração de reforço) são mantidas constantes durante uma determinada condição. Por exemplo, imaginemos que na condição 1 a *alternativa padrão* consiste sempre num reforço de 2s atrasado 10s. Nesta mesma condição, para a *alternativa ajustável*, a duração do reforço é constante (por exemplo 6s) e o seu atraso é constantemente variado consoante as escolhas do sujeito. Por exemplo, o atraso da *alternativa ajustável* pode começar com 10s (igual ao atraso da alternativa padrão). De seguida, se o pombo mostra preferência pela *alternativa ajustável* (porque tem o mesmo atraso que a *alternativa padrão* mas dá mais comida), o atraso da *alternativa ajustável* é aumentado. Este aumento do atraso da *alternativa ajustável* faz com que o valor desta alternativa diminua, ficando mais próximo do valor da *alternativa padrão*. Imaginemos que após alguns aumentos do atraso, o atraso da *alternativa ajustável* é de 40s. O pombo tem de escolher entre 2s de comida atrasados 10s (*alternativa padrão*) e 6s de comida atrasados 40s (*alternativa ajustável*). Se agora o pombo mostra preferência pela *alternativa padrão* (porque apesar de dar menos comida, é menos atrasado), o atraso da *alternativa ajustável* é diminuído. Esta diminuição do atraso da *alternativa ajustável* faz com que o valor desta alternativa aumente, já que um reforço menos atrasado tem mais valor que um mais atrasado. Assim, o valor da *alternativa ajustável* aproxima-se mais uma vez do valor da *alternativa padrão*. Com este ajustamento, o *atraso ajustável* acaba por estabilizar numa duração com a qual o sujeito não mostra preferência por nenhuma das alternativas – o ponto de

indiferença. Imaginemos que nessa primeira condição o pombo mostra indiferença entre 2s de reforço atrasado 10s (*alternativa padrão*) e 6s de reforço atrasado 30s (*alternativa ajustável*). Para esta condição diz-se que o ponto de indiferença é de 30s. Imaginemos ainda que numa segunda condição a *alternativa padrão* é um reforço de 2s mas agora atrasado 20s. Aplicando o mesmo procedimento, pode-se encontrar o ponto de indiferença na segunda condição e descobrir que é, por exemplo, 50s. Assim, em diferentes condições, o procedimento pode ser repetido para outros valores do atraso padrão e para cada atraso padrão encontra-se um ponto de indiferença correspondente.

A partir dos resultados obtidos com este procedimento, Mazur (1984; 1987) descobriu que a relação entre o atraso para um reforço e o seu valor subjectivo era bem descrita por uma função de desconto hiperbólica:

$$V = \frac{A}{1+KD} \quad (\text{Equação 1})$$

em que V é o valor subjectivo do reforço, A está relacionado com a magnitude do reforço, K com a taxa de desconto e D é o atraso entre a resposta de escolha e o reforço (ver figura 1).



**Figura 1** – Função de desconto hiperbólica. V- valor subjectivo do reforço; D- atraso para reforço.

No exemplo anterior, pode-se considerar o atraso para um reforço como um custo; quanto maior o atraso, maior o custo para a obtenção do reforço. Assim, à medida que o custo para um reforço aumenta, o valor subjectivo desse reforço diminui. A esta diminuição do valor subjectivo de um reforço com o aumento do seu custo chama-se desconto. Os estudos de desconto têm-se centrado em três tipos de custo: o atraso (desconto do atraso), a probabilidade (desconto probabilístico) e o esforço (desconto de esforço). De um modo geral, estes estudos têm verificado que quanto maior o atraso, menor a probabilidade ou maior o esforço para um reforço, menor o valor desse reforço.

No mundo natural, os três tipos de custo aparecem frequentemente correlacionados: por exemplo, quanto maior o esforço dispendido na obtenção de uma recompensa, maior o tempo de espera para essa recompensa e também maior é a probabilidade de se receber essa mesma recompensa. Assim, a investigação laboratorial tem procurado separar o impacto de cada um destes tipos de custo na tomada de decisão. Porém, se os estudos de desconto do atraso e probabilístico são abundantes e estão em crescimento, a investigação de como a exigência de esforço afecta o valor de um reforço é

escassa. Este facto é de algum modo surpreendente já que a maioria das escolhas envolve algum tipo de esforço.

Em meados da década de 80, James E. Mazur começou uma série de estudos (ex. Mazur, 1986; Grossbard & Mazur, 1986; Mazur & Kralik, 1990) com pombos com o objectivo de testar as previsões da equação 1 em diferentes situações experimentais. Deste modo, o autor pretendia encontrar uma regra de equivalência que tornaria possível, por exemplo, prever a preferência entre um programa de reforço de tempo fixo (FT- do inglês, *Fixed Time*), em que o reforço é entregue um certo tempo depois da escolha, e um programa de razão fixa (FR- do inglês, *Fixed Ratio*) em que o reforço é entregue após um número fixo de respostas, sendo esse número o parâmetro do FR. Esta regra de equivalência permitiria, por exemplo, prever qual a preferência entre um FT10 com 3s de reforço e um FR10 com 3s de reforço. Assim, o autor procurou compreender qual o efeito da exigência de respostas (i.e., esforço) no valor de um reforço. Porém, ao conceptualizar o esforço pelo número de respostas a emitir, quanto maior é o número de respostas, maior é o atraso para o reforço. Assim, como se verá de seguida, tornava-se difícil distinguir entre os efeitos do esforço e do tempo no valor de um reforço. O principal objectivo do presente estudo foi verificar qual o efeito da exigência de respostas para a obtenção de um reforço, comparando situações em que o tempo para o reforço era o mesmo.

Um dos primeiros estudos (Mazur, 1986) teve como objectivo determinar se a equação 1 se poderia aplicar a programas de razão, usando como D (ver equação 1) o tempo para completar um FR. Porém, a aplicação da equação a programas de razão implica pelo menos dois problemas. Em primeiro, num programa de razão o sujeito tem de dar um certo número de respostas antes do reforço ser entregue, logo, o tempo entre a resposta de escolha e a entrega do reforço é controlado pelo sujeito e não pelo experimentador. Assim, variações na taxa de resposta levam a variações no atraso para o reforço. Em segundo lugar, se durante um atraso o sujeito tem apenas de esperar pelo reforço, num programa de razão são exigidas respostas, e não é muito claro de que modo este factor de esforço afecta o valor do reforço, e de como deve ser incluído na equação de desconto.

Para verificar de que modo o esforço afectava o valor de uma alternativa, em algumas condições deste estudo (Mazur, 1986) a alternativa padrão era um FT (10s, 20s, 30s ou 40s) e a alternativa ajustável era um FR, sendo ambas reforçadas com 3s de acesso a comida (ver tabela 1). Para cada duração do FT da alternativa padrão, o parâmetro do FR era ajustado segundo o procedimento de ajustamento anteriormente descrito (mas adaptado, ajustando-se o número de respostas e não o tempo para o reforço).



**Tabela 1** – Resumo do procedimento nas experiências de Mazur (1986); Grossbard & Mazur (1986) e Mazur e Kralik (1990).

Experiência	Fase	Alternativa padrão		Alternativa ajustável	
		Programa	Duração reforço	Programa	Duração reforço
Mazur (1986)	FT-FR	FT	3s	FR	3s
Grossbard & Mazur (1986)	FR-FR	FR		FR	
	FT-FT	FT	2s	FT	6s
	FT-FR	FT		FR	
Mazur e Kralik (1990)	FR-FT	FR	3s	FT	3s
	VT-FT	VT		FT	

Suponhamos que a condição em que a alternativa padrão era um FT10s e que no FR da alternativa ajustável o pombo responde a uma taxa média de 1 resposta por segundo. Se a exigência de respostas na alternativa ajustável não tem efeito na escolha a não ser pela imposição de um atraso entre a escolha e o reforço, (como sugerido por Neuringer & Schneider, 1968), o ponto de indiferença deveria ser um FR10 (que demora aproximadamente 10s a ser completado).

Por outro lado, se a exigência de respostas diminui o valor da alternativa ajustável, numa escolha entre FT10s e FR10 o pombo já não seria indiferente, preferindo o FT10s. Assim, pelo procedimento de ajustamento, no ponto de indiferença o parâmetro do FR deveria ser um que levasse menos de 10s a ser completado.

Os resultados mostraram que, no ponto de indiferença, o tempo para completar o FR era menor que a duração do FT da alternativa padrão. Apesar de existir grande variabilidade entre sujeitos, nenhum pombo apresentou pontos de indiferença com valor superior ao FT da alternativa padrão. Ou seja, os pombos mostraram uma preferência pelo programa FT em detrimento do programa FR.

Para incluir o esforço na equação de desconto, Mazur (1986) considerou que o efeito das respostas no valor de um reforço podia-se reflectir numa taxa de desconto maior ( $K_{FR}$ ), quando comparada com a taxa de desconto durante um atraso simples ( $K_{FT}$ ). Assumindo que no ponto de indiferença  $V_{FR}=V_{FT}$  (os subscritos FR e FT referem-se às alternativas ajustável e padrão, respectivamente), a partir da equação 1 verifica-se que

$$\frac{A_{FR}}{1+K_{FR}D_{FR}} = \frac{A_{FT}}{1+K_{FT}D_{FT}} \quad \text{e como } A_{FR}=A_{FT}, \text{ segue-se}$$

$$D_{FR} = \frac{K_{FT}}{K_{FR}} D_{FT} \quad \text{(Equação 2)}$$

Usando a equação 2 em que  $D_{FR}$  era o tempo para completar o FR e assumindo  $K_{FT}$  igual a 1, o autor obteve  $K_{FR}$  igual a 1.35. Assim, de acordo com esta equação, a preferência dos pombos por atrasos simples (FT) dever-se-á ao desconto do FR ser mais rápido que o do FT. Dito de outra forma, um reforço obtido após algumas respostas desvaloriza mais rapidamente que um reforço obtido após um simples atraso.

Num estudo subsequente, Grossbard e Mazur (1986) procuraram compreender melhor se a função hiperbólica de desconto podia ser aplicada a programas de FR. O seu principal objectivo era verificar se a equação 1 se poderia aplicar a programas de razão, usando como D (ver equação 1) o parâmetro do FR e não o tempo para completar um FR como no estudo anterior (Mazur, 1986). Para além desta diferença no objectivo experimental, o estudo de Grossbard e Mazur (1986) usou alternativas com diferentes magnitudes de reforço (alternativa padrão - 2s de reforço; alternativa ajustável - 6s de reforço). De seguida descrevem-se as três fases da experiência (ver tabela 1).

Na primeira fase (fase FR-FR) os pombos escolheram entre a alternativa padrão (programa FR com 2s de reforço) e alternativa ajustável (programa FR com 6s de reforço). O parâmetro do FR para a alternativa padrão variou entre condições, e para cada parâmetro foi encontrado um ponto de indiferença.

Os resultados obtidos mostraram que, à medida que o parâmetro do FR na alternativa padrão aumentava, o ponto de indiferença também aumentava. Isto é, quanto maior o parâmetro do FR, menor o valor do reforço obtido. Os pontos de indiferença obtidos foram bem descritos pela equação 1, utilizando como D quer o parâmetro do FR quer o tempo para completar o FR, sugerindo que a mesma função que descreve a relação entre o tempo para um reforço e o valor desse reforço pode ser também aplicada com sucesso para descrever a relação entre o número de respostas necessárias para a obtenção de um reforço e o valor desse reforço. Assim, obteve-se a primeira evidência de que a função de desconto do atraso é semelhante à função de desconto de esforço.

Para a segunda fase (fase FT-FT), em vez de programas FR, usaram-se programas FT. Isto permitiu comparar as escolhas envolvendo dois programas FR e as escolhas envolvendo dois programas FT. Para essa comparação, o atraso da alternativa padrão em cada condição foi escolhido para ser o mesmo que a média (de grupo) de tempo para completar o FR na condição correspondente da fase anterior. Por exemplo, se os pombos respondessem a uma resposta por segundo, e se na Fase FT-FT a alternativa padrão tivesse sido FR5, FR10 e FR30, na fase FT-FT a alternativa padrão seria, em diferentes condições, FT5s FR10s e FT30s. A alternativa ajustável foi também um FT.

Para comparar os pontos de indiferença das fases FR-FR e FT-FT, os autores compararam, para cada atraso padrão (tempo para completar o FR na fase FR-FR e atraso correspondente na fase FT-FT), qual o ponto de indiferença em cada fase. Para isso, assumindo que no ponto de indiferença

$V_6=V_2$  (os subscritos 6 e 2 referem-se às alternativas ajustável e padrão, respectivamente), a partir da equação 1 verifica-se que :

$$D_6 = \frac{A_6 - A_2}{K_i A_2} + \frac{A_6}{A_2} D_2 \quad (\text{Equação 3})$$

Note-se que na equação 3 o K está subscrito, porque esta equação pode ser aplicada quer à fase FR-FR quer à fase FT-FT, e partindo do pressuposto de que num FR a taxa de desconto do reforço é superior do que num FT (Mazur, 1986), o valor de K depende do programa de reforço.

Assim, se a exigência de respostas não afectasse o valor das alternativas, os pontos de indiferença na fase FT-FT deveriam corresponder aos tempos para completar o FR na fase FR-FR ( $K_{FR}=K_{FT}$  e as rectas dadas pela equação 3 deveriam sobrepor-se). Por outro lado, se a imposição de respostas desvalorizasse o reforço ( $K_{FR}>K_{FT}$ ), os pontos de indiferença na fase FT-FT seriam maiores que na fase FR-FR (ver equação 3). Os resultados de grupo não mostraram diferenças entre as fases. Contudo, a nível individual verificaram-se diferenças: para dois pombos os pontos de indiferença foram superiores na fase FR-FR e para os restantes dois foram superiores na fase FT-FT. Assim, poderíamos concluir que para metade dos animais a exigência de respostas desvaloriza o reforço e que para a outra metade há uma valorização. É necessário, contudo, salientar que as durações dos atrasos padrão na fase FT-FT não têm correspondência directa em termos individuais com a fase FR-FR (pois os valores usados foram a média de grupo), e quando há correspondência directa, ou seja, nos dados de grupo, não se encontram diferenças entre as fases.

Na última fase, fase FT-FR, a alternativa padrão era a mesma que na fase FT-FT, e na alternativa ajustável estava em vigor um FR. Comparando as fases FT-FT e FT-FR, para todos os pombos a alternativa padrão era exactamente igual, mudando apenas a alternativa ajustável: um FT na fase FT-FT e um FR na fase FT-FR. Assim, se a imposição de respostas não afectasse o valor do reforço de 6s, os tempos para completar o FR na fase FT-FR deveriam ser os mesmos que os pontos de indiferença na fase FT-FT. Por outro lado, se o FR desvalorizasse o reforço de 6s, na fase FT-FR os pontos de indiferença deveriam ser menores que na fase FT-FT. De facto, para todos os pombos os pontos de indiferença nesta última fase (FT-FR) foram menores que na fase FT-FT, sugerindo que a imposição de respostas desvaloriza o reforço, havendo uma preferência por atrasos simples em relação a programas de razão. Contudo, a variabilidade dos dados individuais mostra também que o efeito da exigência de respostas varia muito entre indivíduos.

Os estudos até agora mencionados sugerem que: (a) a equação hiperbólica de desconto descreve bem a relação entre o valor de um reforço e o seu atraso, tanto em atrasos sem exigência de respostas (FT) como em atrasos com exigência de respostas (FR); (b) há uma preferência por FT em relação a FR e (c) essa preferência poder-se-á dever à maior taxa de desconto no FR. Todavia, no que concerne a esta última afirmação, os resultados nem sempre foram consistentes.

Mazur e Kralik (1990) apontam dois problemas aos estudos anteriores, que poderão, em parte, explicar alguma da inconsistência. Em primeiro, talvez o esforço exigido nos estudos anteriores não tenha sido suficiente para que a preferência por um FT fosse mais saliente. De facto, a partir do último resultado, em que há uma preferência por FT comparativamente a programas FR, Grossbard e Mazur (1986) sugerem que a resposta de bicar (num FR) é mais aversiva do que os outros comportamentos do animal enquanto espera pela comida durante um FT. Dizem ainda que partindo desse pressuposto, a preferência por um atraso simples em relação a um FR com um tempo entre escolha e reforço equivalente, deveria ser tanto mais pronunciada quanto maior fosse o esforço da resposta e, consequentemente, maior a sua aversividade.

Em segundo, nos estudos anteriores a comparação entre uma condição com exigência de respostas e sem exigência de respostas foi sempre feita por meio de programas FR e FT, respectivamente. Ora, o tempo que o sujeito levava a completar um certo FR variava de ensaio para ensaio (por variações na taxa de resposta). Logo, a comparação não era entre dois programas com o mesmo atraso mas um dos programas sem exigência de respostas e o outro com exigência de respostas (FT vs. FR), mas sim entre um atraso fixo sem exigência de respostas e um atraso variável com exigência de respostas. Visto que os animais preferem atrasos variáveis a fixos se a média do atraso para o reforço for igual (Mazur, 1984), os resultados obtidos podem ter sido enviesados.

Assim, Mazur e Kralik (1990) fizeram uma experiência para controlar a variabilidade nos tempos para completar o FR. Para potenciar a diferença entre exigência/não exigência de respostas, exigiram mais força (medida em newtons- N) para uma resposta ser efectuada na tecla padrão. Apresentaram a pombos a opção entre dois programas de reforço, e ambos levavam a 3s de reforço (ver tabela 1). Na fase FR-FT os sujeitos tinham de escolher entre 3s de comida entregues após um FR (alternativa padrão) ou após um FT (alternativa ajustável). Na fase VT-FT, a escolha era entre um tempo variável (VT- do inglês, *Variable Time*) como alternativa padrão e um FT ajustável. Para cada pombo, os componentes do VT foram escolhidos de tal modo que a sua distribuição fosse o mais semelhante possível à distribuição dos tempos para completar o FR na fase FR-FT. Para aumentar o esforço da tarefa, em ambas as fases, era necessária uma força de 0.48 N para dar uma resposta na tecla padrão e apenas 0.10 N para responder na tecla ajustável.

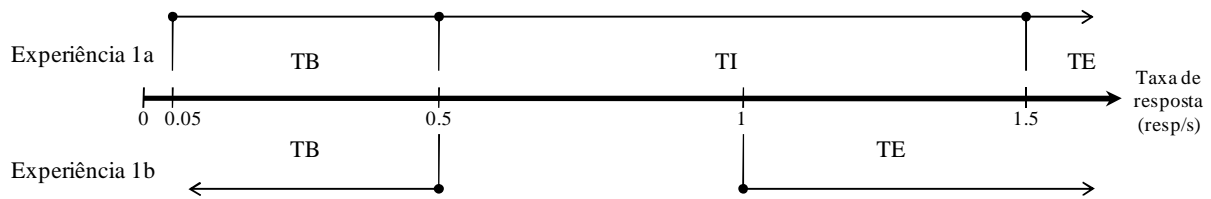
Se os pombos fossem indiferentes entre um programa com exigência de respostas (FR) e um programa sem exigência de respostas mas com o mesmo atraso (VT), os pontos de indiferença deveriam ser semelhantes entre as fases. Por outro lado, se a exigência de respostas desvalorizasse a alternativa padrão, os pontos de indiferença na fase FR-FT deveriam ser maiores que na fase VT-FT. Os autores encontraram pontos de indiferença superiores na fase FR-FT. Para além disso, quanto maior o parâmetro do FR, maior a diferença nos pontos de indiferença entre as fases. Ou seja, concluiu-se que há uma preferência por atrasos simples em comparação a programas de razão com o mesmo

tempo médio entre escolha e reforço, e que esta preferência é tanto mais acentuada quanto maior for o parâmetro do FR.

Os estudos apresentados procuraram, de algum modo, responder à questão: “Qual o efeito do esforço no valor de uma recompensa?”. Em todos estes estudos o esforço foi conceptualizado como o número de respostas exigidas para que o reforço fosse entregue. A conceptualização do esforço pelo número de respostas a emitir, implica que quanto maior é o número de respostas, maior é o atraso para o reforço. Assim, torna-se difícil distinguir entre os efeitos do esforço e do atraso no valor de um reforço. Para além disso, variações na taxa de resposta durante um FR levam a variações no atraso para o reforço, e por isso Mazur e Kralik (1990) comparam a preferência entre reforços entregues após um FR (atrasos variáveis com exigência de respostas) e após um VT (atraso variável sem exigência de respostas).

Porém, para que se possa concluir que os resultados encontrados se devem à variável manipulada, deve-se garantir, tanto quanto possível, que a única diferença entre as condições FR e VT é a variável esforço. Ora, apesar de se comparar um programa FR com um VT, cujos componentes foram escolhidos para emparelhar os tempos entre escolha e reforço no FR, não é certo que o efeito do atraso seja o mesmo nas duas alternativas. Para além disso, nos estudos descritos anteriormente, a sequência de estímulos durante um ensaio com um programa FR foi diferente da sequência de estímulos durante um ensaio FT ou VT. Em todos os programas, no momento da escolha, a luz branca de iluminação geral da caixa estava acesa e duas teclas laterais estavam iluminadas com verde e vermelho. Nos ensaios com programa FR, após a resposta de escolha, a luz branca de iluminação da caixa permanecia acesa, a tecla não escolhida apagava-se e a tecla escolhida permanecia iluminada. Nos ensaios com programas de tempo (FT ou VT), após a resposta de escolha, ambas as teclas de escolha apagavam-se e acendia-se uma luz de iluminação geral de cor correspondente à tecla escolhida. Assim, ao acender-se a luz da caixa com a cor da tecla escolhida, esta luz poderá ter funcionado como reforço secundário (Mazur, 1995). Isto porque independentemente do comportamento do animal, após a luz da caixa ter mudado de cor, era certo que receberia reforço passado alguns segundos. Assim, não é claro se estas diferenças de estímulos durante os atrasos poderão ter influenciado as escolhas dos pombos.

À semelhança das experiências apresentadas, o presente estudo procurou compreender o efeito do esforço no valor de uma recompensa. No entanto, pretendeu-se controlar de modo mais preciso quer os atrasos para os reforços quer os estímulos presentes durante os atrasos. Em vez de se conceptualizar o esforço como o número de respostas para um reforço, tentando depois controlar o atraso para o reforço, a base deste estudo foi conceptualizar o esforço como a taxa de resposta. Deste modo, mantendo sempre o mesmo atraso para o reforço, variou-se o número de respostas exigidas durante este atraso, variando-se assim a taxa de resposta e, conseqüentemente, o esforço.



**Figura 2** – Representação esquemática das taxas de resposta em cada fase na Experiência 1a e 1b. TB- Taxa Baixa; TI- Taxa Intermédia; TE- Taxa Elevada.

A Experiência 1a consistiu em quatro fases: Linha de Base (LB), Taxa Baixa (TB), Taxa Intermédia (TI) e Taxa Elevada (TE) (ver figura 2). Na LB a alternativa padrão era um FT 20s com 2s de reforço e a alternativa ajustável era um FT com 6s de reforço. Nas fases seguintes, a única diferença estava no número de respostas exigidas durante o atraso padrão. Na fase TB, durante o atraso padrão eram exigidas entre 1 e 10 respostas, na fase TI entre 10 e 30 respostas, e por fim, na fase TE, pelo menos 30 respostas.

Os pombos mostraram dificuldades em aprender a tarefa, e por isso o procedimento sofreu pequenas alterações de modo tornar a tarefa mais simples. A Experiência 1b consistiu em três fases: Linha de Base (LB), Taxa Baixa (TB) e Taxa Elevada (TE) (ver figura 2). Na LB a alternativa padrão era um FT 10s com 2s de reforço e a alternativa ajustável era um FT com 6s de reforço. Nas fases seguintes, a única diferença estava no número de respostas exigidas durante o atraso padrão. Na fase TB, durante o atraso padrão não eram permitidas mais de 5 respostas e na fase TE eram exigidas pelo menos 10 respostas.

## EXPERIÊNCIA 1a

### MÉTODO

#### *Sujeitos*

Os sujeitos foram seis pombos *Columba livia*, mantidos a aproximadamente 80% do seu peso em alimentação livre. No biotério os animais tinham sempre à disposição *grit* e água. A colônia de pombos tinha um ciclo de luz:escuridão de 13h:11h, com as luzes a acenderem-se às 8h, e a temperatura a variar entre os 20 e 22°C. Todos os pombos tinham experiência anterior com vários procedimentos, mas nenhum tinha participado em experiências de escolha entre diferentes durações de reforço ou em experiências em que se varia o atraso para a obtenção de comida.

#### *Aparato Experimental*

A experiência foi conduzida em três caixas experimentais para pombos Lehigh Valley, com 34 cm de altura, 35 cm de comprimento e 31 cm de profundidade. Cada caixa estava equipada com uma ventoinha para circular o ar e mascarar sons do exterior. O painel frontal estava equipado com três teclas alinhadas horizontalmente, de 2.5 cm de diâmetro e com uma distância de 9 cm entre elas (do centro para o centro das teclas). A distância da base das teclas ao gradeamento de arame do chão da caixa era de 22.5 cm. As teclas laterais podiam-se iluminar de verde ou vermelho e a central de branco. A abertura do alimentador (com 6 x 5 cm) estava centrada horizontalmente no painel frontal, 8.5 cm acima do chão. Quando activado, o alimentador era iluminado com uma luz de 7.5-W. Na parede oposta ao painel frontal, centrada horizontalmente e a 30 cm acima do chão, encontrava-se uma luz de 4.8-W que iluminava toda a caixa.

Os eventos foram programados em Visual Basic 2008, o qual também registou os dados da sessão.

#### *Procedimento Geral*

A experiência consistiu em quatro fases: Linha de Base (LB), Taxa Baixa (TB), Taxa Intermédia (TI) e Taxa Elevada (TE). Em cada fase, os pombos escolheram entre duas alternativas: a padrão e a ajustável. Com a excepção da LB em que não eram exigidas respostas, as restantes fases diferiram entre si apenas no número de respostas exigidas durante o atraso da alternativa padrão. A ordem de exposição às três fases experimentais foi contrabalanceada (ver tabela 2). De modo a garantir que o procedimento de ajustamento do atraso foi aplicado apenas quando os sujeitos conseguiram cumprir os requisitos exigidos do número de respostas na alternativa padrão, em todas as fases a condição de ajustamento foi precedida por uma condição de treino.

**Tabela 2-** Ordem das fases para cada pombo e número de sessões em cada condição na Experiência 1a. Os asteriscos assinalam as condições que não foram completadas.

Pombo		Ordem das Fases							
		LB		TB		TI		TE	
P8201	Treino 10	Ajust 23	Treino 56*	Ajust	Treino	Ajust	Treino	Ajust	
		LB		TI		TB		TE	
PG15	Treino 62	Ajust 14*	Treino	Ajust	Treino	Ajust	Treino	Ajust	
		LB		TE		TB		TI	
P928	Treino 46	Ajust 18*	Treino	Ajust	Treino	Ajust	Treino	Ajust	
		LB		TB		TE		TI	
P440	Treino 58	Ajust 13*	Treino	Ajust	Treino	Ajust	Treino	Ajust	
		LB		TI		TE		TB	
P366	Treino 10	Ajust 49	Treino 7*	Ajust	Treino	Ajust	Treino	Ajust	
		LB		TE		TI		TB	
P726	Treino 10	Ajust 24	Treino 25*	Ajust	Treino	Ajust	Treino	Ajust	

Para os pombos P8201, PG15 e P928 a alternativa padrão estava associada à tecla verde e para os pombos P440, P366 e P726 estava associada à vermelha. Por uma questão de conveniência as condições serão referidas pela ordem e com as especificações apresentadas ao pombo P8201. Assim, ao longo de toda a experiência a tecla verde representa a alternativa padrão e o reforço para esta alternativa é sempre de 2s atrasado 20s. A tecla vermelha representa a alternativa ajustável e o reforço para esta alternativa é sempre de 6s. O atraso da alternativa ajustável foi sempre de 20s nas condições de treino e foi variado nas condições de ajustamento.

A primeira fase constituiu a Linha de Base: a alternativa padrão tinha um atraso de 20s durante os quais não eram requeridas respostas. Na fase Taxa Baixa durante os 20s do atraso da alternativa padrão eram requeridas entre 1 e 10 respostas, na fase Taxa Intermédia eram requeridas entre 10 e 30 respostas e na fase Taxa Elevada eram requeridas pelo menos 30 respostas. Durante toda a experiência nunca eram requeridas respostas durante o atraso para a alternativa ajustável.

De seguida será descrito o procedimento de cada fase (LB, TB, TI, TE) e de cada condição (treino e ajustamento).



*(a) Linha de Base*

*(a.1) Treino*

As sessões de treino continham 44 ensaios, 20 dos quais para cada alternativa e 4 ensaios em que a primeira bicada na tecla central para começar o ensaio era reforçada. Estes 4 ensaios foram pseudoaleatorizados com os restantes ensaios. Estes ensaios foram necessários para manter a resposta à tecla central e foram introduzidos a partir da 12.<sup>a</sup> sessão. A sessão acabava ao fim de 44 ensaios ou de 90 min. (correspondendo a 60 ensaios), o que ocorresse primeiro. Este critério para terminar a sessão apenas foi implementado a partir da 12.<sup>a</sup> sessão, antes da qual não havia limite temporal fixo para terminar a sessão.

O ensaio começava com a iluminação da tecla central com cor branca. A bicada nesta tecla era exigida para prevenir que o pombo estivesse mais perto de uma das teclas laterais no momento da escolha. Uma bicada na tecla central apagava-a e acendia-se uma das teclas laterais: com cor vermelha ou verde. A posição das cores variou de modo pseudoaleatório ao longo dos ensaios para controlar qualquer preferência por uma das teclas. A primeira bicada numa das teclas laterais constituía a escolha do sujeito dando início ao atraso. No fim do atraso a tecla apagava-se e era accionado o alimentador durante 2s ou 6s, dependendo da escolha. Após o reforço, seguia-se o intervalo entre ensaios (IEE) durante o qual apenas a luz de iluminação geral da caixa estava acesa. A duração do IEE foi variável, de modo a que cada ensaio durasse 90s. Em alguns ensaios de treino poderia acontecer que o tempo decorrido entre o início do ensaio e o fim do reforço fosse superior a 90s. Nesses casos o IEE era de 15s. Tanto a tecla central branca como as teclas laterais de escolha permaneciam acesas durante 10s no máximo. Se não houvesse resposta durante estes períodos seguia-se o IEE e o ensaio era repetido as vezes que fosse necessário até terminar com reforço.

Nas condições de treino o atraso para o reforço de 6s permanecia sempre igual ao atraso do reforço de 2s, ou seja, com 20s.

Note-se que na Linha de Base, para ambas as alternativas de escolha (teclas verde e vermelha), as únicas bicadas exigidas eram na tecla central para iniciar o ensaio e na tecla vermelha ou verde para iniciar o atraso. Assim, um ensaio podia terminar sem reforço se não houvesse resposta na tecla central ou numa das laterais.

A condição de treino terminava após um mínimo de 10 sessões, quando em três sessões consecutivas pelo menos 85% dos ensaios apresentados eram reforçados, i.e., terminavam com comida. Isto assegurou que o pombo estava a desempenhar bem a tarefa.

Para os pombos P8201, P366 e P726, esta fase consistiu apenas no procedimento anteriormente explicado. Porém, para os pombos PG15, P928 e P440 atingirem o critério, foi necessário usar alguns procedimentos intermédios.

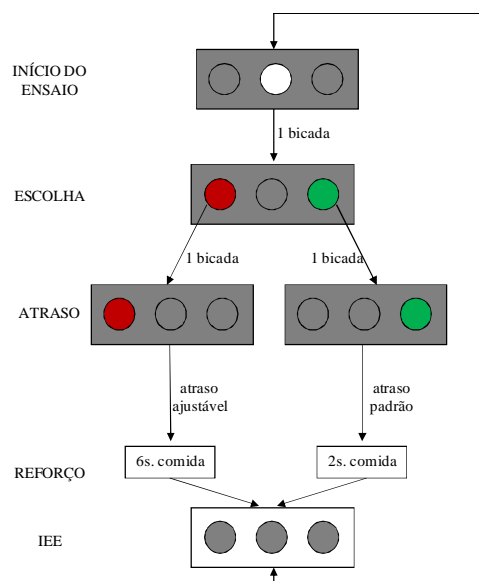
Em primeiro, os pombos foram expostos ao procedimento originalmente planeado, à semelhança dos restantes pombos. Porém, após 8 sessões, os pombos continuavam a perder muitos

ensaios (em média 35% dos ensaios apresentados não eram reforçados) e não era visível uma tendência de melhoria no desempenho. Assim, pensou-se que seria necessário tornar o reforço mais contíguo à resposta, já que num FT20s a resposta que dá origem ao reforço (a resposta de escolha) dista 20s da sua consequência. Para além disso, as mudanças foram feitas em ambas as alternativas para não induzir qualquer viés. Assim, diminuiu-se o atraso para 5s. Contudo, como essa mudança não alterou o desempenho dos pombos, decidiu-se substituir o atraso (FT) por um FR. Primeiro implementou-se um FR1 e depois um FR2. Quando a percentagem de ensaios reforçados era igual ou superior a 85%, substituiu-se o FR por um FT2s. Por fim, este FT foi sendo aumentado gradualmente, começando nos 2s e acabando no objectivo final de um atraso de 20s.

*(a.2) Ajustamento*

Os 40 ensaios foram divididos em 4 blocos de 10 ensaios. Em cada bloco os primeiros dois ensaios eram ensaios de escolha forçada (como os ensaios da condição de treino) seguidos de 8 ensaios de escolha livre, em que após uma bicada na tecla central ambas as teclas laterais se acendiam, uma vermelha e a outra verde. Em cada bloco foi acrescentado um ensaio em que a primeira resposta na tecla central era reforçada, perfazendo um total de 44 ensaios.

A sequência de eventos num ensaio de escolha livre foi idêntica à dos ensaios forçados, com a excepção de que após a resposta na tecla central branca para iniciar o ensaio acendiam-se ambas as teclas laterais: uma com cor vermelha e outra com cor verde (ver figura 3). A primeira bicada numa das teclas laterais constituía a escolha do sujeito. Assim, a tecla não escolhida apagava-se e iniciava-se o atraso.



**Figura 3** - Esquema da sequência de eventos durante um ensaio de escolha livre.

O atraso da alternativa ajustável era iniciado com 20s e podia mudar após cada bloco de 10 ensaios, dependendo das escolhas do pombo. Se o pombo escolhesse a tecla vermelha menos de três vezes num dado bloco de 8 ensaios de escolha, o atraso ajustável diminuía 1s no bloco seguinte. Se escolhesse a tecla vermelha mais de 5 vezes num dado bloco, o atraso aumenta 1s. Se a tecla ajustável fosse escolhida entre 3 e 5 vezes o atraso no bloco seguinte permanecia inalterado. Na Linha de Base, no primeiro bloco da primeira sessão o atraso da alternativa ajustável foi iniciado a 20s. Nas sessões seguintes o atraso era calculado com base no último bloco da sessão anterior como se se tratasse de um contínuo. O ponto de indiferença foi calculado pela média do atraso ajustável nas últimas 5 sessões.

A condição de ajustamento era terminada após um mínimo de 15 sessões quando três critérios de estabilidade eram atingidos: (a) a média do atraso da alternativa ajustável mais alta ou mais baixa de uma sessão não ocorre nas últimas 5 sessões, (b) a média do atraso da alternativa ajustável das últimas 5 sessões não pode ser nem a mais alta nem a mais baixa da média de 5 sessões consecutivas e (c) a média do atraso da alternativa ajustável das últimas 5 sessões não difere mais de 10% da média das 5 sessões antecedentes<sup>1</sup>.

#### *(b) Taxa Baixa*

##### *(b.1) Treino*

Nesta fase era necessário ensinar os animais a responder com uma taxa baixa na alternativa padrão: igual ou superior a 0.05 mas inferior ou igual a 0.5 respostas por segundo, em média.

Em cada sessão, foi imposto durante o atraso padrão um limite mínimo de uma resposta e um limite máximo que foi diminuído progressivamente. A diminuição progressiva do número máximo de respostas foi feita através de um programa de reforço percentílico: Com base na sessão anterior ordenou-se o número de respostas dadas durante o atraso na alternativa padrão em cada ensaio. De seguida calculou-se qual o número de respostas abaixo do qual se encontravam 80% dos ensaios. Assim, esse valor seria usado na sessão seguinte como limite máximo de respostas. Por outras palavras, numa sessão, caso o animal tivesse o mesmo desempenho do dia anterior, em 80% dos ensaios seria reforçado. O limite só diminuía quando numa sessão fossem completados os 44 ensaios com reforço.

Note-se que nesta fase, ao impor um mínimo de 1 resposta durante o atraso padrão, se ao fim dos 20s não tivesse havido respostas, o ensaio era repetido. Por outro lado, devido ao limite máximo, à 11.<sup>a</sup> resposta o ensaio era terminado e repetido.

O critério de estabilidade aplicado foi o mesmo que o da fase de treino na Linha de Base, assegurando que o animal conseguia desempenhar bem a tarefa em ambas as alternativas.

---

<sup>1</sup> Se  $M1$  = média do atraso da alternativa ajustável das penúltimas 5 sessões e  $M2$  = média do atraso da alternativa ajustável das últimas 5 sessões, então  $\frac{M2-M1}{M1} \leq 0.10$ .

### *(b.2) Ajustamento*

O procedimento na condição de ajustamento da fase Taxa Baixa foi idêntico ao correspondente da Linha de Base, com a exceção de que, durante o atraso da alternativa padrão eram requeridas entre 1 e 10 respostas para que fosse entregue o reforço de 2s. O atraso da alternativa ajustável era iniciado com 20s.

### *(c) Taxa Intermédia*

#### *(c.1) Treino*

Nesta fase era necessário ensinar os animais a responder a uma taxa intermédia na alternativa padrão: superior ou igual a 0.5 mas inferior ou igual a 1.5 respostas por segundo, em média.

De modo a aumentar a taxa de resposta durante o atraso padrão, nas primeiras sessões este atraso foi substituído por um FR10. Assim, ao escolher (de modo forçado) a tecla verde, independentemente do tempo, o reforço era entregue imediatamente após a 10<sup>a</sup> bicada. De seguida, quando em pelo menos 80% dos ensaios o pombo estava a completar o FR10 em menos de 20s, introduziu-se o atraso de 20s, durante o qual tinham de ser dadas pelo menos 10 respostas. Deste modo estaria assegurado o cumprimento do limite mínimo de respostas. De seguida, para ensinar o limite máximo, seria aplicado o programa de reforço percentílico do mesmo modo que na fase Taxa Baixa.

Note-se que nesta fase, ao impor um limite mínimo de respostas durante o atraso padrão, se ao fim dos 20s o número de respostas fosse inferior a 10, o ensaio era repetido. Por outro lado, devido ao limite máximo, à 31.<sup>a</sup> resposta o ensaio era terminado e repetido.

O critério de estabilidade aplicado foi o mesmo que o da fase de treino na Linha de Base, assegurando que o animal conseguia desempenhar bem a tarefa em ambas as alternativas.

#### *(c.2) Ajustamento*

O procedimento na condição de ajustamento da fase Taxa Intermédia foi idêntico ao correspondente da Linha de Base, com a exceção de que, durante o atraso da alternativa padrão eram requeridas entre 10 e 30 respostas para que fosse entregue o reforço de 2s.

### *(d) Taxa Elevada*

#### *(d.1) Treino*

Nesta fase era necessário ensinar os animais a responder a uma taxa elevada na alternativa padrão: igual ou superior a 1.5 respostas por segundo, em média.

De modo a aumentar a taxa de resposta durante o atraso padrão, nas primeiras sessões o atraso foi substituído por um FR30. De seguida, o procedimento foi semelhante ao planeado na fase Taxa Intermédia: quando em pelo menos 80% dos ensaios o pombo estivesse a completar o FR30 em menos de 20s, introduziu-se o atraso de 20s, durante o qual tinham de ser dadas pelo menos 30 respostas. Deste modo estaria assegurado o cumprimento do limite mínimo de respostas.

Note-se que nesta fase apenas é imposto um limite mínimo de respostas durante o atraso padrão, pelo que, se ao fim dos 20s o número de respostas fosse inferior a 30, o ensaio era repetido.

O critério de estabilidade aplicado foi o mesmo que o da fase de treino na Linha de Base, assegurando que o animal conseguia desempenhar bem a tarefa em ambas as alternativas.

#### *(d.2) Ajustamento*

O procedimento na condição de ajustamento da fase Taxa Elevada foi idêntico ao correspondente da Linha de Base, com a excepção de que, durante o atraso da alternativa padrão eram requeridas pelo menos 30 respostas para que fosse entregue o reforço de 2s.

## **RESULTADOS E DISCUSSÃO**

### *Pontos de Indiferença*

Apenas os pombos P8201, P928, P366 e P726 terminaram uma condição de ajustamento (LB-ajustamento) e os pontos de indiferença foram 40s, 44.5s, 55.25s e 35.2s, respectivamente. Estes valores significam que os pombos foram indiferentes entre 2s de comida atrasados 20s e 6s de comida atrasados esse valor.

A variabilidade entre os sujeitos e os valores do atraso no ponto de indiferença são consistentes com dados da literatura. Numa condição muito semelhante, os pombos de Grossbard e Mazur (1986) tiveram pontos de indiferença que variaram entre os 32s e 56s, aproximadamente. O facto dos pontos de indiferença encontrados na Linha de Base serem semelhantes aos referidos na literatura comprova que o procedimento foi bem construído e aplicado, e que os pombos aprenderam a tarefa. Porém, como se verá de seguida, a elevada proporção de ensaios não reforçados foi um impedimento à continuação da experiência.

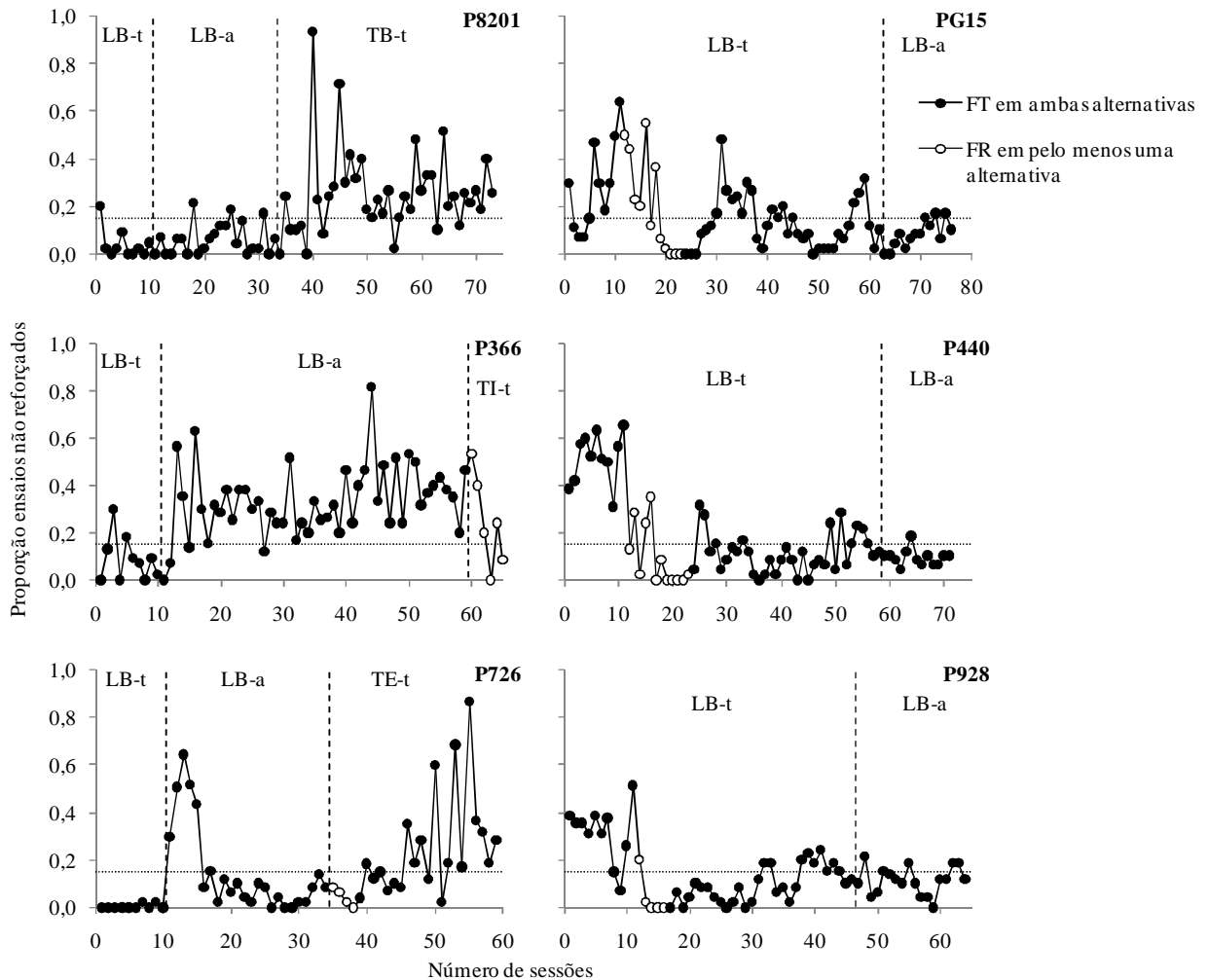
### *Proporção de ensaios não reforçados*

Na figura 4 encontram-se os gráficos individuais da proporção do número de ensaios não reforçados em cada sessão. As linhas verticais separam cada uma das condições e a linha horizontal situa-se no valor 0.15, representando o critério de estabilidade nas fases de treino, segundo o qual em três sessões consecutivas não poderia haver mais de 15% de ensaios não reforçados por sessão.

Note-se que os pombos do painel da direita (figura 4) apenas passaram pela fase LB (condições de treino e ajustamento) e os pombos do painel da esquerda, para além dessa fase, iniciaram a condição de treino da fase seguinte (TB para P8201; TI para P366 e TE para P726).

Em cada gráfico os marcadores preenchidos a preto representam as sessões em que ambas as alternativas foram um FT e os marcadores preenchidos a branco representam as sessões em que pelo

menos uma das alternativas foi um esquema FR ao invés de um FT. Os círculos preenchidos a branco durante a LB-treino significam que em ambas as alternativas estava em vigor um FR e durante o treino na segunda fase de cada pombo (TI-treino para P366 e TE-treino para P726) significam que apenas no atraso padrão estava em vigor um FR.



**Figura 4** – Proporção de ensaios não reforçados em cada sessão. LB-t – treino da Linha de Base; LB-a – ajustamento da Linha de Base; TB-t – treino da fase Taxa Baixa; TI-t – treino da fase Taxa Intermédia; TE-t – treino da fase Taxa Elevada.

Durante a LB-treino, os três animais do painel da esquerda tiveram poucos ensaios não reforçados e terminaram esta condição após 10 sessões. Porém, para os pombos do painel da direita, a proporção de ensaios não reforçados começou elevada e não mostrava tendência para diminuir (primeira série de marcadores pretos). Nesta condição, LB-treino, os ensaios podiam ser não reforçados por dois motivos: ou o pombo não iniciou o ensaio (não bicou na tecla central durante o período de 10s) ou após iniciar o ensaio o pombo não escolheu a tecla lateral. Para além disso, a razão dos ensaios não reforçados não foi a mesma para os três pombos: para o PG15 e P928 a maior parte

dos ensaios não reforçados deveram-se à ausência de resposta de escolha (99% e 90%, respectivamente) e para o P440 55% dos ensaios não reforçados deveram-se à falta de resposta na tecla central para iniciar o ensaio. A partir da sessão 12, com a introdução do reforço da bicada na tecla central em alguns ensaios, e com a substituição do atraso simples por um esquema de reforço FR, o número de ensaios não reforçados diminuiu substancialmente (ver círculos a branco). De seguida, com a substituição do esquema FR por um FT gradualmente maior, permitiu que os animais atingissem o critério de estabilidade e passassem à fase seguinte.

Na LB-ajustamento, apenas os pombos P726 e P366 apresentaram ensaios com uma proporção elevada de ensaios não reforçados. Para o P726, a introdução do reforço esporádico da bicada na tecla central diminuiu a proporção de ensaios não reforçados. Já o P366 permanece toda a fase com uma média de 33% ensaios não reforçados por sessão. Contudo, como nesta fase o critério de estabilidade referia-se à estabilidade do ajustamento do atraso, o número de ensaios não reforçados não foi tido em conta e a experiência prosseguiu sem alterações.

Por fim, apenas três pombos passaram por uma segunda fase: P8201 (TB-treino), P366 (TI-treino) e P726 (TE-treino). Tanto para o P8201 como para o P726 a proporção de ensaios não reforçados aumentou nesta fase em comparação com a fase anterior<sup>2</sup>. Porém, as causas para a perda destes ensaios foi diferente. Para o P8201, 24% dos ensaios não foram reforçados porque o número de respostas durante o atraso foi maior ou menor que o requerido e 76% dos ensaios porque não houve escolha. Para além disso, dos ensaios com a alternativa padrão em que era dada a resposta de escolha, em apenas em 16% dos ensaios é que o pombo não era reforçado por não dar o número de respostas exigidas durante o atraso. Isto significa que quando o pombo escolhia, conseguia dar o número de respostas exigidas; o problema estava em muitas vezes não haver resposta de escolha. Já para o P726, 44% dos ensaios não reforçados deveram-se ao não cumprimento do número de bicadas exigidas, 35% porque não houve escolha e os restantes 21% porque o ensaio não foi iniciado. Por fim, o P366 foi exposto a poucas sessões nesta nova fase em que ainda não tinham sido aplicadas as contingências finais.

Em suma, todos os pombos apresentaram, em algum momento da experiência, uma elevada proporção de ensaios não reforçados o que se revelou um impedimento à continuação da experiência. Contudo, não houve qualquer regularidade do número de ensaios não reforçados: diferentes pombos apresentaram mais ensaios não reforçados em diferentes condições e por diferentes causas.

Por um lado, não é claro porque é que três pombos (P8201, P366 e P726) conseguiram aprender bem a tarefa na condição LB-treino, mas para os restantes três foi necessária uma mudança no procedimento de modo a tornar a aprendizagem mais gradual. Por outro lado, apesar de ter desempenhado bem a tarefa nesta primeira condição, na condição seguinte de ajustamento, o pombo

---

<sup>2</sup> É importante relembrar que nestas fases os ensaios não reforçados podiam dever-se ao pombo: (a) não iniciar o ensaio; (b) não escolher uma tecla lateral ou ainda (c) não cumprir o número de respostas exigidas durante o atraso (dando mais ou menos respostas que o exigido).

P366 começa a não dar a resposta de escolha em muitos ensaios. Por fim, para os pombos P8201 e P726, nas condições de treino das fases TB e TE respectivamente, apesar de a tarefa ter sido ensinada de modo gradual, a proporção de ensaios não reforçados era elevada e após 56 e 25 sessões (para os pombos P8201 e P726, respectivamente) não apresentava tendência para decrescer. Visto que os pombos pareciam ter dificuldade em aprender a tarefa imposta, decidiu-se terminar esta experiência e torná-la mais simples.

De modo a simplificar a experiência, tentou-se perceber quais os factores que poderiam estar a dificultar o seu progresso.

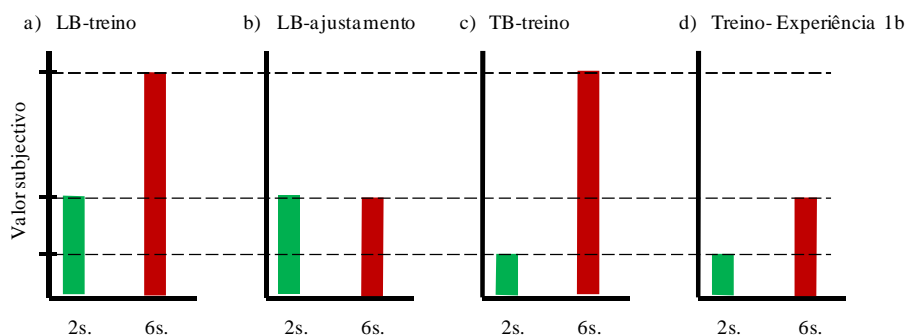
Em primeiro lugar, nos ensaios forçados a transição entre o período de escolha (forçada) e o atraso não era sinalizada, ao passo que nos ensaios de escolha livre, no momento da escolha, a tecla não escolhida apagava-se, iniciando-se o atraso. Assim, pode acontecer que os pombos tivessem percebido não um atraso de 20s, mas um atraso variável (com duração igual à latência de escolha + 20s). Assim, pensou-se que, nos ensaios forçados, poderia haver uma confusão entre o atraso programado (20s) e o atraso percebido pelos pombos. Contudo, não é claro porque razão este facto poderia levar o pombo a deixar de responder, já que em outros estudos, quando o programa utilizado é um FR, esta transição também não costuma ser sinalizada (ex. Grossbard & Mazur, 1986) e os animais aprendem a tarefa. Ainda assim, nos programas FT a transição entre escolha e o atraso costuma ser sinalizada (ex. Grossbard & Mazur, 1986): no momento da escolha, acende-se uma luz de iluminação geral da caixa, com uma cor correlacionada com a escolha (por exemplo luz vermelha quando é escolhida a tecla vermelha e luz verde quando é escolhida a tecla verde). Assim, uma das alterações efectuadas na Experiência 1b foi a sinalização do início do atraso. Porém, devido a impossibilidades de equipamento, ao invés da sinalização com luzes de iluminação da caixa, no período de escolha a(s) tecla(s) estava(m) a piscar, e a sua iluminação tornava-se constante após a escolha e início do atraso.

Na maior parte dos estudos de autocontrolo com animais, a duração dos atrasos é pequena, não ultrapassando, na maior parte dos casos, os 10s. Na Experiência 1a, o atraso da alternativa padrão durou 20s. Esta duração de 20s foi escolhida pois permitia que se exigissem três níveis de taxa de resposta (Taxa Baixa, Taxa Intermédia e Taxa Elevada). O grande intervalo de tempo entre a escolha e o reforço (20s) pode ter dificultado a associação entre a resposta de escolha e a sua consequência. Contudo, a dificuldade dos pombos em aprender a tarefa não se pode apenas dever à duração do atraso padrão, já que existem estudos (Mazur, 1986; Grossbard & Mazur, 1986) que utilizaram um atraso padrão de 20s ou até mais, e os autores não reportam qualquer dificuldade acrescida. Mas é necessário afirmar que os animais desses estudos são geralmente experientes e a sua história experimental poderá explicar a facilidade numa tarefa com atrasos longos. A segunda alteração do procedimento foi a diminuição do atraso padrão. Como consequência da diminuição do atraso, a terceira alteração foi o uso de apenas duas fases experimentais (Taxa Baixa e Taxa Elevada).



Com o objectivo de controlar a taxa de reforço, que pode afectar as escolhas de autocontrolo (Lattal, 2010 para discussão), a duração dos ensaios foi mantida constante. Por conseguinte, foi imposto aos animais uma série de limites temporais dentro de um ensaio: por exemplo, 10s para iniciar o ensaio e 10s para escolher uma das teclas laterais. Assim, alguns ensaios não eram reforçados por não haver uma resposta de escolha no período de 10s. A quarta alteração do procedimento foi o aumento deste período, já que assim poder-se-iam reforçar ensaios em que a latência de escolha era maior.

Por fim, no treino da segunda fase pela qual alguns pombos passaram (TB para P8201 e TE para P726) a maior parte dos ensaios não foram reforçados porque não houve resposta para iniciar o ensaio ou escolher uma tecla lateral. Quando havia resposta na tecla lateral, os animais conseguiam cumprir os requisitos exigidos durante o atraso padrão a maior parte das vezes. Para além disso, para os pombos P8201 e P726 (nas respectiva segunda fase), 99% e 88% dos ensaios em que não houve a resposta de escolha, foram na alternativa padrão. Isto sugere que de alguma forma, os pombos estavam a rejeitar esta alternativa padrão. Supomos que esta “rejeição” poderá ter sido causado pelo facto de nesta segunda fase a alternativa padrão ter um valor muito inferior em comparação com a alternativa ajustável, e por isso a sua escolha (forçada) ser “recusada”. Vejamos, por exemplo, o caso do pombo P8201 (figura 5).



**Figura 5** - Representação esquemática do valor das alternativas padrão (barra verde) e ajustável (barra vermelha) em cada condição para o P8201.

Na figura 5 encontram-se duas barras: a verde representa o reforço de 2s (a alternativa padrão) e a vermelha o reforço de 6s (a alternativa ajustável). Quanto mais alta a barra, maior o valor subjectivo desse reforço. Na LB-treino (figura 5- a), ambos os reforços têm um atraso de 20s e por isso o valor subjectivo do reforço de 6s é maior que o de 2s.

Na LB-ajustamento (figura 5- b), as contingências para o reforço de 2s continuam as mesmas que na condição anterior: os 2s de reforço são atrasados 20s e durante o atraso não são exigidas respostas. Assim, do treino (figura 5-a) para o ajustamento (figura 5-b) da LB, o valor do reforço de 2s mantém-se. Por outro lado, na LB-ajustamento (figura 5-b) através do procedimento de ajustamento

do atraso que faz com que os reforços fiquem com o mesmo valor, o reforço de 6s diminui (porque o seu atraso aumentou de 20s para 40s) para ficar com um valor semelhante ao reforço de 2s.

De seguida, segue-se a condição TB-treino (figura 5-c). Nesta fase o atraso do reforço de 6s passa a ser novamente 20s, e por isso, ao passar da condição LB-ajustamento (figura 5-b) para Taxa Baixa-treino (figura 5-c) o valor do reforço de 6s aumenta. Relativamente ao reforço de 2s as contingências para o reforço são agora diferentes da condição anterior: os 2s de reforço são atrasados 20s e durante o atraso têm de ser dadas entre 1 e 10 respostas. Assim, se esta restrição de respostas não tivesse qualquer efeito no valor do reforço de 2s, o valor deste reforço deveria manter-se em relação à condição anterior (LB-ajustamento). Porém, o facto do pombo “recusar” a responder à alternativa padrão leva-nos a supor que esta restrição desvaloriza o reforço de 2s, isto é, o valor deste reforço diminui ao passar da condição LB-ajustamento (figura 5-b) para Taxa Baixa-treino (figura 5-c). A grande diferença de valor entre as alternativas padrão e ajustável na fase TB-treino (figura 5- c) pode ter feito com que o reforço de 6s fosse muito mais preferido e pode ter levado o animal a não responder à alternativa menos valorizada.

Assim, a última alteração efectuada no procedimento foi a seguinte: ao passar de uma condição de ajustamento (ex. figura 5-b) para uma de treino, em vez do atraso ajustável voltar a ser 20s, aumentando o valor do reforço de 6s (figura 5-c), manteve-se o atraso do ponto de indiferença da fase anterior de ajustamento (comparar figura 5- b com figura 5-d). Assim, mesmo que a imposição/restricção de respostas na alternativa padrão faça com que o valor do reforço de 2s diminua, a diferença entre as alternativas não é tão acentuada (figura 5-d).

Porque nenhum dos factores isoladamente poderia explicar as várias dificuldades dos pombos nas várias fases, ao invés de investigar quais as causas para o elevado número de ensaios não reforçados, preferiu-se modificar todos os factores que se pensou que pudessem estar a afectar a experiência.

## EXPERIÊNCIA 1b

### MÉTODOS

#### *Sujeitos*

Os sujeitos foram os mesmos da Experiência 1a.

#### *Aparato Experimental*

O equipamento utilizado foi o mesmo da Experiência 1a.

#### *Procedimento Geral*

A experiência consistiu em três fases: Linha de Base (LB), Taxa Baixa (TB) e Taxa Elevada (TE). Cada fase era constituída por duas condições: treino e ajustamento (ver tabela 3)

**Tabela 3** - Ordem das fases para cada pombo com o número de sessões em cada condição na Experiência 1b. Nas condições de ajustamento a duração do atraso para a alternativa ajustável no ponto de indiferença encontra-se entre parênteses e é expresso em segundos. Os asteriscos assinalam as condições que não foram completadas.

Pombo	LB		TE		TB		TE		TB		TE	
	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.
P8201	5	15 (34.45)	13	26 (38.9)	12	13 (24.5)	11	11 (31.35)	10	17 (19.85)	10	8*
PG15	4	21 (39.85)	10	16 (31.95)	23	16 (42.25)	14	24 (27.85)	11	22 (46)	7	3*
P928	4	14 (31.85)	9	18 (35)	10	14 (22.2)	7	18 (37.8)	6	19 (30.4)	8	23 (37.1)
P440	5	27 (20)	35	10 (27.5)	61	21*						
P366	5	17 (33.1)	41	29 (54.2)	25	20 (22.35)						

Pombo	LB		TB		TE		TB		TE		TB	
	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.	Treino	Ajust.
P726	4	14 (25.95)	80	15 (22.35)	9	12 (23.35)	23*					

Para cada pombo, a associação entre a cor da tecla e a alternativa manteve-se igual à Experiência 1a. Por uma questão de conveniência as condições serão referidas pela ordem e com as especificações apresentadas ao pombo P8201: a tecla vermelha representa a alternativa ajustável e o

reforço para esta alternativa é sempre de 6s; a tecla verde representa a alternativa padrão e o reforço para esta alternativa é sempre de 2s atrasado 10s.

A primeira fase constituiu a Linha de Base: a alternativa padrão tinha um atraso de 10s durante os quais não eram requeridas respostas e o atraso para a alternativa ajustável foi iniciado com 10s. Na fase Taxa Elevada durante o atraso da alternativa padrão eram requeridas pelo menos 10 respostas e na fase Taxa Baixa eram permitidas no máximo 5 respostas. Na condição de ajustamentos destas fases experimentais o atraso para a alternativa ajustável era iniciado com o valor do ponto de indiferença da fase anterior. Durante toda a experiência nunca eram requeridas respostas durante o atraso para a alternativa ajustável.

A partir da LB estabeleceu-se a ordem de apresentação das seguintes fases: se, nos últimos três dias da LB- ajustamento, a média da taxa de resposta durante o atraso padrão fosse inferior a 1 resposta por segundo, a fase que se seguia era Taxa Elevada e se a taxa de resposta fosse superior a 1, seguia-se a fase Taxa Baixa. Assim, a experiência seguiu um delineamento ABCBCB em que A representa a Linha de Base e B e C as fases Taxa Elevada e Taxa Baixa, respectivamente (excepto para o pombo P726 para o qual a correspondência para B e C é inversa).

De seguida será descrito o procedimento de cada fase.

#### *(a) Linha de Base*

##### *(a.1) Treino*

O procedimento geral foi semelhante ao da Experiência 1a: nas condições de treino as sessões continham igualmente 44 ensaios, 40 dos quais ensaios forçados (metade para cada alternativa) e 4 ensaios em que a primeira bicada na tecla central branca era reforçada. A sessão acabava ao fim de 44 ensaios ou 70 min. (correspondendo a 60 ensaios), o que ocorresse primeiro.

A sequência de eventos dentro de um ensaio foi semelhante à da Experiência 1a. O ensaio começava com a iluminação da tecla central com cor branca. Uma bicada na tecla central apagava-a e uma das teclas laterais começava a piscar (250ms ligadas; 250ms desligadas): com cor vermelha ou verde. A primeira bicada numa das teclas laterais constituía a escolha do sujeito, tornando a iluminação da tecla escolhida permanente e iniciava-se o atraso. No fim do atraso a tecla apagava-se e era accionado o alimentador durante 2s ou 6s, dependendo da escolha. Após o reforço, seguia-se o intervalo entre ensaios (IEE) durante o qual apenas a luz da caixa estava acesa. A duração do IEE foi variável, de modo a que cada ensaio durasse 70s. Caso o tempo decorrido entre o início do ensaio e o fim do reforço fosse superior a 70s, o IEE era de 15s. A tecla central branca permanecia acesa no máximo durante 10s e as teclas laterais de escolha durante 15s no máximo. Se não houvesse resposta durante estes períodos seguia-se o IEE e o ensaio era repetido as vezes que fosse necessário até terminar com reforço.

Nesta primeira condição de treino o atraso para o reforço de 6s foi de 10s.

A condição de treino terminava após um mínimo de 10 sessões, quando em três sessões consecutivas pelo menos 85% dos ensaios apresentados eram reforçados.

#### *(a.2) Ajustamento*

O procedimento geral de ajustamento também foi semelhante ao da Experiência 1a: os 40 ensaios foram divididos em 4 blocos de 10 ensaios. Em cada bloco os primeiros dois ensaios eram de escolha forçada seguidos de 8 ensaios de escolha livre. Para além destes 40 ensaios, em cada bloco acrescentou-se um ensaio no qual a primeira bicada na tecla central levava a 4s de reforço, perfazendo, assim, os 44 ensaios da sessão.

A sequência de eventos num ensaio de escolha livre foi idêntica à dos ensaios forçados, com a exceção de que após a resposta na tecla central branca para iniciar o ensaio, ambas as teclas laterais começavam a piscar: uma com cor vermelha e outra com cor verde. A primeira bicada numa das teclas laterais constituía a escolha do sujeito. A tecla escolhida ficava com iluminação permanente enquanto que a tecla não escolhida apagava-se e iniciava-se o atraso.

O atraso da alternativa ajustável podia mudar após cada bloco de 10 ensaios, dependendo das escolhas do pombo, do mesmo modo que na Experiência 1a. O atraso da alternativa ajustável no ponto de indiferença foi calculado pela média das últimas 5 sessões.

A condição de ajustamento era terminada após um mínimo de 10 sessões, quando os três critérios de estabilidade (idênticos ao da Experiência 1a) eram atingidos.

#### *(b) Taxa Elevada*

##### *(b.1) Treino*

Nesta fase era necessário ensinar os animais a responder a uma taxa elevada na alternativa padrão: igual ou superior a 1 resposta por segundo, em média.

O atraso para a alternativa ajustável tinha a duração do ponto de indiferença da fase anterior.

Nas primeiras sessões de treino, o programa de reforço para a alternativa padrão em vez de um FT10s, foi um FR20 para aumentar a taxa de resposta nesta alternativa. Quando as primeiras 10 respostas eram dadas em 10s (ou menos), em pelo menos 80% dos ensaios, substituiu-se o FR20 por FT10s com exigência de pelo menos 5 respostas, ou seja, o reforço era entregue ao fim de 10s se tivessem sido dadas pelo menos 5 respostas. Este programa de reforço permitiu que o animal fosse exposto a uma condição em que uma resposta nem sempre era imediatamente seguida de reforço, permitindo ainda que a taxa de resposta diminuísse um pouco (para um mínimo de 0.5 respostas/s, em média), sem que deixasse de ser reforçado. Quando numa sessão pelo menos 85% dos ensaios eram reforçados, o número de bicadas exigidas aumentava de 5 para 10.

Note-se que nesta fase apenas é imposto um limite mínimo de respostas durante o atraso padrão, pelo que, se ao fim dos 10s o número de respostas fosse inferior a 10, o ensaio era repetido.

O critério de estabilidade aplicado foi o mesmo que o da fase de treino da Linha de Base.

### *(b.2) Ajustamento*

O procedimento na condição de ajustamento da fase Taxa Elevada foi idêntico ao correspondente da Linha de Base, com a exceção de que, durante o atraso da alternativa padrão eram requeridas pelo menos 10 respostas para que fosse entregue o reforço de 2s.

### *(c) Taxa Baixa*

#### *(c.1) Treino*

Nesta fase era necessário ensinar os animais a responder a uma taxa baixa na alternativa padrão: igual ou inferior a 0.5 respostas por segundo, em média.

O atraso para a alternativa ajustável tinha a duração do ponto de indiferença da fase anterior.

Em cada sessão foi imposto um limite máximo de respostas durante o atraso da alternativa padrão, e esse limite foi diminuído progressivamente. A diminuição progressiva do número máximo de respostas foi feita através de um programa de reforço percentílico: Com base na sessão anterior ordenou-se o número de respostas dadas durante o atraso na alternativa padrão em cada ensaio. De seguida calculou-se qual o número de respostas abaixo do qual se encontravam 60% dos ensaios. Assim, esse valor era usado na sessão seguinte como limite máximo. O limite só diminuía quando numa sessão fossem completados 44 ensaios com reforço até o limite ser de 5 respostas no máximo.

Note-se que nesta fase apenas é imposto um limite máximo de respostas durante o atraso padrão, pelo que à 6.<sup>a</sup> resposta o ensaio era terminado e repetido.

O critério de estabilidade aplicado foi o mesmo que o da fase de treino da Linha de Base.

Este foi o procedimento aplicado a todos os pombos, excepto para o pombo P726, para o qual o procedimento necessitou de algumas alterações. Durante 40 sessões, o pombo foi exposto ao procedimento planeado. Porém, a proporção de ensaios não reforçados era elevada (até esse momento, 32% dos ensaios apresentados não eram reforçados) e não era visível uma tendência de diminuição. Para além disso, a maioria dos ensaios não eram reforçados porque o pombo dava mais respostas do que o estabelecido (em 65% dos ensaios não reforçados). Assim, a primeira alteração foi a introdução de um aviso luminoso (luz da caixa acendia-se durante 250 ms) sempre que o pombo dava a quinta resposta durante o atraso padrão. Após a luz, se o pombo desse mais uma resposta o ensaio era terminado e repetido. Com este aviso pretendia-se que o animal aprendesse que após o piscar da luz, para receber comida, não deveria responder mais. Porém, como após 7 sessões o desempenho não melhorou, pensou-se que seria necessário usar uma punição mais eficaz das respostas. Assim, sempre que o pombo dava a 6.<sup>a</sup> bicada, seguiam-se 10s no escuro (ou seja, não só era retirada a luz da tecla, mas a luz da caixa também era apagada). Ao longo de 5 sessões a proporção de ensaios não reforçados aumentou, sendo que a maior parte dos ensaios nem eram iniciados. Assim, voltou-se ao procedimento original, e ao invés do limite máximo de 5 respostas, utilizou-se um limite de 9 respostas.

*(c.2) Ajustamento*

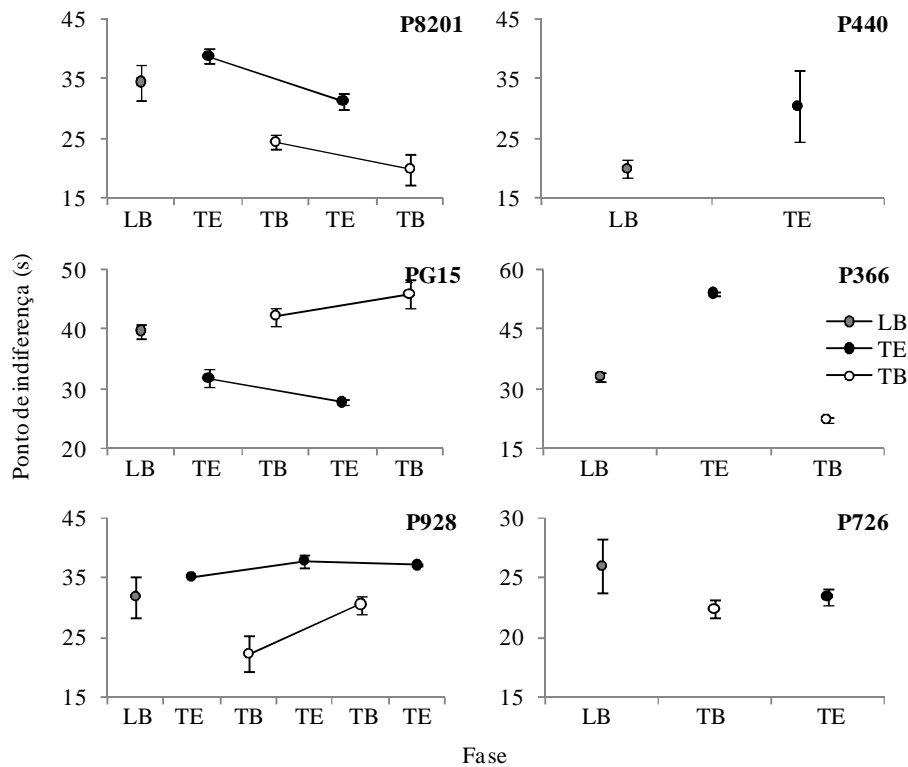
O procedimento na condição de ajustamento da fase Taxa Baixa foi idêntico ao correspondente da Linha de Base, com a exceção de que, durante o atraso da alternativa padrão, eram permitidas no máximo 5 respostas para que fosse entregue o reforço de 2s.

## **RESULTADOS E DISCUSSÃO**

Antes de mais é importante assinalar que a experiência teve que ser interrompida pois iniciaram-se obras no laboratório. Assim, apenas um pombo (P928) completou todas as fases da experiência (a fase LB, três vezes a Fase TE e duas vezes a fase TB). Para além da fase LB, dois pombos (P8201 e PG15) completaram duas vezes as fases TE e TB, dois pombos (P366 e P726) completaram cada fase apenas uma vez e por fim, o P440 completou apenas a fase TE.

O desenho experimental da Experiência 1b incluiu uma fase inicial de LB, a qual teve duas funções: 1) expor os animais ao procedimento mais simples (sem exigência de respostas), e 2) determinar qual a fase seguinte com base na taxa de resposta durante a LB, como foi explicado no procedimento. De seguida, as fases experimentais seguiram um delineamento  $A_1B_1A_2B_2A_3$  com o propósito de eliminar possíveis efeitos de ordem. Assim, os resultados seriam calculados pela média das fases A precedidas por fases B ( $A_2$  e  $A_3$ ) e a média das fases B precedidas por fases A ( $B_1$  e  $B_2$ ). Contudo, visto que apenas o P928 completou a experiência, o agrupamento de dados só foi feito desta maneira para este pombo. Para os restantes pombos que replicaram as fases (P8201 e PG15), nas análises efectuadas agruparam-se os dados das fases iguais.

Para cada condição de ajustamento, a média do atraso ajustável das 5 sessões que satisfizeram o critério de estabilidade foi usada como ponto de indiferença. A tabela 3 contém os valores para cada pombo e em cada fase e os mesmos dados são apresentados em forma de gráfico na figura 6.



**Figura 6-** Pontos de indiferença para cada pombo, para cada fase. LB- Linha de Base; TE- Taxa Elevada; TB- Taxa Baixa. As barras de erro indicam o desvio padrão para cada fase.

No painel esquerdo da figura 6 encontram-se os resultados dos três pombos que repetiram pelo menos uma vez cada fase (TE e TB), e no painel direito encontram-se os restantes pombos. O marcador preenchido a cinzento representa o ponto de indiferença na LB, os círculos preenchidos a preto as fases TE e os círculos preenchidos a branco as fases TB. As fases repetidas estão unidas por uma linha (ver painel da esquerda). Se não existissem efeitos de ordem e a replicação de uma fase fosse perfeita, as linhas que unem fases iguais deveriam ter declive de aproximadamente zero. Para as fases TE os declives das rectas de regressão foram de -1.9, -1.0 e 0.3 (P8201, PG15 e P928, respectivamente) e para as fases TB os declives obtidos foram de -1.2, 0.9 e 2.1 (P8201, PG15 e P928, respectivamente). Assim, os efeitos de ordem não parecem ser sistemáticos, quer entre pombos, quer para o mesmo pombo nas duas fases.

Antes de compararmos os pontos de indiferença obtidos nas fases TE e TB, vejamos a seguinte ordem de ideias para compreender a relação entre os pontos de indiferença obtidos e o valor do reforço.

É importante lembrar que em todas as fases a alternativa ajustável foi um atraso simples (FT) sem exigência de respostas, e a única diferença no procedimento entre as fases TE e TB foi a taxa de resposta requerida durante o atraso padrão: uma taxa elevada na fase TE e uma taxa baixa na fase



TB. Assim, qualquer diferença encontrada entre as fases TE e TB, dever-se-á, à partida, à diferença na taxa de resposta exigida em cada uma delas.

Em cada fase foi encontrado um ponto de indiferença, isto é, a duração do atraso ajustável que torna esta alternativa equivalente à alternativa padrão em termos do seu valor subjectivo. A partir da equação 1, e assumindo que no ponto de indiferença o valor das alternativas é igual, temos

$$V_{pTE} = V_{aTE} \quad , \text{ na fase TE e} \quad (\text{equação 4})$$

$$V_{pTB} = V_{aTB} \quad , \text{ na fase TB} \quad (\text{equação 5})$$

em que  $V_{pTE}$  e  $V_{aTE}$  são os valores das alternativas padrão e ajustável na fase TE e  $V_{pTB}$  e  $V_{aTB}$  são os valores das alternativas padrão e ajustável na fase TB.

Assim, a questão que se coloca é: será que a exigência de uma taxa de resposta elevada diminui o valor do reforço obtido, quando comparado com a exigência de uma taxa de resposta baixa? Ou seja, qual é a relação entre as alternativas padrão das fases,  $V_{pTE}$  e  $V_{pTB}$ ? Existem duas possibilidades: ou as alternativas têm o mesmo valor ( $V_{pTE}=V_{pTB}$ ) ou valores diferentes ( $V_{pTE}<V_{pTB}$  ou  $V_{pTE}>V_{pTB}$ ). Vejamos como estas possibilidades se manifestam nos pontos de indiferença, isto é, nas durações dos atrasos ajustáveis.

- (a) Se  $V_{pTE}=V_{pTB}$  é porque a exigência de uma taxa de resposta elevada não muda o valor do reforço. Assim, tendo em conta que  $V_{pTE}=V_{pTB}$  e a partir das equações 4 e 5:

$$V_{pTE} = V_{pTB} \Leftrightarrow V_{aTE} = V_{aTB} \Leftrightarrow \frac{A_{aTE}}{1+K_{aTE}D_{aTE}} = \frac{A_{aTB}}{1+K_{aTB}D_{aTB}} \quad (\text{equação 6})$$

Como ambas alternativas ajustáveis dão 6s de reforço,  $A_{aTE} = A_{aTB}$ . E porque em ambas as alternativas ajustáveis o programa de reforço é um FT, assume-se  $K_{aTE} = K_{aTB}$  pelo que se segue:

$$V_{pTE} = V_{pTB} \Rightarrow D_{aTE} = D_{aTB} \quad (\text{equação 7})$$

Ou seja, se um reforço obtido após um atraso com exigência de muitas respostas (fase TE) tiver o mesmo valor que um reforço obtido após o mesmo atraso mas com menos respostas (fase TB), os pontos de indiferença das fases TE e TB ( $D_{aTE}$  e  $D_{aTB}$ ) deveriam ser semelhantes.

- (b) Se  $V_{pTE}<V_{pTB}$  é porque a exigência de uma taxa elevada desvaloriza o reforço. Assim, tendo em conta que  $V_{pTE}<V_{pTB}$  e a partir das equações 4 e 5:

$$V_{pTE} < V_{pTB} \Leftrightarrow V_{aTE} < V_{aTB} \Leftrightarrow \frac{A_{aTE}}{1+K_{aTE}D_{aTE}} < \frac{A_{aTB}}{1+K_{aTB}D_{aTB}} \quad (\text{equação 8})$$

Como ambas alternativas ajustáveis dão 6s de reforço,  $A_{aTE} = A_{aTB}$ . E porque em ambas as alternativas ajustáveis o programa de reforço é um FT, assume-se  $K_{aTE} = K_{aTB}$  pelo que se segue:

$$V_{pTE} < V_{pTB} \Rightarrow D_{aTE} > D_{aTB} \quad (\text{equação 9})$$

Ou seja, se um reforço obtido após um atraso com exigência de muitas respostas (fase TE) tiver *menos* valor que um reforço obtido após o mesmo atraso mas com menos respostas (fase TB), os pontos de indiferença nas fases TE ( $D_{aTE}$ ) deveriam ser *maiores* que os pontos de indiferença das fases TB ( $D_{aTB}$ ).

Para a última possibilidade em que  $V_{pTE} > V_{pTB}$ , o raciocínio a aplicar é o inverso. Se  $V_{pTE} > V_{pTB}$  é porque a exigência de uma taxa elevada valoriza o reforço. Assim, seguindo o mesmo raciocínio das equações 8 e 9, conclui-se que:

$$V_{pTE} > V_{pTB} \Rightarrow D_{aTE} < D_{aTB} \quad (\text{equação 10})$$

Ou seja, se um reforço obtido após um atraso com exigência de muitas respostas tiver *mais* valor que um reforço obtido após o mesmo atraso mas com menos respostas, os pontos de indiferença nas fases TE deveriam ser *menores* que os pontos de indiferença das fases. Dito de outra forma, quanto menor valor tiver a alternativa padrão numa dada fase, maior o ponto de indiferença dessa fase.

Voltando à figura 6, verifica-se que para quatro pombos (P8201, P928, P366 e P726) todos os pontos de indiferença da fase TE são superiores aos pontos da fase TB. O pombo P440 não concluiu a condição de ajustamento da fase TB, por isso não se pode tirar qualquer conclusão para este sujeito. Por fim, o pombo PG15 apresenta um padrão oposto aos restantes animais: os pontos de indiferença nas fases TB foram superiores aos da fase TE. As razões que poderão explicar esta preferência por uma taxa de resposta elevada em detrimento de uma taxa de resposta baixa serão discutidas mais adiante.

Para comparar as fases TE e TB, para os pombos do painel da esquerda agruparam-se os pontos de indiferença das condições repetidas, (sendo que para o P928 excluiu-se a primeira fase TE). Para cada pombo do painel da esquerda efectuou-se uma Análise de Variância (ANOVA) de medidas repetidas com a fase (TE e TB) e a ordem (primeira e segunda) como factores. As análises mostraram um efeito principal da fase: para os pombos P8201 [ $F(1,19)=490.40$ ,  $p<.001$ ] e P928 [ $F(1,19)=525.81$ ,  $p<.001$ ], os pontos de indiferença foram superiores nas fases TE e para o pombo PG15 [ $F(1,19)=1256.25$ ,  $p<.001$ ] de indiferença foram superiores nas fases TB. Encontrou-se ainda um efeito de interacção entre os factores para os três pombos [ $F(1,19)= 26.68$ ,  $p<.001$  para P8201;  $F(1,19)= 2738.78$ ,  $p<.001$  para PG15 e  $F(1,19)= 59.34$ ,  $p<.001$  para P928], mas as variações não foram sistemáticas. O efeito da ordem apenas foi significativo os pombos P8201 [ $F(1,19)=1198.29$ ,  $p<.001$ ] e P928 [ $F(1,19)=84.72$ ,  $p<.001$ ]: para o P8201 houve uma diminuição geral dos pontos de indiferença ao longo da experiência, e para o P928 houve um aumento dos pontos de indiferença.

Para comparar as fases TE e TB para os pombos P366 e P726, efectuaram-se testes t de Student para amostras emparelhadas. Para ambos os pombos, a diferença entre os pontos de

indiferença nas fases TE e TB foi estatisticamente significativa [ $t(19)=130.75$ ,  $p<.001$  para P366 e  $t(19)=7.96$ ,  $p<.001$  para P726].

Os pontos de indiferença obtidos em cada fase mostram que a maioria dos animais apresentou uma preferência por atrasos com exigência de poucas respostas em detrimento de atrasos com exigência de muitas respostas. É de salientar que o procedimento aplicado ao pombo P726 apresenta duas diferenças em relação aos restantes animais: a ordem das fases, pois a seguir à LB seguiu-se uma fase TB (e não uma fase TE como para os restantes animais) e o limite da taxa de resposta na fase TB, que foi de 9 respostas e não 5 como para os restantes animais. O facto de a diferença na taxa de resposta exigida entre as fases não ter sido tão grande como para os restantes animais, poderá ser uma das razões para as diferenças nos pontos de indiferença entre as fases também não serem tão grandes no P726.

Se compararmos as exigências comportamentais impostas durante os atrasos padrão da fase LB com os atrasos das fases experimentais (TE e TB), verificamos que na primeira não existe qualquer restrição na taxa de resposta e nas outras fases há uma imposição de limites ao comportamento de bicar. Dada essa restrição, seria razoável pensar que, em relação à LB, em ambas as fases TE e TB o reforço seria desvalorizado, ainda que numa fase pudesse haver uma maior desvalorização que na outra fase.

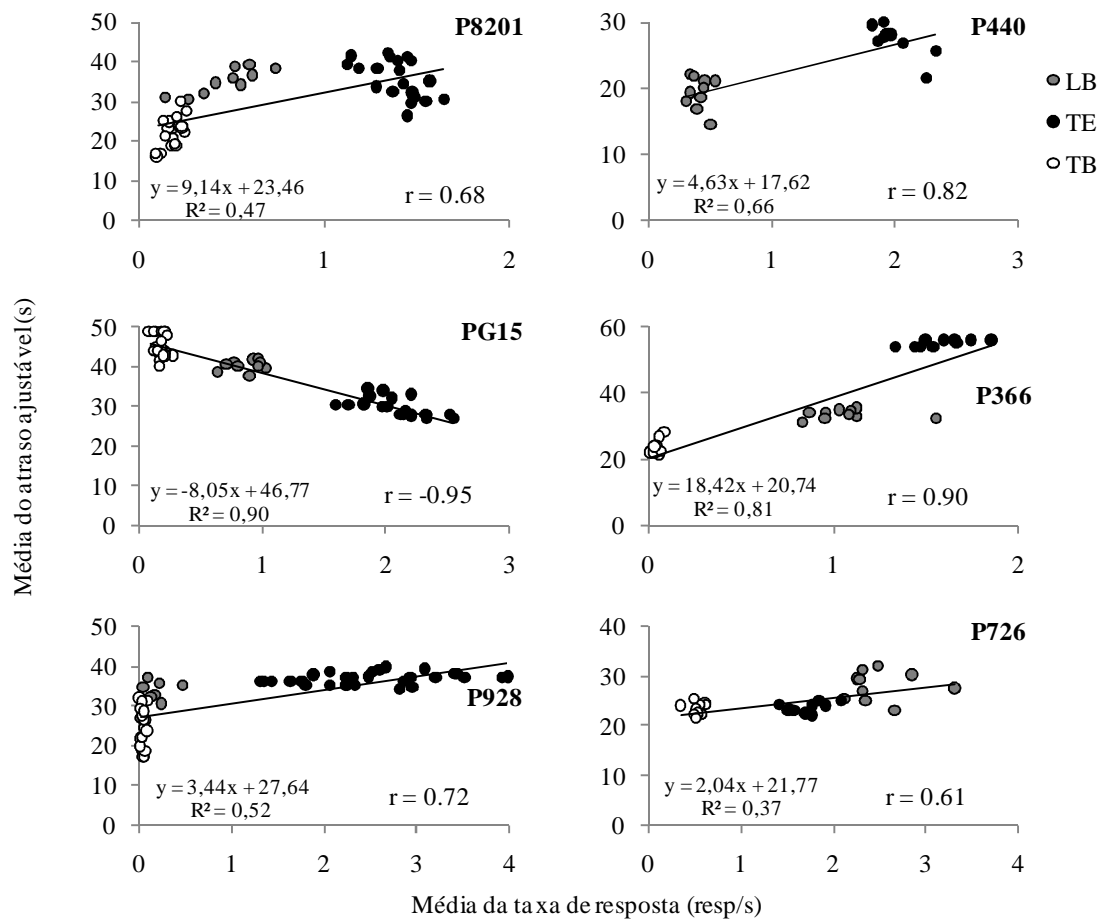
Relativamente a esta possibilidade, o dado que podemos encontrar na figura 6 é que, para todos os pombos que têm pontos de indiferença superiores na fase TE (em relação a TB), 6 dos 8 pontos de indiferença das fases TE estão acima do ponto correspondente da LB e todos os pontos de indiferença das fases TB ( $n=6$ ) estão abaixo do ponto correspondente da LB. Para o pombo PG15, que tem pontos de indiferença superiores na fase TB (em relação a TE), o padrão é o oposto: todos os pontos das fases TB estão acima da Linha de Base e todos os pontos das fases TE estão abaixo da LB. Os resultados parecem sugerir que para a maior parte dos pombos, responder a uma taxa baixa (TB) é preferido a não ter limitações de respostas (LB), e para o PG15, responder a uma taxa elevada (TE) é preferido a um atraso sem limitações de respostas (LB).

Estes padrões mostram que, pelo menos nesta experiência, a ideia de que a imposição de limites ao comportamento desvaloriza uma opção não é necessariamente verdade. Por conseguinte, não faz sentido agrupar as fases TE e TB numa categoria do tipo “há limitações no comportamento” e comparar com a fase LB (categoria “não há limitações no comportamento”).

Como explicar então a preferência pela fase TB em relação à LB, ou a preferência pela fase TE em relação à LB (para PG15). Uma possibilidade é que haja uma variável que seja comum às várias fases e que explique as preferências.

Na figura 7 encontram-se os gráficos da média do atraso ajustável em função da taxa de resposta durante o atraso padrão. Cada ponto representa uma sessão e são apresentados os dados das

últimas 10 sessões de ajustamento de cada fase. Os marcadores preenchidos a cinzento referem-se à fase LB, os marcadores a preto à TE e os marcadores brancos referem-se à fase TB.



**Figura 7** - Média do atraso ajustável em função da média da taxa de resposta no atraso padrão em cada sessão, nas últimas 10 sessões de ajustamento de cada fase. As rectas são as rectas de regressão com melhor ajustamento. Cada painel mostra a equação das rectas de regressão e a percentagem de variância explicada pela recta. Cada painel contém ainda o valor das correlações de Pearson, todos com valor de  $p < .001$ .

Para todos os pombos, com a excepção do P726, todos os pontos da LB (a cinzento) encontram-se entre as outras fases, ou seja, a taxa de resposta durante a linha de base foi inferior à taxa de resposta durante a TE mas superior à TB. O P726 apresenta taxas de resposta mais baixas na LB comparativamente à fase TE (e por este motivo é que a ordem das fases foi diferente para este pombo).

Considerando então as três fases (TB, LB e TB), haverá alguma relação entre a taxa de resposta nestas fases e os pontos de indiferença obtidos (ainda que na fase LB a taxa de resposta não sofra qualquer restrição e nas outras fases sim)?

Para responder a esta questão, para cada pombo calculou-se o coeficiente de correlação de Pearson, para verificar a associação entre a taxa de resposta e o ponto de indiferença. O resultado é apresentado em cada gráfico. Em valores absolutos, o coeficiente de correlação variou entre 0.61 e

0.95. Para além disso, verifica-se que a correlação entre as variáveis é positiva para a maior parte dos animais: quanto maior a taxa de resposta maior o ponto de indiferença. Para o PG15, a relação entre as variáveis é inversa.

Tomemos agora os resultados da figura 7 não como os resultados de três fases distintas (LB, TE e TB), mas como resultados de um contínuo onde a taxa de resposta no atraso padrão varia e para cada taxa de resposta no atraso padrão há uma duração do atraso ajustável correspondente. Consideremos que, pelo próprio procedimento de ajustamento, nas últimas 10 sessões de ajustamento de uma fase os valores das alternativas são aproximadamente iguais:

$$V_a \approx V_p \Leftrightarrow \frac{A_a}{1+K_a D_a} \approx \frac{A_p}{1+K_p D_p} \quad (\text{equação 11})$$

em que os subscritos a e p referem-se às alternativas ajustável e padrão, respectivamente. Como na alternativa padrão o atraso para o reforço foi sempre o mesmo (10s), sendo que o que variou foi a taxa de resposta, em  $D_p$  poder-se-á usar não o atraso entre a escolha e o reforço, mas a taxa de resposta durante o atraso (sigla  $TR_p$  a substituir  $D_p$ ).

Assim, colocando a duração do atraso ajustável em função da taxa de resposta na alternativa padrão temos:

$$D_a \approx \frac{A_a K_p}{A_p K_a} TR_p + \frac{A_a - A_p}{A_p K_a} \quad (\text{Equação 12})$$

Esta equação implica que se  $D_a$  for graficado em função de  $TR_p$ , a função deverá ser uma recta com intercepção no eixo das ordenadas superior a zero e declive também superior a zero (com  $A_a$  superior a  $A_p$ ). Apesar de ser possível que  $K_a$  seja diferente de  $K_p$  pois os programas de reforço são diferentes (Mazur, 1986), os dados obtidos não permitem testar essa hipótese.

Na figura 7, para os dados de cada pombo foi ajustada uma recta, e a sua equação é apresentada no gráfico. Para todos os pombos verifica-se que a intercepção no eixo dos yy é superior a zero, e que para 5 dos 6 pombos o declive é superior a zero.

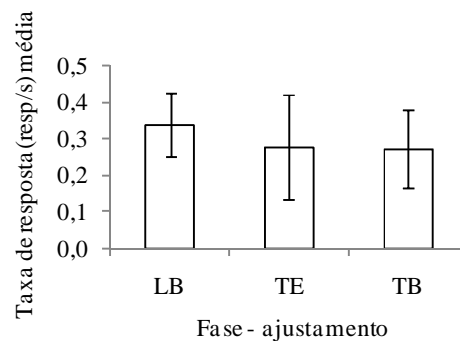
Assim, podemos afirmar que para a maior parte dos pombos existe uma relação positiva entre a taxa de resposta durante um atraso e o valor do reforço obtido no fim desse atraso. Essa relação parece ser bem descrita através dos pontos de indiferença, por meio da função hiperbólica de desconto. É necessário, contudo, enfatizar que esta é apenas uma análise exploratória, já que a experiência não foi planeada para explicar este fenómeno. De modo rigoroso, dever-se ia encontrar um ponto de indiferença para cada taxa de resposta na alternativa padrão (note-se que seria muito complicado exigir sempre a mesma taxa de resposta a um pombo). Isto é, para cada taxa de resposta seria calculado um número (o ponto de indiferença), calculado pela média do atraso ajustável nas últimas 5 sessões. Ora, o que se apresenta na figura 7 não é o ponto de indiferença para cada taxa de resposta, mas sim o valor do atraso de uma sessão para a taxa de resposta dessa mesma sessão, considerando-se as últimas 10 sessões de ajustamento. Em suma, independentemente da imposição de limites (máximos ou mínimos)

de respostas durante o atraso padrão, o valor do seu reforço é determinado pela taxa de resposta obtida durante esse atraso. Quanto maior a taxa de resposta durante o atraso padrão, maior o ponto de indiferença, ou seja, menor o valor subjectivo do reforço obtido.

Para se poder garantir que as diferenças encontradas entre as fases TE e TB se devem apenas à variação na taxa de resposta durante o atraso padrão, é importante verificar que estas fases não diferem noutras dimensões que possam afectar a preferência por uma das alternativas. De seguida serão discutidas duas variáveis que poderiam afectar a preferência por uma das alternativas: diferenças entre fases na taxa de resposta durante o atraso ajustável e diferenças de dificuldade em aprender as exigências de respostas nas fases TE e TB.

Em primeiro lugar, é importante verificar que não existem diferenças na taxa de resposta da alternativa ajustável. Em nenhuma fase são exigidas respostas durante o atraso da alternativa ajustável, mas durante este atraso o animal pode responder na tecla acesa, sem que haja qualquer consequência. Variações entre as fases na taxa de resposta nesta tecla poderiam induzir um viés por alguma das fases.

A figura 8 apresenta a média da taxa de resposta no atraso ajustável nas últimas 10 sessões de ajustamento de cada fase. O pombo P440 foi excluído da análise já que não completou nenhuma fase TB.

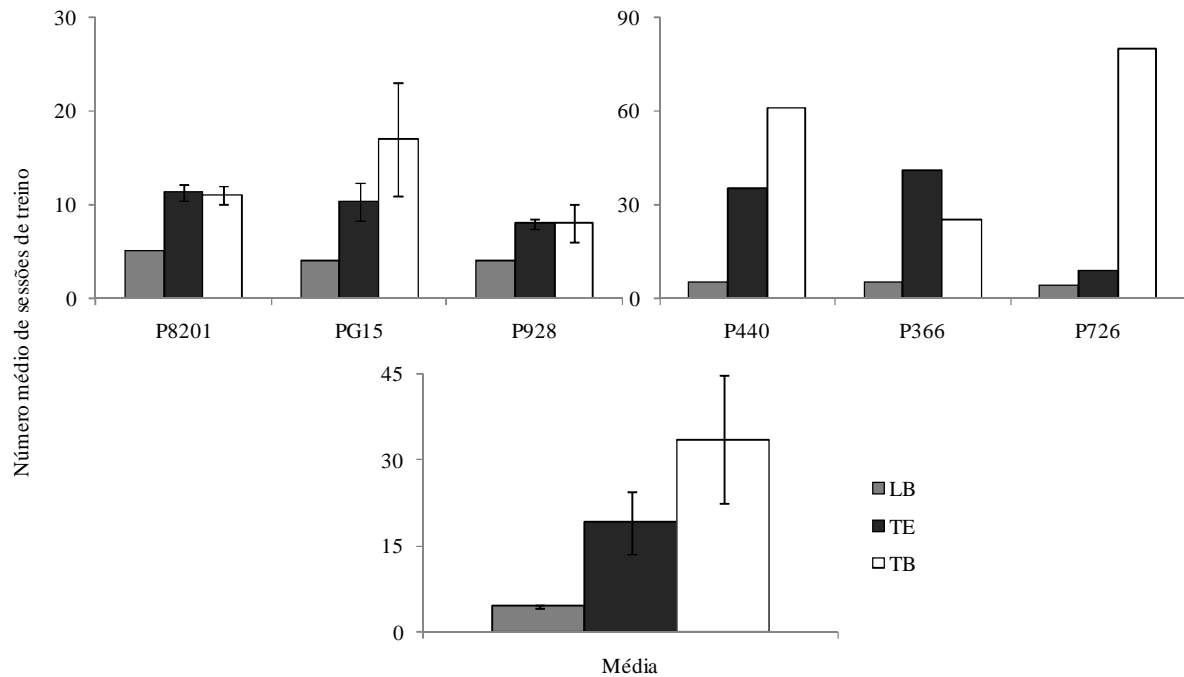


**Figura 8** - Média da taxa de resposta no atraso ajustável nas últimas 10 sessões de ajustamento de cada fase. LB- Linha de Base; TE- Taxa Elevada; TB- Taxa Baixa. As barras de erro indicam o erro padrão da média para cada fase.

Realizou-se uma ANOVA de medidas repetidas para comparar a taxa de resposta durante o atraso ajustável nas fases LB, TE e TB. Não houve efeito da fase na taxa de resposta,  $F(2,8)=0.42$ ,  $p=0.67$ . Assim, podemos excluir a possibilidade de diferenças na taxa de resposta na alternativa ajustável poderem explicar os pontos de indiferença obtidos.

Em segundo lugar, podemos ainda pensar que uma fase foi mais difícil de ser aprendida que outra, e que essa diferença poderia influenciar a preferência dos animais. Para inferir acerca da dificuldade da tarefa, podem-se analisar dois resultados: o número de sessões de treino e a proporção de ensaios não reforçados no ajustamento de cada fase.

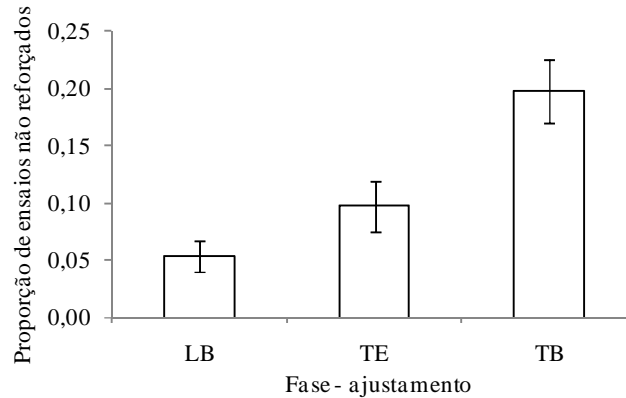
A figura 9 mostra, para cada pombo, a média do número de sessões de treino em cada fase (confrontar com dados da tabela 3). Verifica-se que para 5 pombos (a exceção é o P366) são necessárias tantas ou mais sessões na condição de treino da fase TB do que na fase TE. Ainda assim, em termos da média do grupo, uma ANOVA para medidas repetidas não encontrou um efeito da fase no número de sessões de treino:  $F(2,10) = 3.75, p = .06$ .



**Figura 9** - Média do número de sessões por fase de treino. LB- Linha de Base; TE- Taxa Elevada; TB- Taxa Baixa. As barras de erro indicam o erro padrão da média para cada fase.

Visto que, em termos do número de sessões de treino, parece ser mais difícil ensinar um pombo responder a uma taxa baixa, é necessário verificar se a proporção total de ensaios não reforçados durante o ajustamento foi superior nesta fase. Uma elevada proporção de ensaios não reforçados numa dada fase poderia enviesar os sujeitos a escolherem a tecla ajustável, já que nesta opção estava apenas em vigor um atraso simples.

A figura 10 apresenta a média de grupo da proporção de ensaios não reforçados em cada fase na condição de ajustamento (o sujeito P440 não foi incluído).



**Figura 10-** Média da proporção de ensaios não reforçados em cada fase na condição de ajustamento. LB- Linha de Base; TE- Taxa Elevada; TB- Taxa Baixa. As barras de erro indicam o erro padrão da média para cada fase.

O padrão dos dados é semelhante ao padrão anterior da figura anterior: a proporção de ensaios não reforçados foi superior na fase TB, seguido da fase TE e por fim a LB, onde apenas 5% dos ensaios não foram reforçados. Uma ANOVA de medidas repetidas mostrou um efeito da fase  $F(2,8)=10.39$ ,  $p<.05$ , porém, apenas a diferença entre a LB e a TE atingiu significância estatística ao nível de 0.05.

Em suma, as análises efectuadas permitem concluir que as diferenças encontradas entre as fases TE e TB se devem apenas à variação na taxa de resposta durante o atraso padrão e não a diferenças entre fases na taxa de resposta durante o atraso ajustável ou a diferenças de dificuldade em aprender as exigências de respostas nas fases TE e TB.

Até agora, nos dados apresentados, não existe nenhuma evidência para o pombo PG15 apresentar um padrão de preferências oposto aos restantes animais (ver figura 6). O seu desempenho foi semelhante aos restantes sujeitos na: taxa de resposta na linha de base (figura 7), na taxa de resposta na tecla ajustável (referente à figura 8, apesar dos dados individuais não serem apresentados), no número médio de sessões de treino em cada fase (figura 9) ou ainda na proporção de ensaios não reforçados em cada fase (referente à figura 10, apesar dos dados individuais não serem apresentados).

Assim, uma última variável que poderia explicar o padrão de preferências do PG15 é o padrão de respostas durante o atraso padrão. Para que o reforço fosse entregue, na fase TE eram impostas pelo menos 10 respostas durante os 10s do atraso padrão e durante a fase TB não eram permitidas mais de 5 respostas durante o mesmo atraso. Porém, não foi imposta nenhuma distribuição das respostas durante o atraso padrão (ex., podiam estar todas concentradas no início do atraso, no fim do atraso, etc.).

Para compreender de que modo o padrão de respostas durante o atraso padrão pode influenciar o valor dessa alternativa, é necessário apresentar uma versão mais geral do modelo de desconto hiperbólico (Mazur & Kralik, 1990). Esta extensão do modelo foi proposta para poder lidar com



situações em que após a escolha se seguem eventos de valor negativo (eventos aversivos), como bicar numa tecla.

O modelo de Mazur e Kralik (1990) assume que  $V$ , o valor de uma alternativa, depende de todos os eventos que ocorrem após a resposta de escolha. A contribuição de cada evento para o valor  $V$  depende: (a) da qualidade do evento, (b) da sua duração e (c) da proximidade desse evento do momento da escolha. O modelo afirma que a cada evento, quer seja uma resposta (ex. bicar) quer seja um estímulo (ex. comida), pode ser atribuído um número  $Q$  que representa a qualidade desse evento. Assim,  $Q$  poderá ser um número positivo para comer, um número negativo para um choque e zero para um intervalo de tempo. Tudo o resto sendo igual, o sujeito tenderá a escolher a alternativa com valores positivos e elevados de  $Q$ , em oposição a valores mais baixos ou negativos.

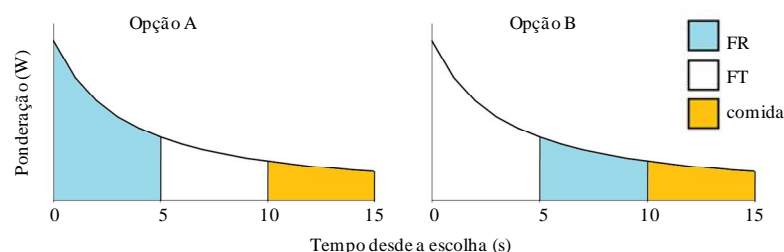
O modelo assume que quanto mais distante no tempo, desde a escolha, estiver um evento, menor a sua contribuição para  $V$ . Assumamos  $t$  como o tempo desde a resposta de escolha. A contribuição do evento que ocorre no tempo  $t$  para o valor  $V$ , é igual a  $Q_t W_t$ , onde  $Q_t$  é a qualidade do evento no tempo  $t$  e  $W_t$  é um factor de ponderação que diminui à medida que  $t$  aumenta (Mazur & Kralik, 1990):

$$W_t = \frac{1}{1+Kt} \quad \text{(Equação 13)}$$

O produto  $Q_t W_t$  é a contribuição instantânea do evento que ocorre ao tempo  $t$  para o valor  $V$ , o valor de uma alternativa. Para calcular o valor total de um evento estendido no tempo, é necessário calcular o integral para a duração desse evento:

$$V = \int (Q_t W_t) dt = \int \left( \frac{Q_t}{1+Kt} \right) dt \quad \text{(Equação 14)}$$

Tomemos o seguinte exemplo. Um pombo tem de escolher 2 opções (ver figura 11). Na opção A tem de responder numa tecla (por ex. FR<sup>5</sup>) durante 5s, segue-se um atraso simples de 5s (FT) e por fim é entregue comida durante 5s. Na opção B, a sequência de eventos é: 5s de atraso, 5s de respostas e 5s de comida. Em ambas as opções o total de tempo quer a responder (FR), quer a esperar (FT), quer a comer é igual: será o animal indiferente entre as alternativas?



**Figura 11-** Ilustração da ponderação de cada evento para o valor de uma alternativa em função da sua distância temporal do momento da escolha.

<sup>5</sup> Para tornar o exemplo o mais simples possível, vamos assumir que o pombo responde a uma taxa constante de uma resposta por segundo.

Assumindo  $K=1$ ,  $Q_{FR}=-1$ ;  $Q_{FT}=0$  e  $Q_{comida}=100$ , a partir da equação 14 temos:

*Opção A*

$$V_A = \int_0^5 \frac{-1}{1+t} + \int_5^{10} \frac{0}{1+t} + \int_{10}^{15} \frac{100}{1+t}$$

$$V_A = -1 \ln(1+t) \Big|_0^5 + 0 \ln(1+t) \Big|_5^{10} + 100 \ln(1+t) \Big|_5^{10}$$

$$V_A = -1.79 + 0 + 37.47 = 35.7$$

*Opção B*

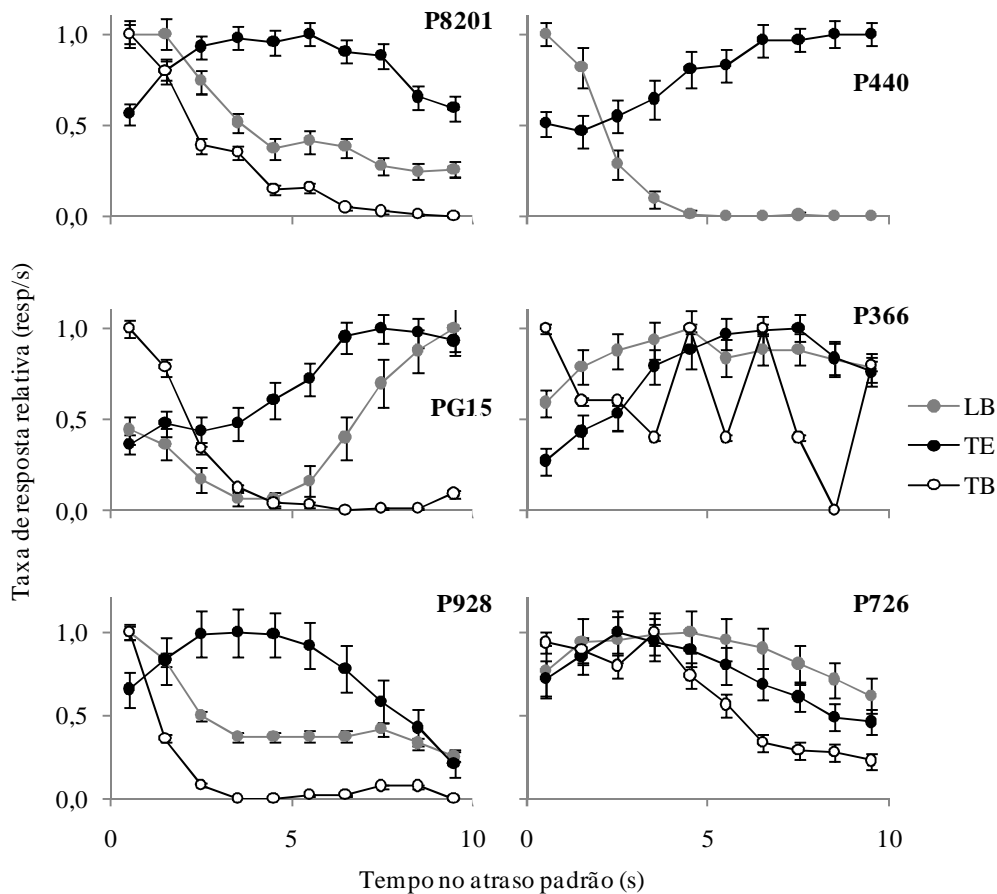
$$V_B = \int_0^5 \frac{0}{1+t} + \int_5^{10} \frac{-1}{1+t} + \int_{10}^{15} \frac{100}{1+t}$$

$$V_B = 0 \ln(1+t) \Big|_0^5 + (-1) \ln(1+t) \Big|_5^{10} + 100 \ln(1+t) \Big|_5^{10}$$

$$V_B = 0 - 0.61 + 37.47 = 36.9$$

Assim, verifica-se que a opção B tem um maior valor e por isso esta opção seria a escolhida. De uma maneira geral, quanto mais longe do momento da escolha estiver um evento aversivo ( $Q < 0$ ; neste caso o FR), menos ponderação esse evento vai ter no valor da alternativa. Os estudos de procrastinação apoiam esta hipótese ao encontrarem que os pombos preferem um requerimento de resposta maior mas mais atrasado, quando comparado com um requerimento de menos respostas mas menos atrasado (Mazur, 1996, 1998).

Será que no presente estudo, a distribuição das respostas durante o atraso padrão teve um efeito no valor do reforço? Na figura 12 encontra-se a média normalizada da taxa de resposta durante o atraso padrão em cada uma das fases. Comparando os padrões de distribuição da taxa de resposta das fases TE e TB, parece que o PG15 apresenta padrões diferentes dos restantes animais. Na fase TB para todos os pombos (apesar de para o P366 não ser tão evidente), a taxa de resposta começa elevada no início do atraso e vai diminuindo à medida que o tempo passa. Já na fase TE, com a exceção dos pombos P440 e PG15, as taxas de resposta tendem a aumentar até cerca de meio do atraso e na parte final do atraso a taxa de resposta começa a diminuir. Ou seja, para o pombo PG15, na fase TB a taxa de resposta é superior no início do intervalo e depois começa a diminuir, e na fase TE a taxa de resposta começa mais baixa e vai aumentando à medida que o tempo vai passando. Para além disso, para este pombo padrão de respostas na LB está mais próximo do padrão da fase TE do que do padrão TB.



**Figura 12** - Média normalizada da taxa de resposta nas últimas 5 sessões de ajustamento cada fase. LB- Linha de Base; TE- Taxa Elevada; TB- Taxa Baixa. As barras de erro indicam o erro padrão da média para cada fase.

Numa tentativa de conseguir compreender porque é que o PG15 apresentou um padrão de preferências oposto aos restantes pombos, foi aplicado o modelo de decaimento hiperbólico (Mazur & Kralik, 1990). Para tal, assumiu-se  $K=1$ ,  $Q_{comida}=100$  e  $Q_{FR}<0$ . Como na fase TE a taxa de resposta média foi sempre maior que na fase TB, qualquer aplicação do modelo em que  $Q_{FR}$  estivesse relacionado com a taxa de resposta absoluta (quanto maior a taxa, menor o  $Q_{FR}$ ) iria prever preferência por valores superiores de  $Q_{FR}$ , ou seja, pela fase TB. Assim, a aplicação com êxito deste modelo teria de presumir que entre as fases, o factor determinante para a preferência não seria a taxa de resposta, mas sim a distribuição das respostas durante o atraso padrão.

Na primeira tentativa considerou-se que quanto maior a taxa de resposta relativa (TR), menor o  $Q_{FR}$ , sendo esta uma relação directamente proporcional ( $Q_{FR}=-TR$ ). Numa segunda tentativa considerou-se que quanto maior a taxa de resposta relativa menor o valor de  $Q_{FR}$ , sendo a relação hiperbólica ( $Q_{FR}=(\frac{1}{1+kTR})-1$ ). Nenhuma destas tentativas conseguiu prever pontos de indiferença superiores na fase TE para todos os pombos (excepto PG15) e inferiores na fase TB para o PG15. Deste modo não foi possível encontrar uma só explicação para os resultados de todos os animais.

Tipicamente, é esperado que animais humanos e não humanos escolham, de entre as alternativas disponíveis, a opção menos custosa. Hull (1943) enunciou a lei do menor esforço da seguinte forma:

“If two or more behavioral sequences, each involving a different amount of energy consumption or work, have been equally well reinforced an equal number of times, the organism will gradually learn to choose the less laborious behavior sequence leading to the attainment of the reinforcing state of affairs” (p. 294).

Contudo, alguns estudos demonstram que a imposição de um custo/esforço nem sempre diminui o valor de uma recompensa, podendo mesmo aumentá-lo.

Por exemplo, nas experiências de autocontrole o sujeito deve escolher entre duas alternativas que diferem tanto no custo como na magnitude da recompensa. Por outro lado, nas experiências que reportam uma preferência pela alternativa que envolve um custo maior, a escolha entre as recompensas não é feita antes do esforço ser dispendido (Clement et al., 2000; Zentall & Singer, 2007a para discussão). Por exemplo, Clement et al. (2000) treinaram pombos em dois tipos de ensaios. Em alguns ensaios uma bicada numa tecla branca mudava-a para vermelho, e depois de uma resposta na tecla vermelha era entregue comida; nos outros ensaios, 20 respostas na tecla branca mudavam-na para verde e depois de uma resposta na tecla verde era entregue comida. Após o treino, era dada aos pombos a escolha entre as teclas verde e vermelha, e eles preferiam a verde. Ou seja, os pombos mostraram uma preferência pelo estímulo que se seguiu a mais esforço. Apesar de inovadores, estes resultados são controversos porque alguns estudos conseguiram replicá-los (Zentall & Singer, 2007a, 2007b; Kacelnik & Marsh, 2002; Aw et al., 2011) mas outros não (Arantes & Grace, 2008; Vasconcelos et al., 2007).

Outro conjunto de estudos que reportam uma preferência por situações de elevado custo medem a preferência após um treino específico em situações em que o custo elevado (ou o custo baixo) é reforçado (Eisenberger et al., 1989; Eisenberger, 1992 para revisão). Por exemplo, no estudo de Eisenberger et al. (1989) ratos receberam comida num corredor após completarem cada corrida (grupo com reforço contínuo) ou após cada 5 corridas (grupo FR). Outros dois grupos, os grupos de controlo, receberam comida ao mesmo tempo que os grupos experimentais, independentemente das suas respostas. De seguida os ratos foram colocados numa caixa operante com duas alavancas, uma requerendo mais força para ser pressionada (alavanca de força elevada) e outra menos força para ser pressionada (alavanca de força baixa). Após uma resposta na alavanca de força elevada era entregue uma grande quantidade de comida, e após uma resposta na alavanca de força baixa era entregue uma pequena quantidade de comida. Quando foi dada a escolha entre as duas alavancas, o grupo FR mostrou maior percentagem de escolha da alavanca de força elevada que os restantes grupos. Os

autores afirmam então que o treino numa situação que exige esforço diminuiu a sua aversividade, e por isso houve uma preferência pela maior quantidade de comida, apesar de esta exigir mais esforço.

Por fim, é necessário ter em conta que os estudos que reportam uma preferência pela alternativa que envolve um maior esforço e aqueles que reportam o seguimento da lei do menor esforço diferem quer nos objectivos experimentais, quer na metodologia utilizada.

## CONCLUSÃO

Este estudo teve como principal objectivo compreender qual o efeito do esforço no valor de uma recompensa. Os estudos anteriores têm conceptualizado esforço como o número de respostas para a obtenção de um reforço. Porém, à medida que se aumenta o número de respostas também se aumenta o tempo entre a escolha e o reforço, e por isso têm sido usadas várias maneiras para controlar o atraso para o reforço. Para além disso, os estudos anteriores compararam situações experimentais em que a sequência de estímulos não era idêntica. Neste estudo operacionalizou-se esforço como a taxa de resposta: manteve-se o atraso para o reforço e variou-se o número de respostas exigidas. Deste modo pretendeu-se comparar o valor de um reforço obtido após diferentes taxas de resposta, sendo a sequência de estímulos idêntica em todas as fases.

A Experiência 1a foi constituída por três fases com diferentes taxas de resposta. Contudo, os pombos tiveram dificuldade em aprender a tarefa e a experiência foi terminada antes de se obter qualquer resultado informativo.

Na Experiência 1b efectuaram-se algumas alterações ao procedimento, que permitiram, com sucesso, que os animais aprendessem e desempenhassem bem tarefa. Nesta experiência implementam-se apenas dois níveis de taxa de resposta: Taxa Elevada e Taxa Baixa.

Estas experiências mostraram, em primeiro lugar, que o procedimento de ajustamento do atraso pode produzir pontos de indiferença sistemáticos para escolhas entre um atraso simples e um atraso durante o qual são exigidas respostas. Mais ainda, parece ser possível aplicar a equação de desconto hiperbólico (equação 1) a uma situação em que em vez do atraso, se usa como variável independente a taxa de resposta. Os pontos de indiferenças obtidos nas fases de taxa elevada e taxa baixa sugerem que para a maior parte dos animais (4 dos 5 animais), a exigência de uma elevada taxa de resposta durante o atraso para uma recompensa desvaloriza essa recompensa. O quinto pombo apresentou uma preferência pela alternativa que exige uma maior taxa de resposta. Das várias análises efectuadas não foi encontrada nenhuma variável que conseguisse explicar a preferência deste pombo. Assim, conclui-se que um só modelo não consegue dar conta da variabilidade de resultados entre os sujeitos.

Dos resultados obtidos nesta experiência ficam várias questões por responder. Por exemplo: “Seriam os resultados diferentes se se tivesse usado outro programa de reforço?”. Mais especificamente, nesta experiência exigiram-se respostas durante um atraso simples. Assim, houve alguma variabilidade na distribuição dessas respostas durante o atraso. Uma forma de diminuir a variabilidade na distribuição das respostas poderia ser, em vez de um atraso simples, usar um programa de intervalo fixo (FI- do inglês, *Fixed Interval*). Num FI, um evento (ex. ligar uma luz) marca o início de um intervalo e a primeira resposta após o fim desse intervalo é reforçada. A vantagem de usar um FI, é que a maior parte dos animais apresenta um padrão de respostas

semelhante: depois de uma pausa inicial, a taxa de resposta aumenta, atingindo o seu máximo no momento do reforço.

Em suma, existe ainda uma série de variáveis relacionadas com o esforço que necessitam de ser estudadas para melhor se compreender como é que o esforço afecta o valor de um reforço. Na verdade, como foi mencionado na introdução deste trabalho, comparando com os estudos de desconto do atraso e probabilístico, os estudos de desconto de esforço são ainda escassos.

Todavia, esta lacuna na investigação poderá ser, pelo menos em parte, explicada pela dificuldade em responder a duas questões metodológicas: (a) O que é o “esforço” e como operacionalizá-lo? e (b) Como desambiguar os efeitos do esforço e do tempo no valor de uma recompensa?

Quanto à primeira questão, sobre a operacionalização de esforço, a maior parte dos estudos refere-se a ele sem o definir. Isto provavelmente acontece porque todos percebemos o que é esforço. Porém, também percebemos que existem diferentes tipos de esforço, e esta variedade é reflectida na investigação. Assim, podemos encontrar estudos em que o esforço é: (a) um esforço mental como fazer cálculos mentais (com humanos; Botvinick et al., 2009); (b) um esforço físico como apertar com força um objecto (humanos- Mitchell, 2003), pressionar uma alavanca pesada (ratos- Chelonis et al., 1998) ou bicar numa tecla com força (pombos- Mazur & Kralik, 1990) ou (c) emitir muitas respostas (pombos- Grossbard & Mazur, 1986).

No que concerne à segunda questão, como separar os efeitos do tempo e do esforço, esta depende em parte da resposta à primeira questão, de como se operacionaliza esforço. Se o esforço for manipulado pela força que é necessária para dar uma resposta, o tempo entre a resposta e a recompensa deverá ser semelhante mesmo quando se varia a força (o esforço). Por outro lado, se o esforço for conceptualizado pelo número de respostas a emitir para a obtenção de uma recompensa, quanto maior o esforço, maior o número de respostas e, consequentemente, maior é o atraso para a recompensa. O principal contributo do presente estudo foi verificar que é possível manipular o esforço em termos do número de respostas, mantendo o atraso para o reforço constante.

O esforço é uma variável que tem sido investigada numa multiplicidade de estudos, com diferentes sujeitos, metodologias e objectivos. Esta multiplicidade reflecte, por um lado, a complexidade desta variável (pois manifesta-se de diferentes formas) e por outro lado a sua abrangência (pois está presente em diferentes situações e organismos). Para uma total compreensão de como o esforço afecta o valor de uma recompensa, é fundamental que no futuro se adopte uma abordagem abrangente que consiga integrar toda esta informação.

## REFERÊNCIAS

- Arantes, J., & Grace, R. C. (2008). Failure to obtain value enhancement by within-trial contrast in simultaneous and successive discriminations. *Learning & Behavior*, *36* (1), 1-11.
- Aw, J. M., Vasconcelos, M., & Kacelnik, A. (2011). How costs affect preferences: experiments on state dependence, hedonic state and within-trial contrast in starlings. *Animal Behaviour*, *81*, 1117-1128.
- Botvinick, M. M., Huffstetler, S., & McGuire, H. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective & Behavioral Neuroscience*, *9* (1), 16–27.
- Chelonis, J. J., Logue, A. W., Sheehy, R., & Mao J. (1998). Effects of response effort on self-control in rats. *Animal Learning & Behavior*, *26* (4), 408-415.
- Clement, T. S., Feltus, J. R., Kaiser, D. H., & Zentall, T. R. (2000). “Work ethic” in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7* (1), 100-106.
- Eisenberger, R. (1992). Learned Industriousness. *Psychological Review*, *99* (2), 248-267.
- Eisenberger, R., Weier, F., Masterson, F. A., & Theis, L. Y. (1989). Fixed-Ratio Schedules Increase Generalized Self-Control: Preference for Large Rewards Despite High Effort or Punishment. *Journal of Experimental Psychology: Animal Behavior Processes*, *15* (4), 383-392.
- Grossbard, C. L., & Mazur, J. E. (1986). A comparison of delays and ratio requirements in self-control choice. *Journal of the Experimental Analysis of Behavior*, *45*, 305-315.
- Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century- Crofts.
- Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour*, *63*, 245-250.
- Lattal, K. A. (2010). Delayed reinforcement of operant behavior. *Journal of the Experimental Analysis of Behavior*, *93*, 129-139.
- Mazur, J. E. (1984). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. Effects of delay and of intervening events on reinforcement value*. Hillsdale, NJ: Erlbaum.
- Mazur, J.E. (1986). Fixed and Variable Ratios and Delays: Further Tests of an Equivalence Rule. *Journal of Experimental Psychology: Animal Behavior Processes*, *12* (2), 116-124.



- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. Effects of delay and of intervening events on reinforcement value*. Hillsdale, NJ: Erlbaum, pp. 55-73.
- 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.
- Mazur, J. E. (1996). Procrastination by pigeons: Preference for larger, more delayed work requirements. *Journal of the Experimental Analysis of Behavior*, 65 (1), 159-171.
- Mazur, J. E. (1998). Procrastination by pigeons with fixed-interval response requirements. *Journal of the Experimental Analysis of Behavior*, 69 (2), 185-197.
- Mazur, J. E., & Kralik, J. D. (1990). Choice between delayed reinforcers and fixed-ratio schedules requiring forceful responding. *Journal of the Experimental Analysis of Behavior*, 53, 175-187.
- Mitchell, S. H. (2003). Discounting the Value of Commodities According to Different Types of Cost. In Heather, N., & Vuchinich, R. (Eds.), *Choice, Behavioral Economics, and Addiction*. Oxford: Elsevier Press, pp. 339-357.
- Neuringer, A. J., & Schneider, B. A. (1968). Separating the effects of interreinforcement time and number of interreinforcement responses. *Journal of the Experimental Analysis of Behavior*, 11 (6), 661-667.
- Rachlin, H., & Green, L. (1972). Commitment, Choice and Self-control. *Journal of the Experimental Analysis of Behavior*, 17 (1), 15-22.
- Vasconcelos, M., Urcuioli, P. J., & Lionello-DeNolf, K. M. (2007). Failure to replicate the “work ethic” effect in pigeons. *Journal of the Experimental Analysis of Behavior*, 87, 383-399.
- Zentall, T. R., & Singer, R. A. (2007a). Within-trial contrast: When is a failure to replicate not a Type I error? *Journal of the Experimental Analysis of Behavior*, 87, 401-404.
- Zentall, T. R., & Singer, R. A. (2007b). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than a less aversive event. *Journal of the Experimental Analysis of Behavior*, 88, 131-149.