



Programa aprovado pelo Conselho Superior de Ensino e Pesquisa da UFPA – Resolução 2545/98. Reconhecido nos termos das Portarias N°. 84 de 22.12.94 da Presidente da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES e No. 694 de 13.06.95 do Ministério da Educação e do Desporto. Doutorado autorizado em 1999.

Aversividade do Timeout em uma Situação de Escolha

Ysláine Lopes Silva

Belém - PA

2026



Aversividade do Timeout em uma Situação de Escolha

Yslaíne Lopes Silva

Tese apresentada ao Programa de Pós-Graduação em Teoria e Pesquisa do Comportamento como requisito parcial para a obtenção do Título de Doutor/a.

Orientador: Prof. Dr. Marcus Bentes de Carvalho Neto

Coorientador: Prof. Dr. Paulo Sérgio Dillon Soares Filho

Belém - PA

2026

Dados Internacionais de Catalogação-na-Publicação (CIP)
UFPA/Núcleo de Teoria e Pesquisa do Comportamento/Biblioteca

S586a Silva, Yslaíne Lopes, 1994-
Aversividade do timeout em uma situação de escolha /
Yslaíne Lopes Silva. — 2026.

178 f.: il.

Orientador: Marcus Bentes de Carvalho Neto

Tese (Doutorado) - Universidade Federal do Pará, Núcleo de Teoria e Pesquisa do Comportamento, Programa de Pós-Graduação em Teoria e Pesquisa do Comportamento, Belém, 2026.

1. Psicologia: pesquisa experimental. 2. Análise do comportamento. 3. Punição. 4. Timeout. 5. DRT (redução de atrasos). 6. RVA (valor agregado). 7. Teoria da informação (transmissão e compressão). 8. SiGN (sinalização). I. Título.

CDD - 23. ed. 150.77

Catálogo na fonte: Maria Célia Santana da Silva – CRB2/780

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 (Processo No. 88887.813885/2023-00).

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (Process No. 88887.813885/2023-00).

**Ysláine Lopes Silva, Programa de Pós-Graduação em Teoria e Pesquisa do
Comportamento, Universidade Federal do Pará, Belém-PA, Brasil.**

Contato: Ysláine Lopes Silva.

E-mail: yslaineledes@outlook.com



Programa aprovado pelo Conselho Superior de Ensino e Pesquisa da UFPA – Resolução 2545/98. Reconhecido nos termos das Portarias N.º. 84 de 22.12.94 da Presidente da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES e No. 694 de 13.06.95 do Ministério da Educação e do Desporto. Doutorado autorizado em 1999.

Defesa de Doutorado

“Aversividade do Timeout em uma Situação de Escolha.”

Aluna: Ysláine Lopes Silva.

Data da Defesa: 04 de novembro de 2025.

Resultado: Aprovada.

Banca Examinadora:

Profº Drº Marcus Bentes de Carvalho Neto (orientador – UFPA).

Profº Drº Paulo Sérgio Dillon Soares Filho (coorientador – UFPA).

Profª Drª Maria Helena Leite Hunziker (membro 1 – USP-SP).

Profº Drº Fábio Leyser Gonçalves (membro 2 – UNESP-Bauru).

Profº Drº Carlos Eduardo Costa (membro 3 – UEL).

Profª Drª Diana Milena Cortés Patiño (membro 4 – UFPA)

Termo de Autorização e Declaração de Distribuição não exclusiva para Publicação Digital no Repositório Institucional da UFPA

IDENTIFICAÇÃO DO AUTOR E DA OBRA

Autor*: Yslaine Lopes Silva

Vínculo com a UFPA: () Servidor; (X) Discente Unidade: Núcleo de Teoria e Pesquisa do Comportamento Sub Unidade: Programa de Pós-Graduação em Teoria e Pesquisa do Comportamento
Tipo do documento: (X) Tese; () Dissertação; () Livro; () Capítulo de Livro; () Artigo de Periódico; () Trabalho de Evento; ()

Outro. Especifique: _____ Título do Trabalho: Aversividade do Timeout em uma Situação de Escolha

Data da Defesa: 04/11/ 2025 Área do Conhecimento: Psicologia

Experimental Agência de Fomento: Capes

*Para cada autor, uma autorização preenchida e assinada.

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO EXCLUSIVA

O referido autor:

- Declara que o documento entregue é seu trabalho original, e que detém o direito de conceder os direitos contidos nesta licença. Declara também que a entrega do documento não infringe, tanto quanto lhe é possível saber, os direitos de qualquer outra pessoa ou entidade.
- Se o documento entregue contém material do qual não detém os direitos de autor, declara que obteve autorização do detentor dos direitos de autor para conceder à Universidade Federal do Pará os direitos requeridos por esta licença, e que esse material cujos direitos são de terceiros, está claramente identificado e reconhecido no texto ou conteúdo entregue.

Se o documento entregue é baseado em trabalho financiado ou apoiado por outra instituição que não a Universidade Federal do Pará, declara que cumpriu quaisquer obrigações exigidas pelo respectivo contrato ou acordo.

TERMO DE AUTORIZAÇÃO

Na qualidade de titular dos direitos de autor da publicação, autorizo a UFPA a disponibilizar de acordo com a licença pública *Creative Commons* Licença 3.0 *Unported*, e de acordo com a Lei nº 9610/98, o texto integral da obra citada, conforme permissões abaixo por mim assinaladas, para fins de leitura, impressão e/ou *download*, a partir desta data. Permitir o uso comercial da obra?

() Sim
(X) Não

Permitir modificações em sua obra?

() Sim, contanto que compartilhem pela mesma licença
(X) Não

O documento está sujeito ao registro de patente? () Sim

(X) Não

A obra continua protegida conforme a Lei Direito Autoral.

Belém(PA), 23/ 01 / 2026

Assinatura do Autor e/ou Detentor dos Direitos do Autor

Documento assinado digitalmente
gov.br YSLAINE LOPES SILVA
Data: 23/01/2026 13:00:51-0300
Verifique em <https://validar.itl.gov.br>

Esta tese é dedicada à Ana Monteiro, minha avó, e à Neomésia Monteiro, minha mãe, que me possibilitaram sonhar e alcançar. Que a força de vocês me acompanhe por onde eu for, e que cada passo meu seja sustentado pelo amor que me benze onde eu chegar.

Silva, Y. L. (2026). Aversividade do Timeout em uma Situação de Escolha. Tese de Doutorado. Programa de Pós-Graduação em Teoria e Pesquisa do Comportamento, Belém, PA, 178 p.

Resumo

O procedimento de timeout (TO) é classicamente descrito como uma forma de punição negativa, na qual a resposta remove temporariamente o acesso a reforçadores. Apesar do seu amplo uso em contextos aplicados e experimentais, ainda existem lacunas conceituais e empíricas acerca dos mecanismos que sustentam seus efeitos sobre o comportamento. O Estudo 1 desta tese apresentou uma análise teórica do TO à luz de quatro modelos contemporâneos de escolha e valor do reforço: Delay Reduction Theory (DRT), Hyperbolic Value-Added (HVA), Teoria da Informação e SiGN. A discussão evidenciou que o TO não atua apenas pela simples retirada do reforço, mas modifica de maneira sistemática a estrutura temporal das contingências, alterando tanto os atrasos médios ao reforço quanto a quantidade de informação transmitida por estímulos sinalizadores. Assim, o TO pode produzir efeitos que se aproximam de punição, mas também desempenhar uma função discriminativa e moduladora da expectativa de reforço. O Estudo 2 apresentou quatro experimentos empíricos que investigaram os efeitos de diferentes durações e taxas de apresentação do TO em esquemas concorrentes com ratos Wistar. No Experimento 1, manipulou-se a duração do TO (5 s, 20 s e 60 s) em concorrentes VI 20 s - VI 20 s, com diferenças sistemáticas entre as durações na taxa de resposta, mas sem grande impacto na preferência pela opção sem TO. O Experimento 2 testou o TO apresentado em VI (5 s) com diferentes durações (5 s e 20 s), confirmando que a presença do TO diminuiu a taxa de respostas na alternativa associada, mas, novamente, sem efeito claro da duração na alocação da resposta. O Experimento 3, conduzido em cadeias concorrentes, replicou esse padrão, mostrando preferência pela

alternativa sem TO, mas sem diferenças robustas entre durações curtas e longas. Por fim, o Experimento 4 variou a taxa de apresentação do TO (em VI 3 s, 5 s e 10 s), demonstrando que quanto maior a frequência de TOs, maior a evasão da opção associada, indicando que a taxa de TO, mais do que sua duração, é a variável crítica para a modulação da escolha. Contudo, a duração do TO se mostrou um fator consistente na alteração da taxa de resposta. De maneira integrada, os dois estudos contribuem para avançar a compreensão do TO em duas frentes complementares. Do ponto de vista teórico, apontam que os modelos de atraso e informação fornecem bases úteis para interpretar os efeitos do TO, mas carecem de refinamento quando aplicados a procedimentos de escolha simples e a manipulações paramétricas do TO. Do ponto de vista empírico, os resultados sugerem que a aversividade do TO não se relaciona linearmente à sua duração, mas sim à sua frequência de ocorrência e ao modo como reorganiza o fluxo temporal de reforços. Esses achados indicam que o TO deve ser compreendido não apenas como uma contingência de supressão, mas também como um evento que altera a estrutura informacional e temporal do ambiente, modulando as escolhas em função do atraso relativo e da previsibilidade de reforços. Essa perspectiva abre caminho para investigações futuras que articulem os efeitos do TO com modelos formais de escolha, explorando a interação entre atraso, sinalização e informação, em busca de uma estrutura conceitual mais abrangente para explicar seus efeitos tanto em contextos experimentais quanto em aplicados.

Palavras-chave: timeout, esquemas concorrentes, escolha, duração, frequência, punição, atraso do reforço, DRT, HVA, Teoria da Informação, SiGN

Silva, Y. L. (2026). Aversiveness of Timeout in a Choice Situation. Doctoral thesis. Graduate Program in Behavior Theory and Research, Belém, PA, 178 pages.

Abstract

The timeout (TO) procedure is classically described as a form of negative punishment, in which the response temporarily removes access to reinforcers. Despite its widespread use in applied and experimental contexts, there remain conceptual and empirical gaps regarding the mechanisms underlying its effects on behavior. Study 1 of this thesis presented a theoretical analysis of TO in light of four contemporary models of choice and reinforcement value: Delay Reduction Theory (DRT), Hyperbolic Value-Added (HVA), Information Theory, and SiGN. The discussion showed that TO does not act solely by simply removing reinforcement, but systematically modifies the temporal structure of contingencies, altering both the average delays to reinforcement and the amount of information conveyed by signaling stimuli. Thus, TO can produce effects that resemble punishment, but also perform a discriminative and modulatory function on reinforcement expectancy. Study 2 presented four empirical experiments that investigated the effects of different durations and rates of TO presentation in competing schedules with Wistar rats. In Experiment 1, the duration of the TO (5 s, 20 s, and 60 s) was manipulated in VI 20 s – 20 s competitors, with s in response rate across durations, but without a major impact on the preference for the option without a TO. Experiment 2 tested the TO presented in VI (5 s) across different durations (5 s and 20 s), confirming that the presence of TO decreased the response rate to the associated alternative, but again with no clear effect of duration on response allocation. Experiment 3, conducted in competing chains, replicated this pattern, showing a preference for the alternative without TO, but without robust differences between short and long durations. Finally, Experiment 4 varied the TO presentation rate (3 s, 5 s, and 10 s), demonstrating that the higher the TO frequency, the

greater the avoidance of the associated option, indicating that the TO rate, rather than its duration, is the critical variable for modulating choice. However, the TO duration proved to be a consistent factor in altering the response rate. Together, the two studies advance understanding of TO on two complementary fronts. From a theoretical point of view, they argue that delay and information models provide useful bases for interpreting the effects of TO but lack refinement when applied to simple choice procedures and parametric manipulations of TO. From an empirical point of view, the results suggest that the aversiveness of TO is not linearly related to its duration, but rather to its frequency of occurrence and the way it reorganizes the temporal flow of reinforcements. These findings indicate that TO should be understood not only as a contingency of suppression, but also as an event that alters the informational and temporal structure of the environment, modulating choices as a function of relative delay and the predictability of reinforcements. This perspective paves the way for future investigations that articulate the effects of TO with formal models of choice, exploring the interaction between delay, signaling, and information, in search of a more comprehensive conceptual framework to explain its effects in both experimental and applied contexts.

Keywords: timeout, concurrent schemes, choice, duration, frequency, punishment, reinforcement delay, DRT, HVA, Information Theory, SiGN

Lista de Ilustrações

Estudo 1

Figura 1 – Ilustração do procedimento de TO	23
Figura 2 – Ilustração da condição de TO e Intervalo Acoplado no Experimento 1 de Toegel et al., (2022).....	28
Figura 3 – Ilustração de um procedimento de cadeia concorrente.....	37
Figura 4 – Ilustração de um procedimento de cadeia concorrente com reforçamento probabilístico.....	47

Estudo 2

Figura 1 – Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 1.....	81
Figura 2 – Média geral da proporção da resposta das fases experimentais do Experimento 1.....	85
Figura 3 - Média individual da proporção da resposta das fases experimentais do Experimento 1.....	87
Figura 4 – Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 2.....	96
Figura 5 – Média geral da proporção da resposta das fases experimentais do Experimento 2.....	101
Figura 6 – Média individual da proporção da resposta das fases experimentais do Experimento 2.....	103
Figura 7 – Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 3.....	109
Figura 8 – Média geral da proporção da resposta das fases experimentais do Experimento 3.....	112

Figura 9 – Média individual da proporção da resposta das fases experimentais do Experimento 3.....	113
Figura 10 – Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 4.....	122
Figura 11 – Média geral da proporção da resposta das fases experimentais do Experimento 4.....	125
Figura 12 – Média individual da proporção da resposta das fases experimentais do Experimento 4.....	126

Lista de Tabelas

Estudo 2

Tabela 1 – Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 1.....	91
Tabela 2 – Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 2.....	104
Tabela 3 – Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 3.....	116
Tabela 4 – Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 4.....	128

Sumário

Estudo 1- Timeout Sob a Perspectiva das Teorias do Atraso: O Efeito do Tempo na Aversividade do Procedimento	17
Parâmetros que Afetam a Eficácia do TO.....	18
Timeout e Atraso do Reforço.....	22
O feito do TO na Redução da Taxa de Reforçamento	24
O efeito da Sinalização dos Intervalos Entre os Reforçadores.....	29
Escolha, Atraso, Reforço Condicionado e Timeout.....	34
Teorias Sobre o Atraso do Reforço e Reforço Condicionado	36
Timeout Sob a Perspectiva das Teorias do Atraso do Reforço	53
Conclusão.....	61
Estudo 2- Efeito da Duração e da Taxa de Apresentação do TO na Alocação da Resposta em Situação de Escolha	64
Efeitos da Duração do TO na Redução da Resposta.....	65
Efeitos da Taxa de Apresentação do TO na Redução da Resposta.....	69
TO em Situação de Escolha Concorrente	72
1. Experimento 1.....	74
1.1 Método.....	75
1.2. Sujeitos.....	75
1.3. Ambiente Experimental	76
1.4. Materiais e Equipamentos.....	76
1.5. Procedimento Geral.....	77
1.6. Análise de Dados	79
1.7 Resultados e Discussão	80
2. Experimento 2.....	94

2.1. Método.....	95
2.2. Sujeitos.....	95
2.3. Ambiente Experimental	95
2.4. Materiais e Equipamentos	95
2.5. Procedimento Geral.....	95
2.6. Análise de Dados	96
2.7. Resultados e Discussão	96
3. Experimento 3.....	105
3.1. Método.....	106
3.2. Sujeitos.....	106
3.3. Ambiente Experimental	106
3.4. Materiais e Equipamentos	106
3.5. Procedimento Geral.....	106
3.6. Análise de Dados	108
3.7 Resultados e Discussão	109
4. Experimento 4.....	120
4.1. Método.....	120
4.2. Sujeitos.....	120
4.3. Ambiente Experimental	121
4.4. Materiais e Equipamentos	121
4.5. Procedimento Geral.....	121
4.6. Análise de Dados	121
4.7. Resultados e Discussão	121
Discussão Geral.....	133
Conclusão.....	150

Referências.....	152
Apêndice A.....	166
Apêndice B.....	169
Apêndice C.....	170
Apêndice D.....	172
Apêndice E.....	173
Apêndice F.....	175
Apêndice G.....	178

Estudo 1

Timeout Sob a Perspectiva das Teorias do Atraso: O Efeito do Tempo na Aversividade do Procedimento

O termo “controle aversivo”, embora sem definição consensual na Análise do Comportamento (Hunziker, 2011, 2017), refere-se ao conjunto de relações comportamentais (respondentes e operantes), em que, dentro das relações operantes, destacam-se o reforçamento negativo, a punição positiva e a punição negativa (Catania, 1999; Pierce & Cheney, 2013). O timeout (TO), foco deste trabalho, é tradicionalmente classificado como um procedimento de punição negativa, uma vez que é caracterizado pela remoção contingente de estímulos reforçadores, com o objetivo de reduzir a probabilidade de ocorrência de uma resposta-alvo (Leitenberg, 1965; Defulio & Hackenberg, 2007).

No campo aplicado, o TO é amplamente utilizado em contextos educacionais e clínicos (Hackenberg & Defulio, 2007; Lydon et al., 2015). Nesses contextos, o procedimento pode envolver a saída do terapeuta ou a remoção do indivíduo do ambiente (Foxx & Shapiro, 1978), sendo operacionalizado em três modalidades principais: TO de exclusão, em que o indivíduo permanece no ambiente, mas é fisicamente removido da área de reforço, sem acesso às atividades; TO de não exclusão, em que o indivíduo permanece no ambiente e pode observar as atividades, mas está impedido de participar; e TO de isolamento, que envolve a remoção do indivíduo para um ambiente separado e sem reforçadores (Harris, 1984; Wolf et al., 2006).

Em contextos experimentais, similar ao que se observa em contextos aplicados, o TO pode envolver a retirada do operando, como a retração da barra em experimentos com ratos (e.g Toegel et al. 2022) ou somente o desligamento do mesmo, como no caso das chaves da resposta em estudos com pombos (e.g Parker et al., 1971). Nesse

contexto, o procedimento de TO também envolve a apresentação de estímulos, como luzes e sons, durante os quais não é apresentado o reforço (Kaufman & Baron, 1968; Parker et al., 1971; Toegel et al., 2022).

Apesar do TO ser considerado um procedimento de alta aceitabilidade social (Harris, 1984), seu uso frequentemente se baseia em diretrizes práticas, e os dados sobre os parâmetros que controlam sua eficácia como procedimento são ainda limitados (Hackenberg & Defulio, 2007). Fatores como a verbalização do motivo, as instruções, o local, a sinalização e a duração do TO têm sido investigados na pesquisa aplicada (Corralejo et al., 2018), mas de forma muitas vezes isolada (Toegel et al., 2022).

Embora as variáveis que afetam a eficácia da punição já tenham sido amplamente estudadas (Azrin & Holz, 1966), como intensidade, duração, taxa e magnitude, ainda são escassas as investigações específicas sobre como essas variáveis operam no contexto do TO (Toegel et al., 2022). Essa lacuna de entendimento é especialmente relevante quando se examina como variáveis experimentais, como a duração dos estímulos, podem influenciar o grau de aversividade desses estímulos.

Parâmetros que Afetam a Eficácia do TO

Os estudos sobre os parâmetros que podem afetar a eficácia do TO enquanto um procedimento aversivo tem demonstrado que, similar ao observado na literatura sobre controle aversivo que utilizou choques como punidores, parâmetros como duração e frequência do TO e a interação com a taxa de reforçamento, podem ser variáveis críticas na determinação de seus efeitos comportamentais (e.g. Dunn, 1990; Ferster e Appel, 1961; Hall, 2015; Thomas, 1968; Toegel et al, 2022).

A literatura sobre controle aversivo indica que a intensidade do estímulo aversivo influencia diretamente a supressão da resposta: quanto mais intenso o estímulo, maior a supressão (Azrin & Holz, 1966). Esse padrão também foi observado em pesquisas sobre

o TO, mais precisamente em relação à sua duração, equivalente à intensidade para este tipo de procedimento (e.g. Ferster e Appel, 1961; Zimmerman & Ferster, 1963; Flory e Lickfett, 1974; Kaufman & Baron, 1968), embora a relação entre esses fatores nem sempre seja linear. Por exemplo, Kaufman e Baron (1968) investigaram o efeito de utilizar diferentes durações do TO (10 s, 60 s, 90 s e 120 s) na taxa da resposta de ratos em uma tarefa de discriminação ($S^D \times S^A$) e concluíram que a supressão de respostas aumentava proporcionalmente ao tempo de TO. No entanto, o estudo de Nader e Morgan (2001), que investigou a autoadministração de cocaína por macacos rhesus, mostrou que essa relação não é tão simples, visto que a duração maior do TO (60 s) não resultou em uma supressão mais intensa da resposta em comparação com as durações de 10 s e 30 s.

Além da duração do estímulo aversivo, os estudos indicam que estímulos aversivos aplicados de forma contínua são mais eficazes na supressão do comportamento do que quando aplicados intermitentemente (Azrin & Holz, 1966). Estudos específicos sobre o TO, como os de Thomas (1968) e Hall (2015), corroboram essa conclusão ao demonstrar que frequências mais altas de TO, como em esquemas de razão fixa (FR 2) ou variável (VR 2), produzem maior supressão das respostas punidas do que esquemas de TOs menos frequentes, tanto em esquemas simples de reforçamento quanto em esquemas múltiplos.

O estudo de Hall (2015) investigou o efeito de diferentes intervalos de TO contingentes e não contingentes (VR: 2 - 40; VI 3 s - 45 s), combinados com um esquema de reforço VI 45 s em componentes múltiplos, sobre a taxa de respostas de pombos. Os resultados indicaram que os TOs contingentes reduziram de forma consistente as taxas de resposta, particularmente nos esquemas de VR de baixa razão

(VR 2) e VI mais frequente (VI 3s), enquanto os TOs de razão e intervalos mais altos ou independentes de resposta produziram supressões mais leves e idiossincráticas.

Essa relação também pode ser encontrada em experimentos concorrentes com esquemas de intervalo variável (VI). Por exemplo, Dunn (1990) expôs quatro pombos a esquemas concorrentes de reforço VI 45 s e VI 90 s para as chaves direita e esquerda, respectivamente. A alternância entre alternativas exigia quatro bicadas consecutivas em uma chave central. Em determinadas condições, foi implementado um TO de 20 segundos, programado em esquemas VI de 45 s, 90 s ou 180 s, seguindo a distribuição de Fleshler e Hoffman (1962), aplicados igualmente a ambas as alternativas. Cada esquema de TO foi apresentado por no mínimo 20 sessões consecutivas, com ordem aleatória entre os sujeitos. Os resultados mostraram uma diminuição da preferência pela chave associada ao VI 45 s quando o TO também seguia um VI 45 s, sugerindo um possível efeito de sobreposição entre esquemas de reforço e de TO. No entanto, a alocação de tempo permaneceu estável para dois sujeitos e diminuiu para os outros dois. Para as demais taxas de TO, não foram observadas diferenças sistemáticas na distribuição de respostas entre as opções. De modo geral, as taxas absolutas de resposta também não variaram sistematicamente em função da frequência de TO.

Outro fator que pode ser relevante para a eficácia do TO é a interação com as propriedades do reforço. Toegel et al. (2022) investigaram essa interação e sugeriram que ambientes enriquecidos durante o *time-in* (período de reforço) poderiam intensificar o efeito punitivo do TO. O experimento 4 desse estudo, no qual os sujeitos foram submetidos a esquemas múltiplos de diferentes taxas de reforço, indicou que taxas mais altas de reforço durante o *time-in* estavam associadas a uma maior supressão de respostas punidas por uma contingência de TO. No entanto, outros estudos, como os de

Thomas (1965) e Van-Haaren e Anderson (1998), apresentaram resultados conflitantes, sugerindo que o TO pode ser menos aversivo em contextos de alta taxa de reforço.

Em resumo, a eficácia do TO como estímulo punitivo parece depender de uma combinação de fatores, como a duração, a frequência e a interação com o reforço. Embora haja evidências de que frequências mais altas de TO e ambientes de reforço mais ricos possam aumentar o efeito supressivo da contingência de TO, a relação entre essas variáveis é complexa e não totalmente compreendida.

A falta de consenso sobre os parâmetros que influenciam a eficácia do TO na supressão de respostas pode, em parte, ser atribuída à concepção do procedimento como um estímulo aversivo estrito, ou seja, como um evento discreto que é apresentado ou removido do ambiente, semelhante a um choque elétrico. No entanto, consideramos que talvez os efeitos do TO poderiam ser melhor entendidos se, respeitando as características do próprio procedimento, o considerássemos como um tempo sinalizado, no qual os reforços não serão apresentados, que é adicionado entre a resposta e disponibilidade do reforçamento.

Naturalmente, considerar o TO como um procedimento que acrescenta um tempo sinalizado que altera os intervalos entre os reforços nos aproxima de uma literatura tradicional na análise do comportamento, que aborda os efeitos de manipulações do atraso dos reforçadores (e.g. Rachlin et al., 1986; Mazur, 2001). O atraso do reforço normalmente é definido como um período entre a emissão de uma resposta e a apresentação do estímulo reforçador produzido por essa resposta (Lattal, 2010). No entanto, ao contrário do atraso do reforço, que interrompe a contiguidade temporal entre a resposta e o reforço (Lattal, 2010), indicando que o reforço será entregue somente após o término desse período, o TO não sinaliza a entrega do reforçador. Em vez de apenas atrasar o estímulo reforçador, a contingência de TO

especifica que não haverá reforço, nem durante o intervalo de tempo estabelecido, nem ao seu término.

Dado que o TO pode ser entendido como um procedimento que introduz um intervalo entre as apresentações dos reforçadores, uma abordagem promissora para identificar os mecanismos e contextos que tornam o TO aversivo pode ser encontrada nas teorias que explicam a relação entre os efeitos de manipulações do tempo/atraso entre a resposta e o reforço e entre os reforçadores (e.g., Rachlin et al., 1986; Mazur, 2001). Nesse sentido, este texto propõe analisar a literatura sobre os efeitos do TO com base nas teorias do atraso do reforçamento, especificamente a partir da Teoria da Redução do Atraso (DRT, em inglês Delay Reduction Theory; Fantino, 1969), Modelo Hiperbólico do Valor Adicionado (HVA, em inglês Hyperbolic Valued-Added Model; Mazur, 1989), Teoria da Informação (Shahan & Cunningham, 2015) e Sinalização para boas notícias (SiGN, em inglês Signal for Good News; McDevitt et al., 2016).

A utilização dessas abordagens teóricas tem como objetivo oferecer um arcabouço conceitual para interpretar como o tempo introduzido pelo procedimento de TO pode alterar o valor relativo das alternativas disponíveis e, conseqüentemente, contribuir para sua função aversiva. Importa destacar, contudo, que este trabalho não se propõe a testar formalmente as equações, modelos matemáticos ou predições quantitativas derivadas dessas teorias. Em vez disso, as teorias são empregadas como instrumentos analíticos para compreender como diferentes concepções do tempo, inclusive o reforço, podem iluminar os efeitos comportamentais observados em procedimentos de TO.

Timeout e Atraso do Reforço

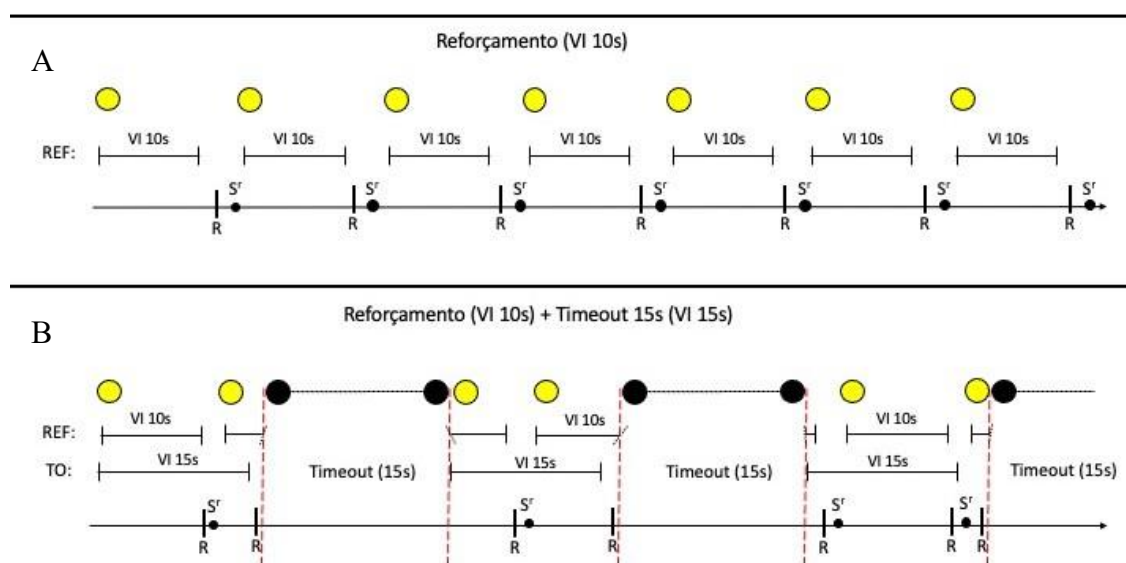
A relação entre o TO e os procedimentos de atraso do reforço é complexa e envolve variáveis como a redução da taxa de reforço e a sinalização de períodos de

ausência de reforço. Para ilustrar esses efeitos, a Figura 1 apresenta dois esquemas comparativos: um em que as respostas são reforçadas no esquema VI 10s (linha superior) e outro em que o mesmo esquema é combinado com a apresentação de TOs de 15s programados no esquema VI 15s (linha inferior).

No primeiro caso, cada resposta após o intervalo VI resulta diretamente em reforço. Já no segundo, embora o reforço continue disponível em VI 10s, algumas respostas produzem a apresentação do TO (15s), durante o qual o reforço fica indisponível, impondo, portanto, um atraso na entrega do reforço. Essa alteração no fluxo da contingência pode reduzir a taxa efetiva de reforçamento e alterar o controle de estímulos com a sinalização diferencial do período de TO.

Figura 1

Ilustração do procedimento de TO



Nota: Ilustração de um procedimento de TO em que o reforço ocorre em VI 10 s na condição de reforçamento (A). Na condição com TO, o reforço também ocorre em VI 10s (*time-in/ luz*), mas tem um TO (sem luz) em VI 15s com duração de 15s (B). O símbolo R equivale a emissão da resposta e S^r equivale a entrega de reforço.

Para compreender melhor essa relação, é necessário aprofundar o entendimento sobre a função de ambos os procedimentos e seus efeitos sobre o comportamento dos organismos.

O efeito do TO na Redução da Taxa de Reforçamento

Uma vez que o TO é um procedimento que impede temporariamente o acesso aos estímulos reforçadores, é natural considerar, em um primeiro momento, que seu efeito aversivo esteja exclusivamente relacionado à redução da taxa global de reforçamento produzida pelo procedimento (Hall, 2015). Esse impacto pode ser claramente observado na Figura 1, que compara dois cenários: um com reforço sob VI 10s e outro com a mesma programação de reforço, acrescida de TOs de 15s programados sob VI 15s.

Embora, em ambos os casos, o reforço continue sendo disponibilizado sob VI 10s, a introdução do TO cria períodos em que o reforço não está acessível, mesmo que a resposta seja emitida. Isso reduz a taxa efetiva de reforçamento.

Apesar de essa interpretação ser plausível e explicar muitos resultados, evidências de diferentes estudos indicam que o efeito do TO nem sempre pode ser compreendido apenas pela redução da taxa global de reforçamento, sugerindo a participação de outros processos. Por exemplo, os trabalhos, como os de Carlson (1972); Branch (1977); Lattal (1984) e Toegel et al. (2022), buscaram avaliar especificamente se a redução da taxa global de reforçamento seria suficiente para reproduzir os efeitos do TO, utilizando, para isso, períodos de indisponibilidade do reforço apresentados de forma sinalizada ou não sinalizada, e sob condições contingentes ou não contingentes à resposta.

O estudo de Carlson (1972) investigou o efeito do TO utilizando um procedimento de cadeias concorrentes. Nesse estudo, macacos foram expostos a duas

cadeias concorrentes com esquemas VI 30 s nos elos iniciais e VI 15 s nos elos finais. Quando um TO de 15 s em VI 15 s foi adicionado ao elo final da opção inicialmente mais preferida, os sujeitos rapidamente transferiram suas respostas para a alternativa sem TO, mesmo quando o elo final dessa alternativa foi alterado para VI 30 s, o que igualou a taxa de reforço entre as opções. Esses resultados indicam que o TO afeta a preferência, produzindo efeitos que vão além da simples redução da taxa de reforçamento.

No estudo de Branch (1977), quatro experimentos foram conduzidos com pombos expostos a esquemas compostos por um primeiro componente de razão fixa (FR 15) e um segundo componente temporal (FI ou FT 25 s), com variações na porcentagem de cadeias reforçadas e na sinalização das consequências. No Experimento 1, 75%, 50% ou 25% das cadeias eram reforçadas com 3 s de acesso ao alimento, enquanto os demais 25%, 50% ou 75% terminavam com um período de 3 s de timeout (blecaute). No Experimento 2, as consequências foram sinalizadas: cadeias não reforçadas eram seguidas por uma luz âmbar, enquanto cadeias reforçadas mantinham a luz verde. No Experimento 3, os TOs passaram a ser apresentados independentemente da resposta, sob condições sinalizadas e não sinalizadas. No Experimento 4, tanto os reforços quanto os TOs não eram contingentes à resposta, e o estímulo da segunda componente podia ser sinalizado (luz branca para reforço, âmbar para timeout) ou não (apenas luz branca). Os resultados mostraram que a presença de estímulos discriminativos para as consequências aumentou o número de cadeias completadas por sessão, mesmo quando a taxa de reforçamento era menor (e.g. 25%). Sob condições não sinalizadas, observou-se uma redução expressiva no número de cadeias completadas, associada a pausas mais longas entre tentativas e menor taxa de respostas no primeiro componente. Essa condição se manteve mesmo quando as consequências não

dependiam da resposta. Contudo, no Experimento 3, em comparação com o Experimento 2, foi observado que o número de cadeias finalizadas foi menor na condição de TO não contingente (Experimento 3).

Esses resultados sugerem que a sinalização das consequências exerce um papel crítico no controle discriminativo do comportamento: estímulos que indicam se uma cadeia será reforçada ou punida (TO) facilitam a alocação de respostas, mesmo quando a taxa de reforço é baixa. Além disso, a contingência entre a resposta e a consequência também influencia a manutenção do responder. Assim, tanto a sinalização quanto a contingência parecem atuar de maneira complementar, afetando diretamente a taxa da resposta.

Complementando esses achados, Lattal (1984, Experimento 2) expôs cinco pombos a um esquema de cadeia VI 50 s seguido por FT 20 s, 75 s ou 150 s, dependendo da fase. Foram comparadas cinco condições principais: (1) atraso sinalizado (blecaute imediatamente após a resposta que inicia o atraso) com intervalo simulado não sinalizado (20 s, 75 s e 150 s) programados de forma independentes em VI 50s e que paralisava o VI de reforço; (2) igual à condição anterior, mas o atraso não era sinalizado (sem blecaute); (3) atraso não sinalizado com blecautes dependentes da resposta (blecautes de 20s produzidos por respostas em um VI 50s que substituíam o intervalo simulado); (4) atraso não sinalizado com blecautes independentes da resposta (blecautes apresentados assim que disponíveis, independentemente da resposta que substituíam o intervalo simulado); e (5) blecautes aleatórios independentes da resposta (blecautes e atrasos podiam se sobrepor no tempo, sem correlação entre si).

Os resultados indicaram que, em quatro dos cinco sujeitos, a taxa de resposta foi mais elevada na condição de atraso sinalizado, sendo que um deles chegou a responder em níveis superiores à própria linha de base. Esse padrão sugere que o sinal associado

ao atraso pode ter funcionado como estímulo discriminativo e/ou como reforçador condicionado, sustentando a taxa de respostas mesmo na presença do atraso. Em contraste, as condições com blecautes não correlacionados ou negativamente correlacionados ao atraso, sejam contingentes ou não contingentes à resposta, produziram taxas de resposta menores.

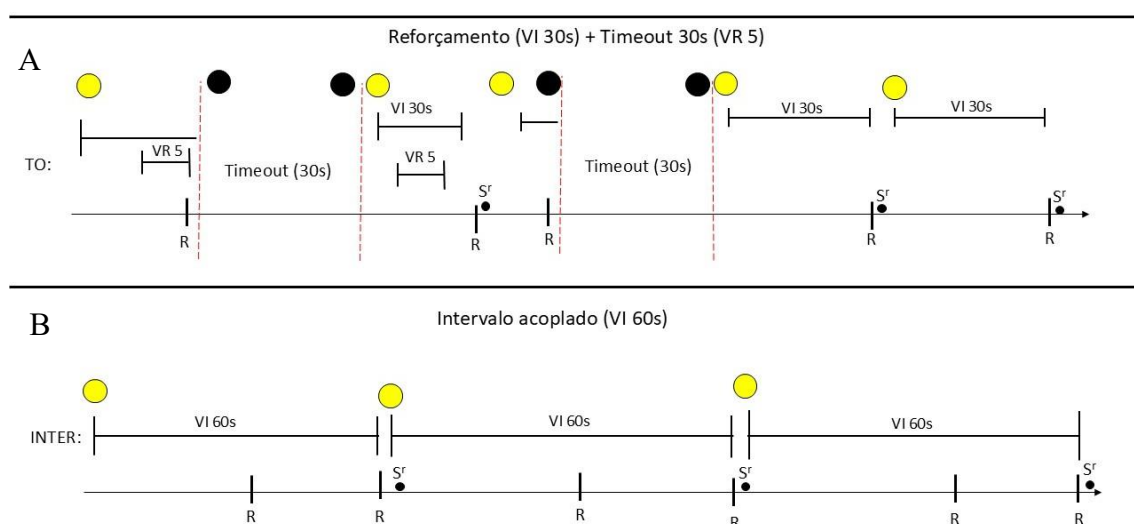
Ainda que, em geral, tenha ocorrido uma redução nas respostas nas condições em que o blecaute não sinalizava o atraso, os resultados variaram consideravelmente entre os sujeitos: para três dos cinco indivíduos, as taxas de resposta foram menores durante os blecautes independentes do que durante os contingentes. Esses achados indicam que, se a alteração na taxa de resposta fosse explicada apenas pela alteração na taxa de reforçamento, os padrões de resposta seriam semelhantes entre condições contingentes e não contingentes, o que não foi observado. Embora Lattal não mencione explicitamente o uso de TO, as condições de blecaute contingente e não contingente à resposta podem ser consideradas equivalentes às condições de TO contingente e não contingente, respectivamente

Mais recentemente, o Experimento 1 de Toegel et al. (2022) submeteu quatro ratos a diferentes condições: uma linha de base com reforço em VI 30s; uma condição de TO (1) de 30s em VR 5 sobreposta ao esquema de reforço; uma condição de TO simulado (2), semelhante à condição anterior, mas com apresentação de reforço não contingente durante o TO; uma condição de intervalo acoplado (3), sem TO, mas com manipulação dos esquemas de reforço de acordo com as apresentações de TO na primeira condição; e uma condição de TO acoplado (4), em que os intervalos sem reforço foram apresentados independentemente da resposta, baseando-se nas apresentações de TO da primeira condição. Os autores observaram que as condições de TO simulado e TO acoplado (não contingente) não foram eficazes em suprimir a

resposta para três dos quatro sujeitos em comparação à linha de base. Em contraste, a condição de TO suprimiu a resposta de três sujeitos na primeira apresentação e de todos os sujeitos na segunda. No entanto, a condição de intervalo acoplado, sem TO, foi a que mais reduziu a taxa de resposta para todos os sujeitos (ver Figura 2). Esses achados indicam que a simples alteração da taxa global de reforçamento não explica, por si só, o efeito do TO, uma vez que condições com taxas semelhantes de reforço (TO e intervalo acoplado) produziram efeitos distintos sobre o comportamento.

Figura 2

Ilustração da condição de TO e Intervalo Acoplado no Experimento 1 de Toegel et al., (2022)



Nota: Ilustração do Experimento 1 de Toegel et al. (2022). A figura ilustra especificamente a condição de TO (A), em que o reforço ocorreu em VI 30 s e um TO de 30 s e VR 5 foram sobrepostos a ele. E a condição de Intervalo Acoplado (B) em que o reforço ocorreu em um intervalo similar aos intervalos entre reforços da condição de TO. O símbolo R equivale à emissão da resposta e S^r equivale à entrega de reforço.

No entanto, algumas lacunas permanecem quando comparamos os resultados desses estudos. Por exemplo, Branch (1977) e Lattal (1984) observaram que períodos de

blecaute não contingentes à resposta levaram a uma redução maior na taxa do comportamento-alvo do que blecautes contingentes, o que contrasta com os achados de Toegel et al. (2022), que mostraram maior supressão da resposta sob TO contingente do que sob TO não contingente. Além disso, o estudo de Carlson (1972), que trabalhou com cadeias concorrentes em condições semelhantes às de TO e de intervalo acoplado de Toegel, encontrou que o TO provocou maior redução na taxa de resposta no elo associado ao TO, mesmo com taxas de reforço equiparadas entre as opções; no entanto, não é possível uma comparação direta entre esses estudos, dadas as diferenças metodológicas.

Embora todos os estudos confirmem que o TO reduz o comportamento alvo, a magnitude e a natureza desse efeito variam conforme os esquemas utilizados (simples, múltiplo e concorrente; contingente e não contingente) e a sinalização apresentada. Essas discrepâncias indicam que ainda há aspectos do funcionamento do TO que precisam ser melhor compreendidos. Em resumo, os dados atuais ainda não oferecem uma resposta conclusiva sobre como mudanças na taxa de reforço afetam a relação entre os intervalos impostos e a distribuição da resposta. A redução na taxa de reforço é um fator relevante, mas a introdução de intervalos sinalizados também parece ser fundamental para entender os efeitos comportamentais dos procedimentos em questão.

O efeito da Sinalização dos Intervalos Entre os Reforçadores

O procedimento TO é caracterizado pela apresentação de um estímulo que sinaliza a indisponibilidade de reforço por um período determinado. Essa sinalização apresenta paralelos com procedimentos de atraso do reforço, nos quais uma mudança de estímulo também indica o intervalo até a apresentação do reforço. De modo geral, estudos que investigaram a função desses sinais em procedimentos de atraso (e.g.,

Kelley et al., 2011; Lattal, 1984; Richards, 1981) mostram que atrasos sinalizados tendem a ser menos aversivos do que atrasos não sinalizados.

No estudo de Richards (1981), nove pombos tiveram a resposta de bicar uma chave reforçada segundo um esquema VI 60 s, no qual a entrega do reforço ocorria após atrasos variáveis de 0,5 a 10 s. Em metade das sessões, esses atrasos eram sinalizados; na outra metade, não havia sinalização. Os resultados indicaram que a presença do sinal reduziu moderadamente a taxa de respostas, independentemente do tamanho do atraso, enquanto, nos atrasos não sinalizados, a supressão da resposta aumentou proporcionalmente à duração do intervalo. Esses achados sugerem que a sinalização atenua o caráter aversivo do atraso, favorecendo a manutenção do comportamento mesmo quando o reforço está distante.

De forma complementar, o Experimento 1 de Lattal (1984) avaliou os efeitos de blecautes contingentes e não contingentes sobre a taxa de respostas de três pombos em um esquema de reforço VI 50 s. Os animais foram expostos a duas condições: blecautes de 20 s contingentes a uma resposta e blecautes aleatórios independentes da resposta. As taxas de resposta aumentaram durante ambas as condições de blecaute em relação à linha de base, e não houve diferença significativa entre as condições contingentes e não contingentes. Esses resultados indicam que sinais de indisponibilidade de reforço, quando não associados de forma consistente à sua ocorrência, exercem pouco efeito sobre a taxa de respostas.

Em um contexto aplicado, o estudo de Kelley et al. (2011) investigou como atrasos sinalizados versus não sinalizados afetam respostas de comunicação em três crianças com deficiências do desenvolvimento e autismo. Cada participante foi treinado para emitir uma resposta comunicativa que produzia reforço imediato sob um esquema de reforço FR-1. Após o treinamento, foram aplicadas quatro condições experimentais:

reforço imediato (0 s de atraso), atraso sinalizado, atraso não sinalizado e extinção. Nas condições de atraso, o reforço foi gradualmente atrasado, iniciando em 2 segundos e chegando a 600 segundos, conforme o desempenho do participante. Na condição de atraso sinalizado, um estímulo discriminativo era apresentado durante o atraso: um recipiente com moedas, um vídeo ou um cronômetro visível. Na condição de atraso não sinalizado, o terapeuta simplesmente aguardava o intervalo. Os resultados indicaram que, embora as taxas de comunicação fossem semelhantes em atrasos curtos para ambas as condições, a presença do estímulo sinalizador manteve a comunicação em níveis mais altos e estáveis em atrasos longos, enquanto, na ausência do sinal, as respostas diminuíram rapidamente.

Kendall (1974) também investigou o papel da sinalização, mas em uma configuração diferente dos estudos anteriores, levando em consideração a probabilidade de reforço em esquemas concorrentes. No Experimento 1, pombos responderam em elos iniciais FR 1 para acessar uma alternativa de alta densidade (reforço sempre após o atraso) ou de baixa densidade (reforço em 50% dos casos; demais resultando em TO). No grupo correlacionado, luzes verdes e vermelhas sinalizavam, respectivamente, reforço e TO; no grupo não correlacionado, esses estímulos não eram preditivos. A correlação perfeita sustentou forte preferência pela alternativa de baixa densidade, enquanto a ausência de correlação aumentou a escolha pela alternativa de alta densidade. No Experimento 2, com VI 20 s nos elos iniciais, manipulou-se sistematicamente a duração do atraso (5 s, 15 s ou 45 s), a correlação dos estímulos na baixa densidade (perfeita ou ausente) e a probabilidade de reforço na alta densidade (100% ou 80%). Atrasos mais longos, correlação perfeita e menor probabilidade de reforço na alternativa de alta densidade aumentaram a preferência pela baixa densidade, enquanto a ausência de correlação reduziu essa preferência.

Esse último estudo converge com análises sobre atraso e probabilidade de reforço, como as de Mazur (1991), que argumenta que reforços probabilísticos funcionam como atrasos variáveis até sua obtenção. Modelos anteriores (e.g., Rachlin et al., 1986) tendiam a superestimar a perda de valor dos reforços incertos, por não considerarem o papel dos estímulos presentes entre a resposta e sua consequência. Esses estímulos, quando associados consistentemente ao reforço, funcionam como reforçadores condicionados, sustentando a preferência mesmo com baixa probabilidade de reforço, similar ao que acontece com atrasos sinalizados do reforço.

O papel dos reforçadores condicionados é amplamente documentado (Fantino, 1977; Williams, 1994; Mazur, 1989, 2006; Fantino & Romanowich, 2007), e mostra que estímulos pareados com reforçadores primários, como alimento, podem manter e aumentar taxas de resposta em relação a condições sem tais estímulos. Além disso, evidências indicam que o valor de um reforço atrasado depende não apenas do tempo total até sua entrega, mas também do tempo passado na presença de estímulos que o predizem (Rachlin et al., 1986; Mazur, 1989, 2001; Shahan, 2010).

Por exemplo, Mazur (1989) demonstrou que a exposição prolongada a estímulos associados ao reforço exerce mais influência sobre o comportamento do que o simples intervalo temporal até sua ocorrência. Nos três experimentos realizados no estudo, pombos escolheram entre duas alternativas: uma alternativa probabilística (luz vermelha) e uma alternativa de ajuste (luz verde). A alternativa probabilística apresentava um atraso fixo de 5 segundos antes do possível reforço, entregue com probabilidade de 0,5 ou 0,2, seguido de um intervalo entre tentativas (IET). A alternativa de ajuste fornecia reforço em 100% das tentativas, com um atraso variável que era modificado até atingir o ponto de indiferença (o atraso no qual o animal escolhia igualmente entre as duas opções). O IET variou entre condições (30 s, 60 s, 90 s ou 3 s,

dependendo do atraso). De forma geral, os resultados mostraram que o IET não afetou a preferência, contrariando previsões de modelos que incluem o tempo total de tentativa (Rachlin et al., 1986), e corroborando as explicações de que a preferência é melhor explicada pelo tempo gasto sob os estímulos associados ao atraso antes do reforço, especialmente quando incorporam a variabilidade desses atrasos (Mazur, 1984). Alterações nos estímulos durante tentativas não reforçadas (como a remoção da luz vermelha) aumentaram a preferência pela alternativa probabilística, confirmando que a duração e natureza do estímulo de espera influenciam mais do que o tempo total da tentativa.

A lógica do reforço probabilístico apresentada por Mazur pode ser estendida para compreender os efeitos do TO, já que esse procedimento adiciona um período de suspensão temporária das contingências, alterando a probabilidade e o padrão temporal da obtenção do reforço, o que aumenta a variabilidade e a incerteza no tempo até o reforço. No entanto, conforme Mazur enfatiza, o intervalo em si pode não ser o fator principal para explicar a preferência; a sinalização associada a esse intervalo desempenha um papel crucial. No caso do TO, essa sinalização diferencial, indicando a indisponibilidade temporária do reforço, pode, dependendo do contexto, reduzir a incerteza temporal sobre quando o reforço estará disponível. Shahan (2010) destaca que estímulos antecedentes relacionados ao reforço modulam o controle do comportamento, atuando tanto como reforçadores condicionados quanto como sinais discriminativos da disponibilidade do reforço. Assim, o TO pode, ao mesmo tempo, diminuir a incerteza ao sinalizar claramente a ausência de reforço, mas também reduzir o valor da alternativa, funcionando como um indicativo de atraso prolongado.

O estudo de Toegel et al. (2022) ilustra esse ponto ao mostrar que o TO com sinalização diferencial entre períodos de reforço e não reforço suprime a resposta em

comparação com a contingência apenas de reforço, contudo, a condição de TO reduziu menos a resposta do que a condição com intervalo acoplado sem sinalização. Essa diferença é coerente com os estudos sobre atraso e reforço probabilístico citados anteriormente, que indicam que contingências que fornecem informações claras, mesmo quando indicam ausência de reforço, tendem a ser preferidas em relação a opções com sinalização ambígua (Sears et al., 2022).

Assim, a aversividade do TO, caracterizada pela sinalização diferencial do período sem reforço, parece depender da clareza informativa das outras opções disponíveis, especialmente em situações de contingências concorrentes. As interações entre sinalização, taxa de reforço e comportamento são complexas e requerem um entendimento integrado para compreender os mecanismos envolvidos na aversividade do TO. Estudos sobre comportamento de escolha têm explorado essa dinâmica, propondo teorias que relacionam a função do estímulo discriminativo não apenas à taxa e magnitude do reforço, mas também ao tempo necessário para a entrega do reforçador primário. Exemplos notáveis incluem a DRT (Fantino, 1969), o HVA (Mazur, 1987), a Teoria da Informação (Shahan & Cunningham, 2015) e o SiGN (McDevitt et al., 2016), que enfatizam a importância desses fatores na avaliação das alternativas disponíveis.

Portanto, neste momento, o presente texto apresentará um breve resumo das teorias quantitativas da escolha, do atraso do reforçamento e do reforço condicionado, enfatizando como essas teorias podem nos ajudar a compreender os resultados observados ao utilizar o procedimento de TO.

Escolha, Atraso, Reforço Condicionado e Timeout

O comportamento de tomar decisões ou fazer escolhas, do ponto de vista comportamental, pode ser entendido como a alocação do tempo de um organismo entre duas ou mais atividades (Baum, 2010). Estudos demonstraram que essa alocação é

diretamente afetada pela taxa dos reforçadores disponíveis (Baum & Rachlin, 1969; Herrnstein, 1970). A formalização dessa relação foi inicialmente apresentada por Herrnstein (1961), sendo denominada Lei da Igualação, conforme descrita pela equação:

$$\frac{B_1}{B_1 + B_2} = \frac{R_1}{R_1 + R_2} \quad (1)$$

na qual B_1 e B_2 correspondem às taxas de resposta e R_1 e R_2 representam as taxas de reforço para a primeira e segunda opção, respectivamente.

Com a avaliação dos efeitos de diferentes propriedades do reforçamento Baum e Rachlin (1969), observaram que a distribuição do comportamento era diretamente proporcional a razão da taxa, magnitude e imediaticidade do reforço. Assim, todas as propriedades do reforço, quando concatenadas, descreveriam o valor das opções. Essa proposta ficou conhecida como a Lei Concatenada da Igualação e pode ser descrita pela equação:

$$\frac{B_1}{B_2} = \left(\frac{R_1}{R_2}\right) * \left(\frac{A_1}{A_2}\right) * \left(\frac{I_1}{I_2}\right) * \left(\frac{X_1}{X_2}\right) = \frac{V_1}{V_2} \quad (2)$$

em que A_1 e A_2 se referem a magnitude dos reforçadores, I_1 e I_2 se refere a imediaticidade e X_1 e X_2 se referem a outros parâmetros do reforço diferentes da taxa, magnitude e imediaticidade. Todos esses valores podem ser resumidos pelo construto de valor da opção nos parâmetros V_1 e V_2 .

Posteriormente, para permitir a descrição de alguns desvios do comportamento em situação de escolha concorrente, Baum (1974) apresentou uma variação da Lei da Igualação de Herrnstein (1961) que ficou conhecida como a Lei da Igualação Generalizada (GLM), que, em sua forma logarítmica, pode ser descrita pela equação:

$$\log\left(\frac{B_1}{B_2}\right) = a \log\left(\frac{R_1}{R_2}\right) + \log b \quad (3)$$

Essa equação é algebricamente similar à equação 1 original, mas com o acréscimo dos parâmetros de sensibilidade, representado por a , e viés, representado por $\log b$.

A sensibilidade (a) mede a resposta do comportamento em relação às taxas de reforço; quando $a = 1$, há correspondência perfeita entre a distribuição das respostas e as taxas de reforço. Valores superiores a 1 indicam *overmatching*, com taxas de resposta maiores do que o esperado, enquanto valores inferiores a 1, conhecidos como *undermatching*, refletem taxas de resposta menores do que o esperado (Baum, 1974; 1979; McDowell, 1989; Reed & Kaplan, 2011). O viés ($\log b$) reflete a preferência do sujeito por uma opção de resposta, influenciada por fatores além das taxas de reforço, como características do organismo e do ambiente. Quando $b = 0$, a distribuição das respostas é exclusivamente determinada pelas taxas de reforço (Baum, 1974; Reed & Kaplan, 2011).

No contexto do desenvolvimento dos modelos de escolha, vêm sendo desenvolvidos modelos que tentam descrever os efeitos de outras variáveis do reforço, como a demora na apresentação do estímulo reforçador. Embora esses modelos não abordem diretamente o procedimento de TO, podem oferecer uma perspectiva alternativa para entender certos aspectos dele, especialmente quando compreendemos o TO como um procedimento que adiciona tempo entre as apresentações dos reforçadores.

Alguns dos modelos desenvolvidos, a partir da Lei da Igualação, para descrever os efeitos de estímulos que sinalizam a demora na entrega dos reforçadores são: DRT; HVA; Teoria da Informação e SiGN.

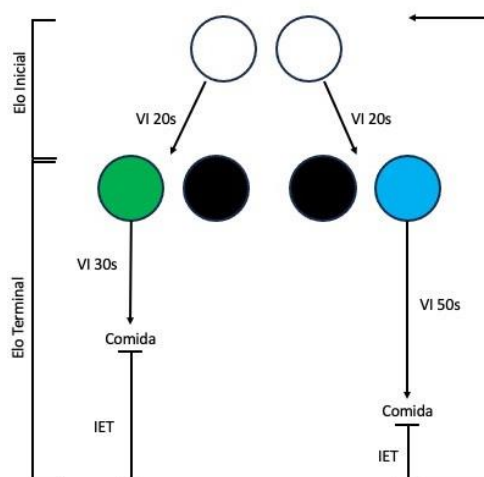
Teorias sobre o Atraso do Reforço e Reforço Condicionado

A DRT, proposta por Fantino (1969), foi desenvolvida originalmente para descrever a distribuição das respostas entre dois esquemas de cadeias concorrentes.

Nestes esquemas, responder a uma das opções concorrentes (elo inicial) produz a apresentação de um estímulo associado ao esquema de reforço primário (elo terminal). Por exemplo, em um esquema de cadeia concorrente, uma tentativa se inicia com a apresentação de duas chaves brancas que operam sob um esquema VI – 20s (elos iniciais). Ao acionar a chave direita, ela muda para azul, enquanto a chave esquerda se apaga e se torna inoperante. Um esquema VI – 50s é aplicado e, ao final do intervalo, a resposta na chave direita recebe um reforço primário (elo terminal 1). Segue-se um intervalo de 10 s com as luzes apagadas e ambas as chaves inoperantes (intervalo entre tentativas) e, em seguida, uma nova tentativa é iniciada. Em uma nova tentativa, se a resposta for na chave esquerda, ela muda para verde e a chave direita fica inoperante. Um esquema VI – 30s entra em vigor e, ao final desse intervalo, a resposta é conseqüenciada por um reforço primário (elo terminal 2) (ver figura 3).

Figura 3

Ilustração de um procedimento de cadeia concorrente



Nota: Ilustração de um procedimento de cadeia concorrente. Os elos iniciais (cor branca) estão em VI 20 s. Após a escolha de uma das opções, o estímulo muda e a opção alternativa fica apagada. Caso o sujeito escolha a opção esquerda, o estímulo

verde sinaliza o elo terminal em VI 30 s e finaliza com a apresentação do estímulo reforçador. Caso o sujeito escolha a opção direita, o estímulo azul sinaliza o elo terminal em VI 50 s e finaliza com a apresentação do estímulo reforçador. Após a apresentação do reforço, um intervalo entre tentativas é iniciado para, em seguida, prosseguir para a fase de escolha novamente (elos iniciais).

Inicialmente, Herrnstein (1964) sugeriu que a preferência pelo elo terminal (i.e. distribuição da resposta no elo inicial) estava exclusivamente direcionada à taxa relativa de reforçamento primário, e, portanto, poderia ser descrita pela lei estrita da igualação, tal qual proposta por Herrnstein (1961). No entanto, contrariando as predições da lei estrita da igualação, Fantino (1969) demonstrou que alterações na duração do elo inicial diminuem a preferência pela opção associada com maior magnitude de reforçamento, o que o levou a proposição da DRT.

De acordo com a proposta da DRT, a preferência no elo-inicial está em função da taxa de reforçamento primário e do valor reforçador condicionado do estímulo associado com o elo-terminal (O'daly & Fantino, 2003). Fantino (1969) sugere ainda que a efetividade de um estímulo como reforçador condicionado está relacionada à quantidade de redução no tempo de espera para o reforço principal em comparação com o tempo total da condição analisada. Assim, o valor reforçador condicionado do estímulo do elo inicial corresponde a diferença entre o tempo total da opção (duração do Elo-Inicial + Elo-Terminal; T) e a duração do Elo-Terminal (t_A). As afirmações da DRT podem ser representadas, em sua forma mais simples, da seguinte maneira:

$$A = f\left(\frac{T - t_A}{T}\right) \quad (4)$$

Portanto, um estímulo que demonstra uma redução mais significativa no tempo até o reforço será mais eficaz como reforçador condicionado do que um estímulo que mostra uma redução menor (Fantino et al., 1993; Killeen & Fantino, 1990). No exemplo

citado anteriormente, o estímulo associado a chave esquerda teria um valor condicionado maior que o estímulo da chave esquerda, visto que, na opção direita, o tempo total até o reforço (T) é a soma do tempo do elo inicial (chave branca com VI 20s) mais o tempo do elo terminal (luz azul com VI 50s), ou seja: $T = 20 + 50 = 70$ s. O tempo do elo terminal (t_A) é 50 s. Aplicando a equação da DRT:

$$A = f\left(\frac{70 - 50}{70}\right) = f\left(\frac{20}{70}\right) = 0,28$$

Enquanto na opção esquerda, o tempo total até o reforço (T) é $20 + 30 = 50$ s. O tempo do elo terminal (t_A) é 30 s. Aplicando a equação da DRT:

$$A = f\left(\frac{50 - 30}{50}\right) = f\left(\frac{20}{50}\right) = 0,4$$

Nesse caso, o estímulo da opção esquerda (luz verde) tem maior valor porque sinaliza uma redução maior do tempo até o reforço (40%) do que o da opção direita (luz azul), que reduz em 29%. Portanto, a DRT prevê que o sujeito mostrará uma preferência maior pela alavanca esquerda.

Posteriormente, Squires e Fantino (1971) revisaram o modelo matemático para que o atraso sinalizado passasse a ser ajustado com base na taxa de reforço das alternativas disponíveis, sendo descrito assim:

$$\frac{B_L}{B_L + B_R} = \frac{r_L (T - t_L)}{r_L (T - t_L) + r_r (T - t_R)} \quad (5)$$

Para $t_L < T$ e $t_R < T$,

= 1, para $t_L < t$, $t_R > T$,

= 0, para $t_L > T$, $t_R < T$,

onde B_L e B_R representam a taxa de respostas nos elos iniciais esquerdo e direito, respectivamente; r_L e r_R são as taxas gerais de reforço primário nas opções

esquerda e direita; T é o tempo médio total até o reforço primário a partir do início da fase de escolha e t_L e t_R são os tempos até o reforço medidos a partir dos elos terminais nas teclas esquerda e direita (Squires & Fantino, 1971).

O modelo proposto por Fantino pode ser reduzido à Lei da Igualação quando não há elos terminais; contudo, diferente da versão de Herrnstein (1961), a DRT considera que a redução relativa do atraso para o reforço primário modula o efeito da taxa relativa do reforçamento primário na escolha (O'daly & Fantino, 2003).

A DRT prevê que o valor de um estímulo do elo terminal é influenciado pelo atraso geral ao reforço primário, que resulta da combinação dos atrasos dos elos iniciais e terminais. De acordo com o modelo, o valor dos elos terminais diminui à medida que o atraso médio geral aumenta. Assim, um aumento no atraso dos elos iniciais reduz o valor relativo do elo terminal (Fantino et al., 1993).

A teoria sugere ainda que a escolha não é diretamente afetada pela taxa do reforço condicionado, mas sim pelo efeito deste reforçador condicionado na taxa geral de reforço primário. Portanto, aumentar a frequência do reforço condicionado não altera significativamente a preferência, a menos que haja um aumento correspondente na taxa de reforço primário (Fantino et al., 1991). Outras previsões da teoria estão relacionadas a preferência determinada exclusivamente pela taxa de reforço primário quando os elos terminais das opções são iguais, e à preferência pelo elo terminal mais curto quando os elos terminais são desiguais entre as opções (Fantino et al., 1993).

Shahan (2010) afirma que, com a proposta da DRT por Fantino (1969), foi desenvolvido um conjunto de modelos explicativos diferentes para explicar o efeito do reforçamento condicionado na escolha, como o HVA (Mazur, 1989).

O HVA, proposto por Mazur (1989), parte da descrição de como o valor de um reforçador diminui com o aumento do atraso na sua entrega. Esse modelo é baseado na

ideia de que o valor de um reforçador decai de forma hiperbólica conforme o tempo de atraso aumenta, conforme descrito pela equação:

$$V = \frac{A}{(1 + KD)} \quad (6)$$

em que V representa o valor do reforçador entregue após um certo atraso (D); A é o valor do reforçador se fosse entregue imediatamente e K é um parâmetro que indica a rapidez com que o valor decai com o aumento do atraso (Mazur, 1989).

Inicialmente, esse modelo teve como parâmetro o desempenho em procedimentos concorrentes, simples ou encadeados, em que variava apenas o atraso na entrega do reforçamento entre as opções (e.g., McDiarmid & Rilling, 1965; Mazur et al., 1985). Apesar do debate sobre os ajustes do modelo, essa equação descreveu satisfatoriamente o fenômeno quando o atraso era fixo (Mazur, 1984, 2001).

Contudo, estudos têm demonstrado que os sujeitos mostram preferência por alternativas com atrasos variáveis, mesmo quando o atraso médio é o mesmo que o de alternativas fixas (e.g. Cicerone, 1976, Mazur, 1986; Bateson & Kacelnik, 1995). Assim, o modelo hiperbólico foi estendido para reforçadores entregues após atrasos variáveis e para reforçadores probabilísticos. Quando os atrasos são variáveis, a equação hiperbólica foi generalizada para calcular o valor total de uma alternativa como uma média ponderada dos valores dos possíveis atrasos, considerando suas probabilidades de ocorrência (Mazur, 2001). Sendo apresentado da seguinte forma:

$$V = \sum_{i=1}^n P_i \left(\frac{A}{1 + kD_i} \right) \quad (7)$$

Nessa equação, V é o valor de uma opção que pode apresentar n atrasos possíveis em uma tentativa e P é a probabilidade de que um atraso de D_i segundos ocorra (Mazur, 1989).

Esse modelo permite explicar por que em situações nas quais os sujeitos podem escolher entre atrasos fixos e variáveis, acabam apresentando uma preferência por atrasos variáveis, mesmo com a média de atrasos sendo equivalente entre as opções (Mazur, 2001). Assim, em um experimento hipotético em que o sujeito tem a opção de escolher entre uma opção fixa que entrega 2 pellets de reforço a cada 8s e a outra opção que também entrega 2 pellets de reforço, mas com um atraso variável de 3s ou 15s, cada um com 50% de chance de ocorrência. A equação seria:

$$V_F = \sum_{i=1}^n P_i \left(\frac{A}{1 + kDi} \right) = \frac{2}{(1 + 1 \times 8)} = \frac{2}{9} = 0,22$$

$$V_V = \sum_{i=1}^n P_i \left(\frac{A}{1 + kDi} \right) = 0,5 \frac{2}{(1 + 1 \times 3)} + 0,5 \frac{2}{(1 + 1 \times 15)} = 0,5 \frac{2}{4} + 0,5 \frac{2}{16} = 0,31$$

Prediz que o sujeito preferirá a alternativa variável, pois o valor médio é maior, mesmo que um dos atrasos seja longo.

Em experimentos com reforçadores probabilísticos, em que o reforçador é entregue em apenas uma fração das tentativas, o modelo foi adaptado para tratar esses reforçadores como se fossem entregues após atrasos variáveis. Isso permitiu que o modelo previsse com precisão os pontos de indiferença em experimentos em que o reforço era incerto (Mazur, 2001).

O autor também observou que essa reformulação do modelo conseguia lidar com situações envolvendo reforçadores probabilísticos. Rachlin et al. (1986) já tinham proposto que reforçadores probabilísticos poderiam ser considerados análogos a reforçadores atrasados, pois, nesses contextos, a entrega do alimento não ocorre imediatamente após cada resposta de escolha, produzindo um atraso efetivo até a obtenção do reforçador (Mazur, 1991). Contudo, o modelo de Rachlin et al. (1986) propunha que a escolha entre reforçadores certos e probabilísticos poderia ser prevista

considerando apenas a probabilidade e o atraso médio do reforço, o que não foi observado de fato (Mazur, 1989).

Para explicar os dados com reforço probabilístico com mais precisão, Mazur (1989) introduziu a ideia de que o valor de um reforço probabilístico depende do tempo cumulativo em que o sujeito está exposto aos estímulos que precedem o reforço, os chamados estímulos condicionados. Esses estímulos funcionam como indicadores da entrega futura do reforço e modulam a preferência, de modo que a presença ou ausência desses estímulos em tentativas não reforçadas altera o ponto de indiferença, mesmo que a probabilidade e o atraso médio do reforço não mudem.

Logo, essa abordagem mostrou-se mais precisa na previsão do comportamento em escolhas envolvendo reforço probabilístico (Mazur, 1991) e gerou previsões contraintuitivas, confirmadas empiricamente, como o aumento da preferência por reforçadores probabilísticos quando a duração dos estímulos condicionados é reduzida, mesmo mantendo-se constantes os atrasos e as probabilidades de ocorrência do alimento.

Outro modelo relevante na área da escolha é denominado Teoria da Informação e foi revisado por Shahan e Cunningham (2015). A teoria original foi desenvolvida como uma alternativa aos modelos explicativos tradicionais de reforço condicionado baseados em princípios pavlovianos. Ela se apoiava na ideia de que a aquisição de informações sobre eventos biologicamente relevantes atuaria como um reforçador intrínseco pela redução da incerteza, independentemente de qualquer associação direta com reforços primários, como comida ou água (Dinsmoor, 1983). Contudo, esse modelo falhou em explicar os dados de experimentos em que os estímulos indicavam a ausência de reforço (S-), visto que esses estímulos não funcionavam como reforçadores

condicionados, mesmo que reduzissem a incerteza sobre a disponibilidade do reforço (Dinsmoor, 1983; Shahan & Cunningham, 2015).

Posteriormente, foram realizados refinamentos na Teoria da Informação aplicada ao condicionamento respondente (Shahan & Cunningham, 2015). Nessa nova versão, o estímulo neutro torna-se condicionado (CS) quando fornece informações sobre quando um estímulo incondicionado (US) ocorrerá, ou seja, quando reduz a incerteza temporal sobre o US. Dessa forma, os estímulos que proporcionam maior redução de incerteza temporal sobre o momento de reforço (S+) mantêm o comportamento por mais tempo do que os estímulos que apenas indicam que o reforço não está disponível (S-) (Shahan & Cunningham, 2015).

Por exemplo, a quantidade de informação que um estímulo condicionado (CS) fornece sobre o tempo até a apresentação do reforço pode ser medida como a diferença entre a incerteza inicial e a incerteza restante após a apresentação do CS. Essa diferença é medida em bits e pode ser formalizada como:

$$H_{com} = \log_2 \frac{C}{t} \quad (8)$$

Onde H_{com} é a quantidade de informação que o CS transmite sobre o tempo até a apresentação do US; C é o tempo médio entre as apresentações dos US; e t é o tempo médio entre as apresentações do CS e do US.

Dessa forma, esse modelo prevê que em procedimentos em que um sujeito pode emitir uma resposta de observação para obter informação sobre a disponibilidade de reforço (por exemplo, se está em um período de reforço ou extinção), o sujeito vai preferir respostas que produzem estímulos que reduzem a incerteza sobre o tempo de reforço. Por exemplo, em um procedimento de observação, períodos de reforço sob um esquema VI 50 s alternam com períodos de extinção, com probabilidade de ocorrência do reforço $p = 0,4$. Nessa condição, a aplicação da equação seria:

$$Hcom = \log_2 \left(\frac{C}{t} \right) = \log_2 R \frac{50}{50} S = \log_2 \left(\frac{125}{50} \right) = \log_2 (2,5) = 1,32$$

A quantidade de informação transmitida seria $Hcom = 1,32$ bits. Mas caso a probabilidade de ocorrência do reforço diminua para $p = 0,1$, a equação ficaria:

$$Hcom = \log_2 \left(\frac{C}{t} \right) = \log_2 R \frac{50}{0,1} S = \log_2 \left(\frac{500}{50} \right) = \log_2 (10) = 3,32$$

E a quantidade de informação transmitida seria $Hcom = 3,32$ bits, indicando que o estímulo condicionado agora fornece uma redução ainda maior da incerteza temporal, o que, segundo o modelo, o torna um reforçador condicionado mais potente. Portanto, conforme a probabilidade do reforço sinalizado diminui, o valor informativo do estímulo aumenta, potencialmente fortalecendo o controle do comportamento por esse estímulo.

Já em procedimentos mais complexos, como os de cadeias concorrentes, o valor relativo da opção é uma função da razão dos valores de $Hcom$ para os estímulos de cada opção (Shahan & Cunningham, 2015). Assim, o modelo matemático consistiria em:

$$\frac{Hcom1}{Hcom2} = \frac{\log_2 \left(\frac{C}{t_1} \right)}{\log_2 \left(\frac{C}{t_2} \right)} = \frac{\log_2 C - \log_2 t_1}{\log_2 C - \log_2 t_2} \quad (9)$$

Com esses parâmetros, a teoria prevê que em contextos em que a taxa relativa de reforço é variada no elo terminal da cadeia, a razão dos $Hcom$ diminui à medida que as durações dos elos iniciais aumentam e tem seu valor aumentado à medida que as durações dos elos terminais aumentam (Shahan & Cunningham, 2015).

Cunningham e Shahan (2018) aplicaram o modelo da Teoria da Informação a procedimentos de observação e constataram que animais, especialmente pombos, preferem estímulos informativos, mesmo quando essas opções fornecem menos reforço. Os pombos se destacam nessas escolhas, pois dependem mais de estímulos visuais,

demonstrando uma preferência acentuada por sinais que indicam a presença ou ausência de reforço, em comparação com animais que dependem menos de estímulos visuais, como os ratos. Outra comparação feita pelos autores foi em relação ao ambiente em que essas opções são apresentadas. No ambiente natural, a preferência por estímulos informativos é adaptativa, pois auxilia na identificação de recursos necessários para a sobrevivência. No entanto, em ambientes controlados, como os de laboratórios, esse mesmo comportamento se torna mal adaptativo, levando à perda de recursos. Assim, o comportamento evolutivo que é benéfico para a sobrevivência depende do contexto em que é emitido.

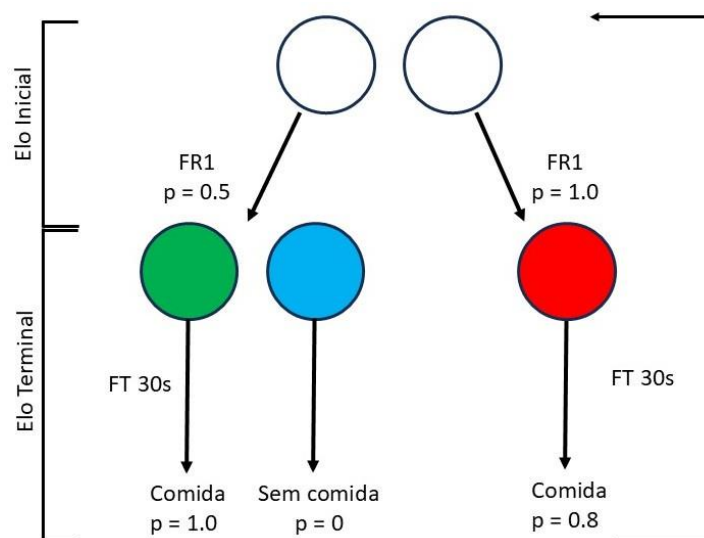
Uma outra abordagem que destaca a função dos estímulos condicionados em procedimentos de escolha é a SiGN. Esse modelo foi derivado diretamente da DRT, mas estende a teoria para situações com resultados probabilísticos e sinalização explícita do resultado. Dessa forma, ele consegue explicar a preferência por alternativas que fornecem menos reforço, fenômeno conhecido como escolha subótima (Dunn et al., 2024).

Esse tipo de preferência foi observado no estudo do Experimento 3 de Dunn e Spetch (1990), em que cinco pombos foram submetidos a um procedimento de cadeias concorrentes para avaliar os efeitos de estímulos sinalizados versus não sinalizados na escolha. Os elos iniciais consistiam em chaves brancas apresentadas sob esquemas de resposta FR 1 ou em VI 80 s, enquanto os elos terminais tinham duração de 30 a 50 s. Nas condições sinalizadas, a presença de estímulos discriminativos indicava a disponibilidade de reforço em 50%, 60% ou 80% das tentativas, enquanto nas condições não sinalizadas nenhuma informação adicional era fornecida. Os resultados mostraram que os pombos exibiram preferência marcante pelas alternativas sinalizadas, respondendo mais frequentemente às chaves que forneciam informação sobre a

disponibilidade de reforço, mesmo quando a taxa de reforço era igual entre as alternativas. Alterações no esquema do elo inicial (FR 1 vs. VI 80) não modificaram substancialmente essa preferência, indicando que o valor informativo dos estímulos do elo terminal foi o principal determinante da escolha (ver Figura 4).

Figura 4

Ilustração de um procedimento de cadeia concorrente com reforçamento probabilístico



Nota: Ilustração de um procedimento de cadeia concorrente com reforçamento probabilístico. Os elos iniciais (cor branca) estão em FR1. Após a escolha por uma das opções, o estímulo muda e a opção alternativa fica apagada. Caso o sujeito escolha a opção esquerda, o estímulo verde tem probabilidade de 50% de ocorrer no elo terminal e sinaliza que o estímulo reforçador será apresentado ao final do tempo fixo de 30 s. Contudo, essa opção também tem a probabilidade de 50% de apresentar o estímulo azul no elo terminal, que sinaliza que o reforço não será disponibilizado ao final do tempo fixo de 30 s. Caso o sujeito escolha a opção direita, o estímulo vermelho será apresentado no elo terminal em 100% das vezes, mas a apresentação do estímulo reforçador ao final do tempo fixo de 30 s ocorrerá em 80% das tentativas.

A partir disso, esses autores buscaram explicar a escolha subótima com base na função dos estímulos discriminativos de reduzir o tempo para a apresentação do estímulo reforçador. Essa função acabaria por aumentar o valor da opção com menor probabilidade de reforço em comparação com a opção alternativa que não fornecia nenhuma sinalização diferencial.

O SiGN parte do pressuposto de que, no contexto de escolhas subótimas, um sinal que indica uma redução no tempo até o reforçador (mesmo que o reforço não seja garantido) tem um valor significativo. Esse valor adicional ocorre porque o sinal indica que a recompensa está mais próxima, tornando a opção mais rica. Assim, a característica-chave do modelo é o conceito de "bônus de boas notícias". O modelo considera que, quanto maior a incerteza sobre o reforço, maior será o impacto positivo de um sinal que indica que o reforço está a caminho. Esse "bônus" faz com que as opções sinalizadas se tornem mais atraentes, mesmo quando a chance real de obter a recompensa é menor. O SiGN prevê ainda que a escolha subótima aumenta à medida que o atraso entre a emissão da resposta e o reforçador aumenta (McDevitt et al., 2016; Dunn et al., 2024).

Os autores destacam que o termo "boas notícias" é definido operacionalmente com base em parâmetros experimentais específicos, não sendo usado para representar construtos teóricos, como "informação", como ocorre na Teoria da Informação (Dunn et al., 2024).

O modelo também considera o papel dos sinais que indicam não-recompensa (S-), que não alteram efetivamente a contingência nem têm capacidade de suprimir a resposta. Esses sinais recebem um valor fixo de 1 s na equação do modelo, refletindo apenas o processamento do estímulo sem contribuir para a redução de atraso (Dunn et al., 2024).

O SiGN parte da ideia de que a escolha entre alternativas depende tanto do reforço primário (a quantidade e rapidez com que o reforço é entregue) quanto do reforço condicionado, que é o valor que os sinais (como estímulos que indicam “boa notícia”) têm ao antecipar a chegada do reforço. Para calcular isso, primeiro determina-se o tempo médio até o reforço a partir do início do procedimento, considerando a probabilidade de cada alternativa fornecer reforço e a duração dos elos iniciais e terminais. A redução média do atraso é obtida subtraindo-se o tempo esperado até o reforço do tempo total médio e, se houver sinais de “boas notícias”, adiciona-se um bônus que aumenta o valor condicionado da alternativa. Esse bônus é ajustado por um fator β , que leva em conta o tempo dos elos iniciais e terminais e como eles influenciam a força do reforço condicionado. Por fim, a proporção de escolha prevista pelo modelo é calculada como uma relação entre a taxa de reforço primário e o reforço condicionado total de cada alternativa, combinando assim o efeito do reforço real e do sinal que antecipa o reforço, resultando na equação final do SiGN:

$$\frac{R_a}{R_a + R_b} = \frac{r_a \delta_a}{r_a \delta_a + r_b \delta_b} \quad (10)$$

$$= 1 \quad \text{quando } \delta_a > 0, \delta_b > 0$$

$$= 0 \quad \text{quando } \delta_a > 0, \delta_b < 0$$

$$\quad \text{quando } \delta_a < 0, \delta_b > 0$$

onde R_a e R_b representam o número de respostas nas opções A e B, respectivamente; r_a e r_b representam a taxa de reforço para as opções A e B; e δ_a e δ_b representam o reforço condicionado das opções A e B determinado pela soma da redução média dos atrasos (D_{avg}) e da redução do atraso bônus (D_{bonus}) ajustado pelo parâmetro β .

$$\delta = D_{avg} + \beta \cdot D_{bonus} \quad (11)$$

Para exemplificar o modelo, o Experimento 1 do estudo de McDevitt et al., (2024) sobre escolha subótima pode ser utilizado como base. Nesse estudo, dez pombos participaram de um procedimento de cadeias concorrentes. Os animais tinham que completar um esquema no elo inicial, que era FR 1 para um grupo ou VI 30 s para o outro e, posteriormente, tinham acesso a duas alternativas: a subótima, associada a um elo terminal sinalizado por cores distintas (por exemplo, branco ou verde), que fornecia alimento apenas 20% das vezes, e a ótima, cujos elos terminais não diferenciavam os resultados e entregavam alimento 50% das vezes. Cada elo terminal durava 20 s, seguido de acesso ao comedouro ou de blecaute.

Aplicando as fórmulas necessárias para o modelo no contexto em que o elo inicial é FR1, o cálculo seria: considerando a alternativa subótima (A) que apresenta probabilidade de reforço de 20%, com elo terminal (TL) de 20 s seguido de 5 s de acesso ao reforço quando ocorre e 1 s quando não ocorre, devido à sinalização do S⁻, e elo inicial (IL) FR 1 de 1 s. O tempo médio no IL até o reforço é $IL_A = 1/0,2 \times 1 = 5$ s. O tempo médio no TL é $TL_A = 0,2 \times 20 + 0,8 \times 1 = 4,8$ s por entrada, totalizando TL_A até reforço = $1/0,2 \times 4,8 = 24$ s. Assim, a taxa de reforço da alternativa subótima é $r_A = 1/(5 + 24) \approx 0,0345$. A alternativa ótima (B) tem probabilidade de reforço de 50%, TL com e sem reforço de 20 s sem sinalização e IL FR 1 = 1 s; o tempo até reforço é $IL_B = 1/0,5 \times 1 = 2$ s e no $TL_B = 20 \times 1/0,5 = 40$ s, resultando em $r_B = 1/(2 + 40) \approx 0,0238$. As proporções programadas são $P_A = 0,2/(0,2 + 0,5) = 0,286$ e $P_B = 0,714$. O tempo médio até o reforço é $T = P_A/r_A + P_B/r_B \approx 8,29 + 30,00 = 38,29$ s. O valor do reforço condicionado para a alternativa subótima é δ_A : $D_{avgA} = T \times p - TL_reforço = 38,29 \times 0,2 - 20 = -12,34$ s; $D_{bonusA} = T \times (1 - p) - TL_não_reforço + TL_reforço = 38,29 \times 0,8 - 1 + 20 = 49,63$ s; $\beta = \log_{10}(1 + 20/1) \approx 1,322$. O valor do reforço condicionado final seria:

$$\delta_A = D_{\text{avg}} + \beta \cdot D_{\text{bonus}} = -12,34 + 1,322 \times 49,63 = 53,27$$

$$\delta_B = D_{\text{avg}} + \beta \cdot D_{\text{bonus}} = 19,14 - 20 = -0,86$$

Assim, o modelo SiGN prediz que:

$$\frac{R_a}{R_a + R_b} = \frac{r_a \delta_a}{r_a \delta_a + r_b \delta_b} = \frac{0,0345 \times 53,27}{0,0345 \times 53,27 + 0,0238 \times (-86)} = 1,011$$

Logo, a opção subótima, no estudo de de McDevitt et al., (2024), deve receber o maior número de alocação de resposta na condição em que o elo inicial é FR1. O que corrobora com o resultado observado pelos autores, em que a média da proporção foi de 0,92 nessa condição.

No entanto, o SiGN tem algumas limitações, especialmente quando se trata de prever comportamentos fora dos experimentos clássicos com pombos, por exemplo, estudos com ratos (e.g. Chow et al., 2017; Cunningham & Shahan, 2019) e humanos (e.g. Lalli et al., 2000; Molet et al., 2012) demonstraram resultados mistos em relação a preferência pela alternativa subótima. No geral, os dados apontam que os pombos são mais propensos a apresentar esse padrão de comportamento (Dunn et al., 2024).

Em síntese, as teorias analisadas (DRT, HVA, Teoria da Informação e SiGN) convergem na ideia de que o tempo até a obtenção do reforço influencia o valor de uma opção. Além disso, todas reconhecem que estímulos que sinalizam a proximidade de uma recompensa podem se tornar reforçadores condicionados, na medida em que reduzem a incerteza quanto à entrega do reforço.

Contudo, existem algumas diferenças. A DRT e a HVA diferem quanto aos determinantes centrais do comportamento de escolha: a DRT propõe que a escolha é guiada pela redução relativa do atraso ao passar do elo inicial ao terminal, enquanto a HVA sugere que depende do valor adicionado nessa transição. Apesar dessas diferenças, os modelos são conceitualmente próximos e frequentemente fazem previsões

semelhantes. No entanto, há situações em que os dados favorecem um modelo em detrimento do outro. Por exemplo, a DRT não distingue entre esquemas de reforço fixos e variáveis, enquanto a HVA não prevê o efeito de diferença constante, em que a proporção de escolha permanece a mesma quando a diferença entre os elos terminais é constante, independentemente de suas durações (Mazur, 2001).

Nesse contexto, o SiGN pode ser entendido como uma extensão da DRT, projetada para lidar com as limitações observadas por Mazur, especialmente na previsão da escolha entre alternativas com reforço probabilístico e sinalizado. Além disso, o SiGN modela detalhadamente como estímulos sinalizadores do reforço (S+) ou da não entrega do reforço (S-) influenciam o valor da alternativa, diferentemente do modelo DRT.

No caso do HVA, embora o modelo também não descreva explicitamente o papel do S-, sua aplicação em contextos de reforçadores probabilísticos permite inferir efeitos indiretos do sinal de não-reforço. O valor de uma alternativa é determinado pelo tempo gasto na presença de estímulos condicionados que antecedem o reforço, e não pelo tempo total até a entrega da recompensa. Dessa forma, quando a comida não é fornecida, o impacto se reflete na ausência ou na redução do tempo de exposição a esses estímulos, o que modifica a preferência do sujeito, mesmo que a probabilidade de reforço permaneça constante.

Na mesma direção do SiGN, a Teoria da Informação oferece uma base formal para quantificar o papel orientador dos estímulos, entendendo-os como sinais que guiam o comportamento do organismo para os reforços distribuídos no tempo e, potencialmente, em outras dimensões, como o espaço ou as relações entre eventos. Diferentemente da DRT, que diz, de forma não formalizada, que um estímulo condicionado tem valor porque reduz o atraso até o reforço, a Teoria da Informação

mostra que essa redução de atraso pode ser vista como informação; ou seja, o valor do estímulo é proporcional à quantidade de incerteza temporal que ele elimina e propõe uma medida quantitativa formal, em bits, do quanto o estímulo informa sobre o reforço (Shahan & Cunningham, 2015).

Diferentemente do SiGN, a Teoria da Informação explica o valor de um estímulo em termos de quanto ele orienta o comportamento do organismo em direção aos reforços distribuídos no tempo. Nessa perspectiva, a sinalização de reforço (S+) pode reduzir diretamente a incerteza temporal ao estar associada à apresentação do estímulo reforçador, enquanto a sinalização de não reforço (S-) pode afetar indiretamente a contingência ao contribuir para a estimativa do tempo até a entrega do próximo reforço. Dessa forma, ambos os tipos de sinal podem influenciar a taxa e a alocação de respostas. Em contraste, o modelo SiGN não prevê mudanças comportamentais significativas em função do S-, assumindo que sinais de não reforço contribuem pouco para explicar a distribuição da resposta entre alternativas.

Diante dessas considerações, o presente estudo pretende utilizar esses quatro modelos (DRT, HVA, SiGN e Teoria da Informação) como ferramentas teóricas complementares para interpretar os efeitos do TO sobre o comportamento. Ao comparar como cada modelo prevê o impacto de atrasos e de estímulos sinalizadores de reforço ou de não-reforço, será possível compreender de forma mais detalhada os mecanismos que determinam a aversividade do TO. Dessa forma, a integração dessas abordagens pode oferecer uma base robusta para analisar e interpretar os dados experimentais, permitindo avaliar quais aspectos do TO são mais bem capturados por cada modelo e como eles se relacionam.

Timeout Sob a Perspectiva das Teorias do Atraso do Reforço

Embora as principais teorias sobre atraso do reforço não incluam referências explícitas ao procedimento de TO, suas previsões acerca de como a inserção de um intervalo entre a resposta e a entrega do estímulo reforçador afeta o comportamento podem oferecer um arcabouço útil para interpretar os efeitos do TO. O presente trabalho não pretende desenvolver um modelo formal específico para esse tipo de contingência, dado que os modelos de atraso foram originalmente formulados para contextos particulares (e.g., cadeias concorrentes) e exigiriam adaptações substanciais para serem aplicados diretamente ao TO. Em vez disso, buscamos destacar aspectos conceituais centrais dessas teorias, explorando como eles podem contribuir para a compreensão do impacto do TO sobre a alocação de respostas.

O TO pode ser concebido como uma variável que introduz um atraso adicional até a entrega do reforço, aumentando o tempo total entre reforçadores para a alternativa em que é programado. Sob a perspectiva das principais teorias sobre atraso do reforço (DRT, HVA; Teoria da Informação e o modelo SiGN), a introdução do TO deveria reduzir o valor relativo da alternativa em que é aplicado. Contudo, ressalta-se ainda que, como colocado na seção anterior, cada uma dessas teorias sobre atraso aborda o papel do tempo até o reforço de maneira distinta, o que pode gerar previsões diferentes sobre os efeitos do TO. A seguir, são descritos os mecanismos centrais de cada modelo e suas implicações para o TO:

1- A DRT postula que a escolha é guiada pela redução relativa do atraso ao transitar do elo inicial ao elo terminal. Assim, qualquer intervalo adicional, como o TO, aumenta o atraso total até o reforço, diminuindo a redução relativa do atraso e, portanto, o valor da alternativa em que ocorre.

2- A HVA, por sua vez, descreve o valor de estímulos condicionados como sujeito a um desconto hiperbólico em função do atraso. Nesse modelo, a introdução do TO também

reduz o valor da alternativa ao aumentar o atraso efetivo até o reforço. Contudo, devido à forma hiperbólica da função, aumentos sucessivos no atraso produzem reduções progressivamente menores no valor, resultando em efeitos de saturação quando o TO se torna muito longo. Embora não trate explicitamente do papel de estímulos sinalizadores, a HVA também implica que o S⁻ teria pouco efeito direto sobre a escolha, uma vez que o que determina o valor da opção é o tempo na presença do S⁺.

3- A Teoria da Informação enfoca não apenas o atraso em si, mas a quantidade de informação temporal que um estímulo fornece sobre a entrega do reforço. Nessa perspectiva, o TO não é avaliado exclusivamente pelo tempo que acrescenta, mas pelo quanto contribui para a incerteza sobre quando o próximo reforço ocorrerá. Ainda assim, o sinal de não reforço (S⁻) é considerado apenas um contribuinte indireto, na medida em que afeta o tempo total até o próximo reforço, mas não exerce influência direta robusta sobre a escolha.

4- De modo semelhante, o modelo SiGN atribui papel central aos estímulos discriminativos, especialmente ao S⁺, que sinaliza a disponibilidade de reforço. Já o S⁻ tem valor limitado, pois apenas o momento em que o organismo identifica que se encontra sob S⁻ é relevante; o tempo restante sob esse estímulo não é incorporado ao cálculo de valor. Assim, o S⁻ tende a exercer pouco impacto direto sobre a alocação da resposta.

Apesar das diferenças nos mecanismos explicativos, todos os modelos convergem na previsão de que a introdução de um atraso adicional, como o TO, tende a diminuir o valor da alternativa e, portanto, reduzir sua escolha. Essa interpretação é consistente com achados de estudos sobre TO em contingências simples, nos quais a introdução de um período sinalizado sem acesso ao reforço reduziu a taxa de resposta. Por exemplo,

Kaufman e Baron (1968) expuseram pombos a diferentes durações de TO, observando que a supressão da resposta aumentava proporcionalmente à duração do atraso imposto.

Sob a perspectiva da DRT, tal efeito decorre da redução do valor relativo da alternativa à medida que o atraso total para o reforço aumenta. Resultados similares foram observados em estudos que manipularam a frequência de aplicação do TO. Thomas (1968), por exemplo, programou TOs frequentes sob um esquema de razão fixa (FR 2) e verificou maior supressão da resposta, o que pode ser interpretado como resultado da desvalorização cumulativa da alternativa devido à repetida inserção de atrasos.

Sob a perspectiva da HVA, o TO pode afetar essas duas variáveis principais para estabelecer o valor de uma opção: o atraso e a probabilidade. Primeiro, ao impor um intervalo sem acesso ao reforço, o TO aumenta o atraso efetivo entre a resposta e a próxima oportunidade de reforço, o que, segundo o modelo, reduz o valor da alternativa. Segundo, quando o TO é sinalizado, ele pode indicar que o reforço não ocorrerá naquele ciclo, o que pode ser interpretado como uma redução na probabilidade percebida de reforço. Assim, o TO pode levar a uma redução acentuada do valor subjetivo da alternativa afetada.

No entanto, nem todos os resultados empíricos se alinham a essa previsão linear. Nader e Morgan (2001), ao manipular a duração do TO, verificaram que incrementos na duração do TO não necessariamente aumentam a supressão da resposta. Embora a DRT preveja uma redução inicial da taxa de resposta sempre que o estímulo sinaliza um atraso adicional para o reforço, ela não fornece uma explicação clara para a estabilização desse efeito diante de TOs mais longos.

Uma possível interpretação para esse padrão é oferecida pelo modelo HVA. Esse modelo postula que o valor de um reforçador decai de forma hiperbólica à medida que

aumenta o atraso até sua entrega. Nessa perspectiva, a introdução de um TO equivale a um incremento no atraso total, reduzindo o valor relativo da alternativa. Contudo, devido à forma da função hiperbólica, aumentos adicionais no atraso, como no caso de TOs progressivamente mais longos, produzem reduções cada vez menores no valor do reforço, resultando em um efeito de saturação. Esse padrão é consistente com resultados experimentais em que durações mais longas de TO não geraram supressão adicional proporcional.

Em procedimentos concorrentes, Dunn (1990) observou que a manipulação da taxa de TO entre alternativas provocou uma mudança na alocação de respostas em favor da opção com menor taxa de TO, sem, entretanto, reduzir a taxa global de respostas. Quando o TO é aplicado igualmente a ambas as alternativas, como no experimento de Dunn, a DRT sugere que a redução relativa do atraso permanece semelhante entre elas, de modo que o valor do estímulo condicionado não é substancialmente afetado em termos absolutos. Nesse contexto, pequenas diferenças relativas entre os elos são suficientes para deslocar a proporção de escolhas, mas não para reduzir a força absoluta do comportamento.

Previsões semelhantes emergem a partir do modelo HVA, que considera o desconto hiperbólico do valor do reforço em função do atraso e o papel do tempo cumulativo de exposição ao estímulo condicionado (CS). Quando ambas as alternativas incluem TO, o atraso total aumenta para ambas; porém, devido à natureza hiperbólica da função de desconto, incrementos adicionais no atraso produzem efeitos cada vez menores (saturação), de modo que o valor subjetivo global não se reduz significativamente e as taxas absolutas permanecem estáveis. Pequenas diferenças na variabilidade do atraso ou na duração do CS, contudo, podem alterar a média ponderada do valor, deslocando a preferência sem suprimir a resposta.

Em contraste, em cadeias concorrentes, como no estudo de Carlson (1972), em que o TO foi aplicado apenas a uma alternativa e a taxa de reforço foi igualada, observou-se não apenas uma preferência pela alternativa sem TO, mas também uma redução na taxa de resposta da alternativa com TO. Pela perspectiva da DRT, o TO modifica o perfil temporal do elo terminal daquela alternativa, reduzindo o seu valor condicionado, mesmo com a taxa de reforço semelhante entre as opções. Em cadeias concorrentes, a alocação de respostas pode depender também do valor do estímulo condicionado do elo terminal; portanto, uma depreciação significativa no tempo restante para o reforço ($T - t$) pode levar não apenas à realocação de escolha, mas também à diminuição da taxa absoluta de respostas nesse elo.

Essa lógica também encontra suporte na HVA: o TO, em apenas uma cadeia, cria um aumento focalizado do atraso daquela alternativa (maior atraso efetivo), que, pelo desconto hiperbólico, reduz substancialmente seu valor subjetivo. Além disso, se o TO prolonga o tempo cumulativo de exposição ao CS, mantendo o sujeito mais tempo sob estímulos que não sinalizam reforço, o HVA prevê uma redução adicional do valor condicionado, resultando em menor alocação de respostas nessa opção.

Embora, na maioria dos casos, seja previsível que o TO, enquanto intervalo adicional entre reforçadores, reduza a taxa de respostas ou, ao menos, desloque as respostas em direção à alternativa sem TO, nem todos os resultados podem ser explicados apenas pela extensão do intervalo ou pela alteração da taxa de reforço. Por exemplo, no Experimento 1 de Toegel et al. (2022), a condição de TO, embora tenha diminuído a taxa de respostas em relação à linha de base, apresentou uma taxa maior do que a da condição de intervalo acoplado, que possuía a mesma taxa de reforço, mas não incluía estímulos sinalizando a ausência de reforço.

De forma semelhante, Lattal (1984) observou, no Experimento 2, que blecautes não contingentes produziram uma maior supressão da resposta do que blecautes contingentes, mesmo quando ambos resultavam em períodos sem reforço. Já no Experimento 3, tanto os blecautes contingentes quanto os não contingentes aumentaram a taxa de respostas em relação à linha de base, que não incluía nenhum estímulo sinalizando a ausência de reforço, sugerindo que a presença de sinalização, mesmo associada a eventos não reforçadores, pode exercer controle discriminativo e alterar a taxa de resposta.

No estudo de Toegel et al. (2022), a presença de sinais discriminativos no TO parece atenuar a supressão da resposta. A Teoria da Informação sugere que o TO funciona como um estímulo informativo, indicando quando o reforço será retomado e, assim, reduzindo a incerteza temporal. Em contraste, o intervalo acoplado não sinalizado não fornece informação adicional, aumentando a incerteza e provocando uma redução mais intensa na taxa de respostas. Assim, os estímulos associados ao TO podem atuar como reforçadores condicionados, permitindo que o sujeito antecipe que o reforço ocorrerá após o período de TO.

De maneira complementar, o modelo SiGN enfatiza que estímulos que sinalizam ausência de reforço (S^-) exercem efeito supressivo limitado sobre a escolha, independentemente de sua duração (Dunn et al., 2023; McDevitt et al., 2016). Entretanto, Sears et al. (2022) observaram que, mesmo que um S^- tenha pouco efeito supressivo frente a um estímulo de reforço, ele pode ser preferido em situações em que não há informação sobre a entrega de reforço. No contexto de Toegel et al. (2022), isso explica por que a condição de TO sinalizado reduziu menos a taxa de respostas do que a condição de intervalo acoplado não sinalizado: embora o TO funcione como um S^- ,

ainda fornece informação que aumenta seu valor relativo frente à alternativa sem sinalização.

Para facilitar a interpretação dos dados de Toegel et al. (2022), pode-se conceber a situação como um esquema encadeado. Nessa extrapolação, a alternativa com TO pode ser entendida como o elo inicial de uma cadeia, enquanto o estímulo de *time-in* (por exemplo, a luz que sinaliza o acesso ao reforço) funciona como o elo terminal, indicando a entrega do reforço primário. Embora o estímulo diferencial presente durante o TO sinalize a ausência de reforço, o estímulo de *time-in* pode adquirir função de reforçador condicionado, pois reduz o atraso percebido até o reforço efetivo. Por outro lado, a condição de intervalo acoplado não inclui estímulo discriminativo que indique a redução do atraso, deixando o sujeito sem referência clara sobre quando o reforço ocorrerá. Isso torna a alternativa menos reforçadora e ajuda a explicar a supressão mais intensa da taxa de respostas nessa condição.

Explicações análogas podem ser aplicadas aos achados de Lattal (1984). Quando o blecaute é contingente à resposta, ele se torna mais previsível, funcionando como um sinal consistente que reduz a incerteza temporal sobre a disponibilidade de reforço. Em contraste, condições em que o blecaute é não contingente, isto é, ocorre independentemente do comportamento, podem aumentar a imprevisibilidade do ambiente, favorecendo a supressão da resposta, como observado no Experimento 2. Por outro lado, a sinalização da contingência com blecaute pode não ter tido impacto na taxa de resposta, como defende o SiGN em relação à função do S-, como foi o caso do Experimento 1 de Lattal.

Os estudos sobre o TO indicam que a presença de estímulos que sinalizam a redução do atraso até o reforço pode alterar significativamente o efeito do procedimento sobre a taxa e a alocação de respostas. Estímulos discriminativos claros, que indicam a

proximidade do reforço, parecem aumentar o valor da alternativa, mesmo quando o TO está presente. Nesse sentido, a interpretação do TO à luz das teorias do atraso sugere que, embora o intervalo adicional imposto pelo procedimento contribua para a desvalorização da opção, os estímulos associados ao TO e à presença ou ausência de reforço também desempenham um papel determinante na resposta dos sujeitos. Esse mecanismo explicaria por que, em algumas condições, o TO não resulta em supressão significativa da resposta, o que aponta para a necessidade de investigações futuras sobre os fatores que modulam a eficácia do procedimento.

Embora as teorias analisadas ofereçam importantes ferramentas conceituais para compreender os efeitos do TO, é fundamental reconhecer suas limitações. A DRT e a HVA, por exemplo, foram originalmente formuladas para explicar o comportamento em cadeias concorrentes com reforçamento positivo, o que levanta questionamentos sobre a validade de suas aplicações diretas ao TO, que envolve componentes aversivos e supressivos. Nesses casos, extrapolações precisam ser feitas com cautela, já que os mecanismos de escolha podem funcionar de maneira distinta quando o atraso ao reforço é mediado por estímulos negativos ou sinais de não reforço. Por outro lado, tanto a Teoria da Informação quanto o modelo SiGN têm o mérito de incorporar formalizações específicas sobre o papel dos estímulos sinalizadores. No entanto, essas abordagens também enfrentam limitações, visto que nenhuma delas trata de eventos como o TO.

Portanto, embora os modelos discutidos contribuam para avançar a compreensão dos efeitos do TO, é necessário um desenvolvimento teórico mais abrangente e integrado que considere a complexidade das contingências reais e das variáveis que modulam o valor dos estímulos sinalizadores em contextos como o procedimento do TO.

Conclusão

Os fatores que modulam a aversividade do TO, como sua duração e a presença de sinais que indicam o período sem reforço, são cruciais para compreender como ele influencia o comportamento. Embora o TO seja amplamente empregado para reduzir comportamentos inadequados na prática aplicada, evidências sugerem que a simples imposição de um intervalo sem reforço nem sempre é suficiente para explicar ou garantir seus efeitos punitivos.

Ao considerar o TO não como um estímulo discreto, como um choque elétrico, mas como um tempo adicional e sinalizado entre apresentações de estímulos reforçadores, as teorias do atraso do reforço oferecem uma perspectiva alternativa para entender seus mecanismos. Essas teorias permitem avaliar não apenas a redução na taxa de reforço ou o adiamento da entrega do reforçador, mas também como a sinalização da contingência pode alterar o valor condicionado de um estímulo, modulando a preferência e a taxa de resposta.

Em termos teóricos, o TO adiciona atraso entre a resposta e o reforço, o que, segundo essas abordagens, deveria reduzir o valor da opção e, conseqüentemente, a taxa de resposta. No entanto, os resultados experimentais mostram variação considerável. Fatores como a presença de estímulos discriminativos claros e a forma como o tempo de espera é sinalizado podem atenuar ou até mesmo neutralizar a supressão esperada, indicando que a eficácia do TO depende da interação entre múltiplos elementos contextuais.

Essa complexidade evidencia que o TO é um procedimento multifacetado, cujo impacto não pode ser explicado apenas pelo intervalo imposto. Futuras pesquisas devem explorar mais profundamente a aplicação desses modelos teóricos a situações com TO, incluindo ajustes nos modelos existentes para capturar melhor contextos típicos de aplicação prática. Estudos que utilizem cadeias concorrentes são particularmente

promissores, pois permitem testar previsões dos modelos do atraso de forma mais direta.

Em resumo, a eficácia do TO depende não apenas da duração do intervalo, mas também da sinalização e da interação com o reforço. A compreensão completa desses mecanismos é essencial para que o TO possa ser aplicado de maneira precisa e consistente em diferentes contextos de controle comportamental. Dessa forma, a pesquisa futura desempenha um papel central para elucidar os fatores que otimizam a supressão de respostas e ampliar a integração entre a teoria experimental e a prática clínica ou aplicada.

Estudo 2

Efeito da Duração e da Taxa de Apresentação do TO na Alocação da Resposta em Situação de Escolha

O controle aversivo inclui três das quatro relações operantes tradicionalmente descritas pela análise do comportamento: reforçamento negativo, punição positiva e punição negativa (Pierce & Cheney, 2013). O procedimento de timeout (TO), tema do presente trabalho, é considerado uma forma de punição negativa, pois consiste em um procedimento no qual, contingente à emissão de uma resposta-alvo, ocorre a remoção da disponibilidade de estímulos reforçadores, com o objetivo de reduzir sua frequência futura. Formalmente, na Análise do Comportamento, o TO é definido como um período de tempo sinalizado em que os estímulos reforçadores não estão disponíveis, contingente a uma resposta-alvo (Defulio & Hackenberg, 2007; McMillan, 1967).

O TO tornou-se uma técnica amplamente empregada em contextos aplicados, como clínicas e escolas, sendo considerado o procedimento punitivo mais utilizado para reduzir comportamentos problemáticos, especialmente em populações com desenvolvimento atípico, transtornos de atenção e dificuldades de regulação comportamental (Lydon et al., 2015).

Apesar de sua aceitação prática, ainda são pouco compreendidos os mecanismos que tornam o TO um evento aversivo eficaz (Hackenberg & Defulio, 2007). Em ambientes experimentais, diferentes estudos têm buscado isolar variáveis envolvidas nesse procedimento. Toegel et al. (2022), por exemplo, propuseram cinco atributos centrais que caracterizam o TO como punição: 1) contingência à resposta-alvo; 2) suspensão da contingência de reforçamento; 3) mudança de estímulos que diferenciem o TO do time-in; 4) eliminação do estímulo reforçador; 5) redução da taxa geral de reforço. A presença desses atributos contribui para que o TO funcione como uma

contingência aversiva eficaz. No entanto, sua eficácia também depende de variáveis adicionais.

Entre os parâmetros que modulam a eficácia do TO como estímulo punitivo, destacam-se a duração e o esquema de apresentação. A duração do TO, por exemplo, é frequentemente manipulada como indicador de sua intensidade ou magnitude, de forma análoga ao uso de choques em estudos sobre punição positiva. Embora alguns dados sugiram que aumentos na duração produzem maior supressão da resposta (Ferster & Appel, 1961; Kaufman & Baron, 1968), a relação entre esse parâmetro e a eficácia punitiva do TO tem sido inconsistente (Toegel et al., 2022).

Em relação aos esquemas de apresentação, há maior consenso de que esquemas mais densos de punição, como razões fixas mais baixas (e.g., FR 2), produzem maior supressão do que esquemas mais espaçados (Thomas, 1968). Nos esquemas de intervalo variável, os dados são mais ambíguos, mas resultados recentes (Dunn, 1990; Hall, 2015) sugerem que a frequência de aplicação do TO, independentemente do tipo de esquema, é um fator determinante para sua função punitiva. Assim, quando o TO é aplicado com maior frequência, ele se torna um punidor mais efetivo.

Em síntese, embora o TO seja amplamente utilizado e respaldado por evidências aplicadas, sua eficácia como procedimento punitivo depende de múltiplos fatores que vão além da simples suspensão do reforço. A identificação dos parâmetros centrais do TO é um passo importante para compreender seus efeitos sobre o comportamento. Dado o papel potencialmente modulador dos parâmetros “duração” e “taxa de apresentação” do TO, as seções seguintes discutirão em maior profundidade como variações nessas variáveis afetam o comportamento em contextos experimentais e aplicados.

Efeitos da Duração do TO na Redução da Resposta

Na literatura sobre controle aversivo, estudos com estímulos como choques elétricos indicam uma relação linear entre a magnitude do estímulo e a supressão da resposta: quanto maior a intensidade do estímulo, maior a redução da resposta punida (Azrin & Holz, 1966). Alguns estudos indicam que esse padrão também se aplica à duração do TO, sugerindo que TOs mais longos seriam mais eficazes em suprimir respostas indesejadas (Ferster & Appel, 1961; Kaufman & Baron, 1968; Flory & Lickfett, 1974). Contudo, os resultados encontrados até hoje apresentam inconsistências, especialmente quando diferentes procedimentos experimentais são considerados.

Por exemplo, Ferster e Appel (1961) investigaram o efeito da duração do TO em pombos submetidos a uma tarefa de *matching-to-sample*. No primeiro experimento, as respostas corretas em uma chave produziam reforço em um esquema de razão fixa 1 (FR1), enquanto as respostas incorretas produziam TOs com durações variando de 0,5 a 600 segundos, também em FR1. Os autores observaram que, no geral, o TO teve pouco efeito sobre a precisão das respostas, com desempenhos semelhantes aos observados em condições sem TO. Apenas o valor mais extremo de 600 s apresentou algum efeito. No segundo experimento, o esquema de reforçamento foi alterado para um intervalo variável de 3 minutos (VI 3 min), mantendo-se a apresentação do TO após respostas incorretas em FR1, com durações variando entre 1 e 120 segundos. Nesse caso, TOs entre 10 s e 60 s aumentaram a precisão das respostas, enquanto os valores extremos de 1 s e 120 s interferiram negativamente, com o TO de 120 s reduzindo o desempenho geral, afetando tanto respostas corretas quanto incorretas.

Já no estudo de Kaufman e Baron (1968), três ratos foram treinados a emitir respostas de pressão à barra na presença de um estímulo discriminativo sonoro (ruído). As respostas na presença do ruído eram reforçadas em FR1, enquanto respostas emitidas

na presença de um estímulo delta (um tom de 500 cps) resultavam em TOs com diferentes durações (10 s, 60 s, 90 s ou 120 s), também apresentados em FR1. Os resultados mostraram que a supressão das respostas na presença do estímulo delta aumentava com a duração do TO, sugerindo uma relação crescente entre a duração do TO e sua eficácia punitiva.

Diferentemente do que foi encontrado por Kaufman e Baron (1968), no estudo de Nader e Morgan (2001) não foi encontrada essa relação direta. Neste estudo, macacos rhesus foram treinados a se autoadministrar cocaína sob um esquema de reforço de intervalo fixo de 5 minutos (FI 5 min). Após estabilização, o procedimento foi alterado para um esquema múltiplo com quatro componentes. Em dois desses componentes, foram aplicados TOs de 10 s, 30 s e 60 s produzidos a partir de um esquema de VI 30 s, sobrepostos ao esquema de reforço (FI 5 min). Embora os TOs tenham reduzido significativamente as respostas de autoadministração (mais de 65%), não foi observada nenhuma diferença na redução em função da duração dos TOs, sugerindo que a eficácia do TO não aumentou com sua duração.

Estudos sobre o comportamento de esquiva também avaliaram as durações do TO na produção de respostas de esquiva, que seriam uma outra forma de avaliar a aversividade do estímulo para além das contingências de punição. Nessa configuração, o experimento IV de Thomas (1965), utilizando pombos, verificou como a duração de TOs de 30 s, 2 min, 5 min e 15 min afetava as respostas de esquiva, e os dados obtidos mostraram que o maior número de respostas de esquiva ocorreu durante o TO de 5 min, enquanto a duração mais longa prejudicou a taxa de resposta geral em um dos sujeitos analisados. Já Ono et al. (2021) compararam o procedimento de TO com o procedimento de custo de resposta em humanos e verificaram que as respostas de esquiva aumentavam à medida que a perda de pontos aumentava, mas diminuía à

medida que a duração do TO aumentava. Assim, o TO de 20 segundos foi o que obteve mais respostas de esquiva e o de 100 segundos foi o que obteve menos respostas.

Outros estudos aplicados também avaliaram os efeitos da duração do TO, como White et al. (1972) e McGuffin (1991), em crianças com comportamentos desafiadores em instituições. Ambos utilizaram diferentes durações de TOs, observando que durações curtas (1 min) tendem a ser menos efetivas na supressão do comportamento agressivo ou desafiador, enquanto durações intermediárias e mais longas (5-20 min) apresentaram maior eficácia. No geral, não foram observadas diferenças significativas entre as durações mais longas, sugerindo um possível teto no efeito da duração.

Dentre os poucos estudos que testaram a duração do TO em esquemas concorrentes, no experimento 2 de Dunn (1990), pombos foram expostos a esquemas concorrentes de reforços VI 45 s e VI 90 s nas chaves direita e esquerda, respectivamente. O TO (VI 60 s) com durações variadas (0, 10 s e 40 s) foi aplicado igualmente nas opções, sendo apresentado aleatoriamente para cada sujeito. Os resultados mostraram que a alocação da resposta diminuiu para a opção VI 45 à medida que a duração do TO aumentou, contudo, a taxa absoluta de respostas e o tempo alocado em cada chave não variaram sistematicamente em função da duração do TO, o que contrasta com o padrão observado em manipulações com choque elétrico (e.g. de Villiers, 1980). Esses achados sugerem que a duração do TO não é análoga à intensidade do choque, pelo menos em termos de produzir supressão da taxa de resposta ou de sua distribuição.

Em resumo, os dados existentes mostram que TO de curta duração (1 s ou 1 min) são geralmente ineficazes na redução do comportamento-alvo, enquanto durações moderadas (10 a 15 min) parecem produzir maior supressão. No entanto, não há consenso sobre os efeitos incrementais de TOs mais longos, com vários estudos

mostrando que aumentos além de certo ponto não resultam em maior supressão. Além disso, os efeitos da duração do TO variam conforme o tipo de procedimento experimental, o tipo de comportamento-alvo e as contingências associadas.

Efeitos da Taxa de Apresentação do TO na Redução da Resposta

Azrin e Holz (1966) relataram que um estímulo aversivo aplicado de forma constante apresenta maior eficácia na supressão do comportamento do que quando aplicado intermitentemente. Essa relação entre a constância da punição e o grau de supressão do comportamento tem sido amplamente investigada no âmbito dos esquemas de TO, especialmente em contextos de escolha com contingências concorrentes.

Thomas (1968), por exemplo, investigou os efeitos do TO sobre a taxa relativa da resposta de três pombos submetidos a um esquema concorrente. Neste experimento, a resposta de bicar era reforçada com um VI 3 min em ambas as alternativas, cada uma sinalizada por uma luz distinta. Um COD de 3 segundos foi inserido para controlar as transições entre alternativas. Após estabilização das respostas, foi aplicado um TO de 30 segundos contingente a uma das alternativas, utilizando diferentes esquemas FR (FR 50, FR 20, FR 10 e FR 2). Posteriormente, os sujeitos retornaram à linha de base e foram novamente expostos aos mesmos esquemas de razão de TO, mas com duração ampliada para 2 minutos. O estudo demonstrou que o TO funcionou como um punidor efetivo, reduzindo a taxa relativa da resposta na opção com o TO. Além disso, observou-se que a supressão da resposta punida estava diretamente relacionada ao esquema FR: taxas mais altas de punição, como no FR 2, produziram maior supressão e, concomitantemente, um aumento na resposta da alternativa impune. Por outro lado, a duração do TO (30 s ou 2 min) não apresentou efeitos sistemáticos na supressão da resposta, indicando que a frequência de aplicação do TO (taxa) foi a variável determinante na supressão da resposta (Thomas, 1968). Resultados semelhantes aos de

Thomas (1968) foram obtidos em outras pesquisas que manipularam tanto esquemas de razão fixa e variável (Zimmerman & Ferster, 1963; Clark et al., 1973; Jackson & Calhoun, 1977) quanto esquemas de intervalo fixo (Lerman et al., 1997; Ishii & Sakagami, 2007).

No estudo de Ishii e Sakagami (2007), quatro pombos foram expostos a um procedimento concorrente de reforço em FI 30 s e um RI 30 s (independentes), no qual respostas em cada chave colorida (verde para FI e vermelha para RI, ou vice-versa) resultaram na apresentação do alimento. Na linha de base, todos os 65 ciclos do FI programados por sessão resultavam em reforço, enquanto, na condição de omissão, 16 dos 65 ciclos do FI foram substituídos por um período de timeout de 2,5 s, durante o qual as chaves permaneciam iluminadas, mas nenhuma consequência era fornecida. Os resultados mostraram que a taxa geral de respostas no FI (com TO) foi reduzida em comparação à linha de base para três sujeitos e apresentou inicialmente um aumento, seguido de queda, em um deles. No entanto, a análise local indicou que, nos ciclos subsequentes a um TO, a taxa de respostas no FI era maior do que após um reforço, especialmente nas primeiras sessões da condição de omissão. A alocação da resposta também diminuiu na opção com TO em comparação com a linha de base.

No entanto, a literatura apresenta evidências ambíguas quanto à efetividade do TO quando manipulado em esquemas de intervalo variável. No Experimento 1 de Lattal (1984), três pombos foram expostos a um esquema de reforço VI 50 s na linha de base; posteriormente, foram sobrepostas duas condições de blecaute: uma delas, um blecaute de 20 s, era apresentado em VI 50 s, durante o qual o esquema de reforço permanecia suspenso (similar ao TO). Na outra condição, o blecaute de 20 s ocorria independentemente da resposta (VT 50 s), também acompanhado de períodos sem reforço. Os resultados mostraram que tanto os blecautes dependentes da resposta quanto

os independentes aumentaram as taxas de resposta durante o esquema de reforço, em comparação com a linha de base, caracterizando um contraste comportamental positivo, sem diferenças marcantes entre os dois tipos de blecautes.

Estudos mais recentes, como o de Hall (2015), investigaram diretamente as diferenças entre esquemas de razão e de intervalo aplicados a TO contingentes a respostas em esquemas múltiplos de VI 45 s. No Experimento 2A, Hall aplicou TO de 20 segundos em vários esquemas VI e VR, observando que o TO em VR 2 foi o mais eficaz na supressão da resposta, enquanto esquemas de VIs mais frequentes (VI 2,5 s e VI 5 s) apresentaram redução intermediária. Esquemas de TO menos frequentes ou não contingentes resultaram em aumento ou manutenção das taxas de resposta, possivelmente devido a pausas prolongadas após o TO contingente (Hall, 2015).

Ainda nessa linha, Dunn (Experimento 1, 1990) realizou um experimento com pombos submetidos a esquemas concorrentes de reforço VI 45 s - VI 90 s, com TO aplicado igualmente nas opções com duração de 20 segundos e em diferentes taxas de VI (45 s, 90 s e 180 s). Observou-se que, quando o TO estava em VI 45 s, houve uma diminuição da alocação de respostas na alternativa associada ao reforço VI 45 s, em comparação com as outras taxas de TO, embora a taxa de respostas absoluta tenha se mantido constante.

Adicionalmente, o Experimento 1 de Carlson (1972) investigou o TO em esquemas de cadeias concorrentes com macacos rhesus. Cada uma das duas barras disponíveis estava associada a uma cadeia de dois elos: o primeiro operava sob esquemas VI 30 s independentes, e o segundo, após a resposta, sob VI 15 s, com reforço apenas na cadeia escolhida. Após 10 sessões de linha de base, foi introduzido um TO de 15 segundos ao final do primeiro elo em uma das alternativas (a preferida). Posteriormente, para equilibrar as taxas de reforço entre as opções, o VI do segundo elo

da alternativa sem TO foi aumentado para 30 s. Todos os sujeitos apresentaram supressão da resposta no elo inicial da opção punida e aumento da resposta na alternativa, indicando que o efeito do TO não se explicava apenas por diferenças nas taxas de reforçamento.

Assim, a taxa de aplicação do TO emerge como variável crítica para sua eficácia como estímulo punitivo. De modo geral, quanto mais frequente a aplicação do TO, maior a supressão do comportamento-alvo. Essa relação parece válida independentemente do esquema utilizado, seja razão ou intervalo. Contudo, a eficácia dos esquemas de intervalo variável é menos consistente, com estudos apontando tanto para a ausência quanto para a supressão intermediária da resposta.

Embora estudos clássicos sobre punição tenham enfatizado variáveis como intensidade, frequência, imediatividade e sinalização do estímulo aversivo (Azrin & Holz, 1966), ainda há pouca investigação sistemática sobre como essas variáveis se aplicam ao TO. A maioria dos estudos aplicados trata o procedimento de forma pragmática, com ênfase em sua implementação, mas sem clareza sobre os parâmetros que o tornam efetivamente punitivo. Além disso, seu funcionamento permanece pouco compreendido em contextos que envolvem múltiplas fontes de reforço.

Historicamente, a maior parte dos estudos sobre TO foi conduzida em esquemas simples ou múltiplos, nos quais uma única classe de resposta é submetida a diferentes contingências em momentos distintos. Apenas alguns poucos estudos avaliaram o TO em esquemas concorrentes, que permitem a avaliação simultânea de duas ou mais opções de resposta (e.g., Thomas, 1968; Carlson, 1972; Dunn, 1990; Ishii & Sakagami, 2007).

TO em Situações de Escolha Concorrente

Embora a maior parte da pesquisa sobre TO tenha sido conduzida em esquemas simples ou múltiplos, sua aplicação em contextos de escolha concorrente oferece uma oportunidade única de investigar como contingências aversivas influenciam a alocação de respostas entre alternativas simultâneas. Em tais situações, o TO não apenas suprime a resposta alvo, mas também altera o valor relativo das alternativas disponíveis, desviando a escolha para a opção com menor custo aversivo (Carlson, 1972).

A formulação da Lei da Igualação por Herrnstein (1961) abriu espaço para interpretações quantitativas da escolha, que são particularmente úteis para entender como as variáveis de reforço podem influenciar a alocação de respostas.

Especificamente sobre o TO, pode possibilitar o entendimento sobre como a adição de um período para a entrega do estímulo reforçador pode afetar o valor relativo da alternativa, desviando a escolha para opções associadas a maior densidade de reforço.

Nesse contexto, diversos modelos teóricos têm sido propostos para explicar o comportamento de escolha. Além da Lei da Igualação (Herrnstein, 1961), surgiram outras propostas, como a Lei Generalizada da Igualação (GLM, em inglês Generalized Matching Law; Baum, 1974); Teoria da Redução do Atraso (DRT, em inglês Delay Reduction Theory; Fantino, 1969); Modelo Hiperbólico do Valor Adicionado (HVA, em inglês Hyperbolic Valued-Added Model; Mazur, 1987); Teoria da Informação (Shahan & Cunningham, 2015); Sinalização para boas notícias (SiGN, em inglês Signal for Good News; McDevitt et al., 2016). Esses modelos, embora originalmente desenvolvidos para descrever efeitos do reforço positivo, podem fornecer ferramentas importantes para compreender como atrasos, sinais e variações no valor do reforço influenciam decisões, e podem contribuir para a análise do TO enquanto estímulo que altera o tempo de acesso ao reforço.

Estudos experimentais sobre TO em esquemas concorrentes mostram padrões consistentes e outros menos claros. Em termos de duração do TO, resultados sugerem que quanto maior a duração do TO, menor vai ser a preferência pela alternativa punida, contudo, não foi observada uma redução na taxa de resposta absoluta em função da duração do TO (Dunn, 1990). Já a taxa de aplicação do TO sugere que, dependendo do esquema utilizado, ela pode reduzir a taxa de respostas da opção com TO (Carlson, 1972; Ishii & Sakagami, 2007) ou afetar a alocação relativa de respostas, enquanto a taxa absoluta de resposta pode permanecer estável (Dunn, 1990).

Esses achados reforçam a ideia de que o TO pode modular o valor relativo das alternativas sem necessariamente atuar como um estímulo punidor no sentido clássico de reduzir a frequência da resposta-alvo. No entanto, os fatores que determinam se o efeito do TO se manifesta apenas como mudança na alocação de respostas ou também como redução da frequência ainda não são totalmente compreendidos. Diante disso, o presente estudo buscou avaliar a aversividade do TO em um contexto de escolha, manipulando variáveis diretamente relacionadas ao procedimento. Para tanto, foram conduzidos quatro experimentos com os objetivos de investigar: (a) o efeito da duração do TO (Experimentos 1, 2 e 3) e (b) o efeito da taxa de apresentação do TO (Experimento 4).

Experimento 1:

Efeito da Duração do TO em Um Esquema Concorrente Simples

Embora alguns estudos tenham apontado que durações mais longas de TO podem aumentar sua eficácia punitiva (Ferster & Appel, 1961; Kaufman & Baron, 1968; White et al., 1972), os resultados são inconsistentes quando se consideram diferentes procedimentos experimentais, espécies e comportamentos-alvo (Nader & Morgan, 2001; Ono et al., 2021). Alguns achados indicam efeitos não lineares da

duração, sugerindo que TOs muito curtos tendem a ser ineficazes, enquanto durações moderadas (entre 10 segundos e 15 minutos) promovem maior supressão da resposta (Flory & Lickfett, 1974; Thomas, 1965), mas aumentos adicionais além desse intervalo não resultam necessariamente em maior eficácia, o que indica um possível teto (McGuffin, 1991).

Além disso, há evidências de que os efeitos da duração do TO variam conforme o tipo de procedimento utilizado (e.g., esquemas simples vs. concorrentes). Dentre os poucos estudos que investigaram diretamente a variável duração do TO em esquemas concorrentes, destaca-se o experimento 2 de Dunn (1990), cujos resultados mostraram uma relação inversa entre a duração do TO e a alocação relativa de respostas, mas não entre a duração e a taxa absoluta de respostas. Esses achados sugerem que a duração do TO pode afetar preferências em situações de escolha, ainda que seus efeitos sobre a supressão global da resposta não sigam o mesmo padrão observado em estudos com estímulos mais intensos, como choques elétricos (de Villiers, 1980).

Diante desse cenário, o presente estudo teve como objetivo avaliar o efeito da duração do TO em uma situação de escolha sob esquemas concorrentes simples, oferecendo uma abordagem sistemática para analisar como esse parâmetro modula a alocação de respostas entre alternativas de reforçamento.

Método

Sujeitos

Foram utilizados 9 ratos albinos machos (*Rattus norvegicus*), linhagem Wistar, experimentalmente ingênuos, provenientes do Instituto Evandro Chagas, e mantidos no biotério do Núcleo de Teoria e Pesquisa do Comportamento (NTPC) da Universidade Federal do Pará (UFPA). Os sujeitos estavam com aproximadamente 120 dias de vida no início do experimento e ficaram alojados individualmente em gaiolas-viveiro,

fornadas com maravalha, em condições de ventilação e iluminação controladas, com ciclo claro-escuro de 12/12 horas e fornecimento *ad libitum* de água. Os animais foram submetidos à privação de alimento, com alimentação após a sessão experimental, de forma a mantê-los a 80% do seu peso livre. O presente estudo foi aprovado pelo Comitê de Ética no Uso de Animais (CEUA) da UFPA, e protocolado sob o nº 4963241023.

Ambiente Experimental

O experimento foi realizado na Sala 4 do Laboratório de Processos Comportamentais Complexos (LPCC), no prédio do Núcleo de Teoria e Pesquisa do Comportamento 1 (NTPC 1). A sala, de formato retangular, contava com iluminação artificial de duas lâmpadas fluorescentes, ar-condicionado, lavatório e janelas vedadas com blackout. Duas mesas retangulares foram posicionadas lado a lado, e, sobre elas, as caixas de condicionamento também estavam dispostas lado a lado, juntamente com o computador de controle.

Materiais e Equipamentos

Foram utilizadas três caixas de condicionamento operante MED Associates (ENV-0800-VP), modelo standard. As caixas tinham 30 cm de comprimento, 24 cm de largura e 21 cm de altura. No interior das caixas havia duas barras de resposta em um painel de alumínio, separadas por 13 cm de distância e equidistantes de uma abertura central onde está localizado o bebedouro. Acima de ambas as barras havia uma luz. No painel oposto a barras, também de alumínio, havia, na parte superior, a luz central da caixa de condicionamento. A parede da frente, o teto e a parede traseira eram de acrílico, e toda a câmara experimental ficou alojada em um compartimento de som e luz atenuante.

Como estímulo reforçador, era disponibilizado 0,1 ml de uma solução de sacarose com concentração de 10% de sacarose/água por 3 segundos, utilizando um

dispensador de líquidos (*liquid dipper* ENV-202M-S). Todos os eventos da sessão experimental foram programados e registrados automaticamente por meio do software MED-PC IV, em um computador de controle.

Procedimento Geral

As sessões experimentais ocorreram diariamente, no período da manhã e da tarde, aproximadamente no mesmo horário para cada sujeito. Em todas as fases do experimento, os intervalos variáveis dos esquemas de reforçamento e do TO foram programados de acordo com a distribuição de Fleshler e Hoffman (1962), com dez intervalos, apresentados em ordem aleatória e sem repetição dentro de cada ciclo.

Durante o período de TO, todas as luzes foram apagadas e nenhuma barra disponibilizou reforçadores, mas todas as respostas de pressão às barras ocorridas durante esse período foram registradas. Após o TO, todas as luzes foram ligadas e os esquemas de reforço para ambas as alavancas foram mantidos. O tempo de time-in e de TO foi contabilizado separadamente.

Antes do procedimento experimental, os sujeitos passaram por sessões de treino ao bebedouro, de modelagem e de fortalecimento da resposta de pressão à barra. O procedimento experimental teve três fases: duas de reforço e uma de reforço com TO sobreposto ao reforçamento. Nos próximos tópicos, será descrito em detalhe cada um dos procedimentos implementados durante o experimento.

Treinamento ao Bebedouro e Modelagem da resposta

Com o objetivo de emparelhar o som da ativação do bebedouro e o próprio bebedouro, com gotas de sacarose, e de modelar a resposta de pressão à barra, foi realizado o procedimento em que gotas de sacarose foram apresentadas independentemente da resposta, em um esquema de tempo variável de 30 segundos (VT 30 s). Adicionalmente, independentemente da finalização do intervalo, respostas de

pressão na barra ativa, que foi aleatorizada entre os sujeitos, foram imediatamente seguidas por reforço (FR 1). Foram realizadas duas sessões com duração de 30 minutos. Durante as sessões, a luz da caixa e das barras ficou ligada. Quando um sujeito não apresentava a resposta-alvo ao final da segunda sessão, era realizada uma sessão de modelagem manual da resposta.

Treino da Resposta

Nessa etapa, com o objetivo de fortalecer a resposta de pressão à barra, foram realizadas duas sessões de 30 minutos, uma para cada barra, nas quais a resposta de pressão à barra foi reforçada em um esquema FR 1. A sequência de treinos entre as barras foi alternada aleatoriamente entre os sujeitos.

Fase 1 (Linha de Base)

Durante essa fase, foi utilizado um esquema concorrente simples, no qual os sujeitos produziram reforço em ambas as barras em um esquema VI 20 s, independentemente de cada barra. Um *Changing Over Delay* (COD) de 3 segundos foi programado de modo que uma resposta à barra não fosse reforçada se uma resposta à barra oposta tivesse sido emitida nos 3 s anteriores. As sessões tiveram duração de 30 minutos cada e foram realizadas pelo menos 25 sessões, devido à identificação de um problema na programação do COD das sessões iniciais do experimento, e até os sujeitos atingirem a estabilidade na distribuição da resposta. O critério de estabilidade adotado exigiu que, nas últimas 6 sessões, a diferença entre a proporção média de respostas nas três primeiras e nas três últimas sessões variasse em até 10% em relação à média dessas seis sessões, sem tendência crescente ou decrescente, avaliada por inspeção visual.

Fase 2 (TO)

Os sujeitos continuaram a produzir o reforço em ambas as barras em um esquema VI 20 s (esquema concorrente, independente para ambas as barras), contudo,

foi sobreposto um TO de duração variada em um esquema de VI 20 s a uma das barras (a escolha da barra foi aleatorizada entre os sujeitos) contingente à resposta de pressão à barra. A duração do TO de 5 s, 20 s e 60 s variou a cada sessão de forma semi-aleatória para cada sujeito, de modo que uma duração nunca foi apresentada dois dias seguidos, e também variou entre os sujeitos que eram testados no mesmo horário. O COD também foi aplicado às apresentações de TOs de modo que as respostas à barra punida não produziram TOs se uma resposta na barra oposta tivesse sido emitida nos 3 s anteriores. As sessões terminaram após 30 min de time-in. Cada sujeito permaneceu nessa fase por 15 sessões, de cada duração do TO.

Fase 3 (Linha de Base)

Nessa fase, os sujeitos retornaram aos procedimentos adotados na Fase 1 logo após o último dia de TO. Eles permaneceram nessa condição por no mínimo 15 sessões e até os sujeitos atingirem a estabilidade na distribuição da resposta.

Análise de Dados

As variáveis dependentes (VDs) a serem analisadas foram as taxas de resposta (resp/min) em cada opção e a proporção da resposta entre as opções (R1/R2). As taxas de reforçamento local e total e as taxas de TO também foram analisadas, assim como o tempo de *time-in* e o tempo total da sessão. Para a análise dos dados da Fase 1 e Fase 3, foram utilizadas as últimas seis sessões que indicavam estabilidade da proporção da resposta.

Para avaliar um possível efeito das durações do TO, bem como a exposição prolongada a essas durações na Fase 2, os dados foram agrupados em blocos de 5 sessões (Bloco 1: sessões 1-5, Bloco 2: sessões 6-10 e Bloco 3: sessões 11-15) e realizada uma análise estatística utilizando o Modelo Linear Generalizado Misto (GLMM, do inglês). O GLMM é uma ferramenta estatística que permite analisar

variáveis dependentes que não seguem uma distribuição normal e considera simultaneamente efeitos fixos (fatores experimentais) e aleatórios (como participantes ou contextos específicos). Essa flexibilidade é importante quando os dados frequentemente apresentam assimetrias, múltiplas fontes de variabilidade e estruturas hierárquicas ou longitudinais, como em estudos com medidas repetidas (Bono et al., 2021).

Diante das características do presente estudo, em que todos os sujeitos passaram por todas as condições experimentais, o uso de GLMMs permite respeitar a complexidade dos dados reais, evitar distorções estatísticas e obter inferências mais precisas e generalizáveis.

Assim, o estudo utilizou como variáveis fixas a “Duração do TO”, “Bloco de sessões” e “Opção da resposta”; como variável aleatória, utilizou-se “Sujeitos”. O estudo também realizou a análise das médias marginais, que permitem visualizar como a VD é afetada em cada nível dos fatores fixos, bem como os contrastes planejados, que permitem comparações específicas entre as variáveis analisadas. Com base nessa análise, o Bloco 3 foi considerado uma fase de estabilização dos efeitos, servindo de base para as análises descritivas da Fase 2.

Além disso, foi realizada uma Regressão Linear Simples para analisar os efeitos das variáveis independentes (a distribuição relativa da taxa local de reforçamento, a distribuição relativa da taxa total de reforçamento, a taxa do TO no time-in e a taxa do TO no tempo total) sobre a proporção da resposta.

Todas as análises estatísticas foram realizadas por meio do *software* de domínio público e código aberto JASP, versão 0.19.3.0.

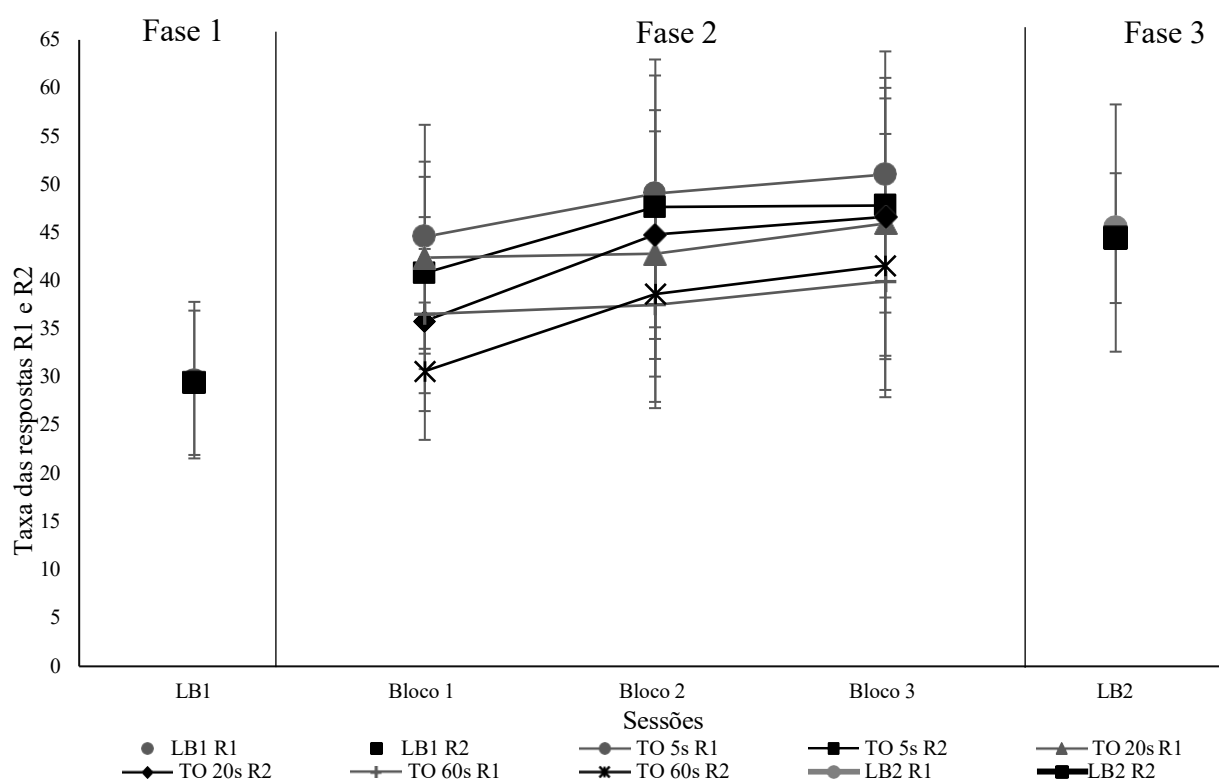
Resultados e Discussão

Na Fase 1, em média, os participantes alcançaram a estabilidade em 26 sessões (SD = 1,09). Ao final dessa fase (i.e., média das últimas seis sessões), os sujeitos apresentaram uma taxa de respostas de 30 resp/min na opção R1 (SD = 7,95) e 29 resp/min na opção R2 (SD = 7,46) e produziram, em média, 6 ref/min (R1: M = 2,81; SD = 0,17 / R2: M = 2,85; SD = 0,11) (Ver Figura 1). A proporção de respostas (Figura 2) foi, em média, de 1,02 (SD = 0,2). O tempo médio total por sessão foi de 41 min (SD = 0,53). Para detalhes dos dados do grupo e individuais, ver Figura 1 e Apêndice A.

Figura 1

Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do

Experimento 1



Nota: A figura 1 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e Fase 3 e dos blocos de sessões da Fase 2 do Experimento 1. O TO foi apresentado contingente apenas à opção R2 da Fase 2. As barras sobrepostas a cada ponto de dados representam os respectivos desvios padrão.

Efeito da Duração do TO na Taxa da Resposta

Com base nos resultados do Bloco 3 da fase 2, a maioria dos sujeitos apresentou um aumento na taxa de resposta em comparação com a fase anterior, independentemente da duração do TO. Na Fase 1, os sujeitos respondiam, em média, 30 resp/min na opção R1 e 29 resp/min na opção R2 e, ao final desta fase, apresentavam, em média, durante o TO de 5s: 51 resp/min na opção R1 (SD = 12,76) e 48 resp/min na opção R2 (SD = 11,11); TO de 20s: 46 resp/min na opção R1 (SD = 14,07) e 47 resp/min na opção R2 (SD = 14,41); TO de 60s: 40 resp/min na opção R1 (SD = 11,27) e 42 resp/min na opção R2 (SD = 13,65), considerando-se o período de time-in. Apenas o sujeito Y2 apresentou redução na taxa de resposta para R1 e R2, em relação à linha de base, quando o TO tinha duração de 20 s e 60 s (ver dados no Apêndice A). A análise das taxas de resposta com base no GLMM, revelou efeito significativo para o fator Duração do TO ($\chi^2_{(1)} = 11.145$, $p = .004$).

Apesar do procedimento de TO ser amplamente conceituado como uma técnica punitiva com base na definição de punição como uma contingência que reduz a probabilidade de uma resposta (Skinner, 1953; Pierce & Cheney, 2013), os resultados do presente experimento indicaram que a maioria dos sujeitos apresentou aumento na taxa de resposta durante a Fase 2 em todos os TOs, em comparação com a linha de base. Esse aumento, entretanto, foi sensível à duração do intervalo, sendo menor para TOs mais longos. Mesmo assim, o padrão observado sugere que, ao menos do ponto de vista da taxa de respostas, o TO não exerceu função punitiva. Resultados semelhantes foram observados por Dunn (1990), que observou que TOs aplicados de maneira contingente à resposta em esquemas concorrentes não produziram queda sistemática na taxa de resposta absoluta, o que contradiz modelos punitivos clássicos que utilizam choque

elétrico como punidor (de Villiers, 1980), sugerindo que o TO pode não produzir uma redução da probabilidade do comportamento nesses contextos.

Outro exemplo dessa dinâmica foi relatado por Ferster (1953) no Experimento III, que avaliou os efeitos de blecautes, semelhantes ao TO no sentido de remover estímulos por um tempo limitado, na taxa de resposta em uma contingência simples. No estudo, três pombos foram submetidos a um esquema de reforço com VI 60 s e, posteriormente, os blecautes (duração não especificada) foram sobrepostos ao reforço sob um esquema de VI 60 s. Os resultados mostraram que esse tipo de apagão contingente à resposta não suprimiu a taxa de resposta dos pássaros, mesmo após longos períodos de exposição, sugerindo que as interrupções no ambiente podem não exercer controle punitivo sobre o comportamento. Essa evidência é relevante para interpretar o padrão observado no presente estudo, em que o TO pode funcionar mais como uma interrupção contextual do que como punição, especialmente em contextos ricos em reforço.

Além disso, a análise dos dados revelou que, durante a apresentação do TO, todos os sujeitos diminuíram a taxa de resposta (ver Apêndice B). Essa diferença entre os períodos de TO e time-in sugere que o TO atuou como um estímulo discriminativo, sinalizando a indisponibilidade de reforço e, por consequência, promovendo a supressão transitória da resposta enquanto estava em vigor (Skinner, 1953; Sidman, 1989). Contudo, o aumento subsequente da taxa de resposta no time-in indica que o TO, em vez de funcionar como punição, produzindo uma redução duradoura na frequência da resposta, pode ter exercido outras funções comportamentais.

Uma possível explicação é que o TO gerou uma alternância entre contextos de reforço e de não reforço, promovendo uma diferenciação temporal do comportamento. Tal como discutido por Ferster (1953), o blecaute contingente à resposta, assim como o

TO, não necessariamente diminui o responder geral, mas acrescenta um “tempo morto” (Ferster, 1953) à contingência. Essa interpretação sugere que o período de blecaute atua como uma pausa discriminativa, alterando temporariamente a probabilidade de emissão da resposta. Nessa lógica, o TO pode ter servido como um evento de sinalização temporal, aumentando o controle do tempo sobre a resposta, de modo que os sujeitos aprendem a alocar suas respostas nos períodos em que o reforço está disponível (time-in), elevando a taxa nesse momento.

Resultados semelhantes foram encontrados por Lattal (1984, Experimento 1), que, ao replicar o estudo de Ferster (1953), também não observou supressão das taxas de resposta em esquemas simples. Pelo contrário, verificou-se até mesmo um aumento na taxa de respostas em comparação à linha de base, na qual não havia blecaute.

Diferentemente de Ferster (1953), entretanto, o autor interpretou esse efeito como um contraste comportamental positivo. Esse fenômeno é descrito na literatura, em esquemas múltiplos, pelo aumento da taxa de respostas em um componente estável, resultante da redução da taxa de reforçamento no componente alternativo (Boyle et al., 2018; Killeen, 2014; Teixeira, 2023). Considerando a sinalização diferencial entre os períodos de TO e de time-in, ainda que o presente experimento tenha empregado um esquema concorrente, a alternância entre essas condições pode ser entendida como análoga à mudança de componentes, o que explicaria a ocorrência de um contraste positivo, ou seja, o aumento da taxa de resposta durante o time-in, mesmo sem alteração na taxa de reforçamento dessa condição, semelhante ao relatado por Lattal.

O estudo de Ishii e Sakagami (2007), embora tenha registrado uma redução na taxa de respostas na alternativa com TO, revelou um padrão consistente com a explicação proposta por Lattal (1984). A análise detalhada mostrou que, nos ciclos

imediatamente seguintes a um TO, a taxa de respostas era maior do que após um reforço, especialmente nas primeiras sessões da condição experimental.

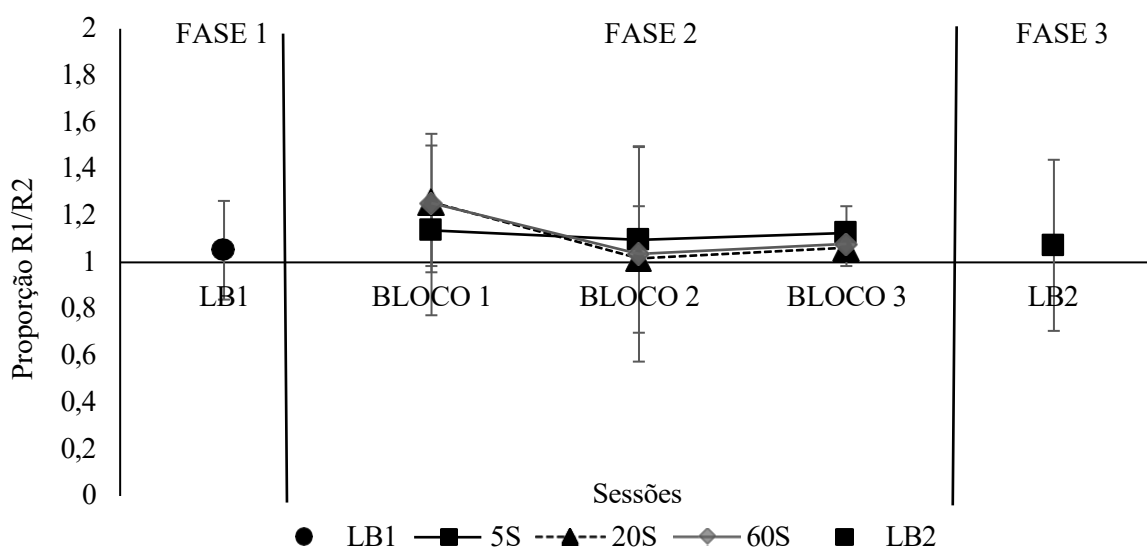
Efeito da Duração do TO ao Longo dos Blocos de Sessões

Com base na análise descritiva da Fase 2, observa-se que, ao longo da exposição ao TO, houve uma redução progressiva na média da proporção de respostas em relação à linha de base inicial (LB1). No Bloco 1, os sujeitos apresentaram maior proporção de respostas em R1 do que nos Blocos 2 e 3, o que pode sugerir um efeito de habituação ao TO.

Contudo, a análise da proporção da resposta por meio do GLMM não indicou efeito significativo do fator Blocos de sessões ($\chi^2_{(2)} = 1.68$, $p = .43$) (ver Figura 2 e Apêndice A).

Figura 2

Média geral da proporção da resposta das fases experimentais do Experimento 1.



Nota: A figura 2 apresenta a média geral da proporção da resposta 1 (R1) sobre a resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da

Fase 2 do Experimento 1. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Duração do TO na Proporção da Resposta

Na Fase 2, os sujeitos apresentaram uma proporção média de 1,12 durante o TO de 5 s (SD = 0,39); 1,06 durante o TO de 20 s (SD = 0,39) e 1,07 durante o TO de 60 s (SD = 0,46) no último bloco de sessões. Apesar dos resultados da alocação da resposta representarem um pequeno aumento da proporção em direção à opção sem TO, quando comparados com a Fase 1 (Prop = 1,02 durante a Fase 1) (ver Figura 2), o resultado da GLMM não revelou efeito significativo para o fator Duração do TO ($\chi^2_{(2)} = 0.470$, $p = 0.703$).

A análise individual mostrou que, para seis dos nove sujeitos (Y1, Y2, Y3, Y4, Y5 e Y7), a introdução do TO na Fase 2, independentemente de sua duração, alterou a distribuição das respostas em direção à opção sem TO no início da exposição. Entretanto, no bloco final, esses sujeitos apresentaram uma diminuição na proporção de respostas nessa opção, padrão semelhante ao observado na média geral.

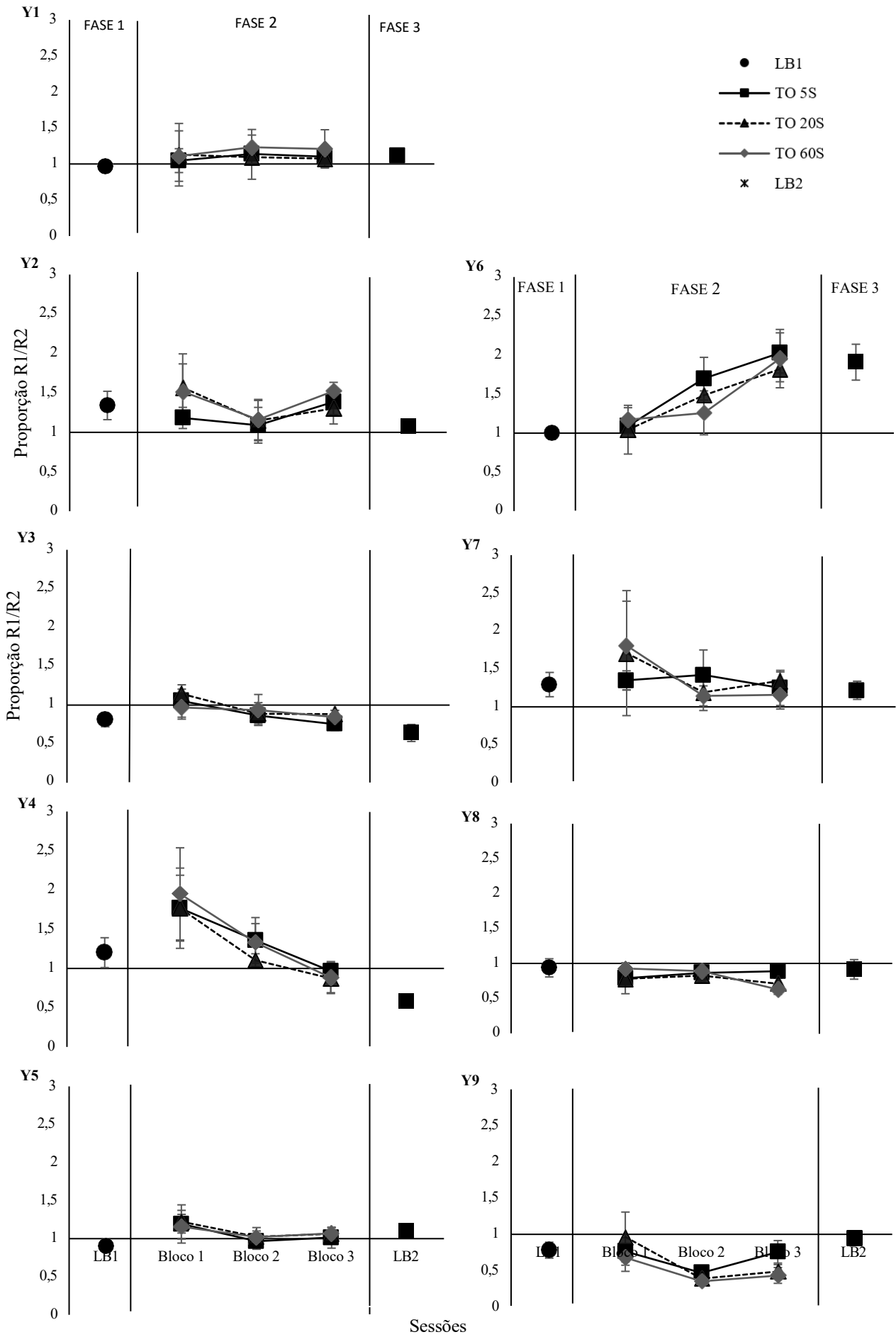
Três sujeitos exibiram padrões distintos da tendência geral: o sujeito Y6 apresentou um aumento progressivo na proporção de respostas ao longo da exposição ao TO; o sujeito Y8 não mostrou alterações expressivas, mantendo-se, no entanto, abaixo do nível observado na linha de base; e o sujeito Y9 apresentou uma redução na proporção desde o início da fase, indicando preferência consistente pela opção com TO (R2).

Ao analisar especificamente o último bloco, percebe-se que seis dos nove sujeitos (Y1, Y2, Y3, Y5, Y6 e Y9) preferiram a opção sem TO em pelo menos duas durações. Desses, cinco alocaram mais respostas na opção sem TO quando a duração foi de 5s (Y1, Y2, Y5, Y6 e Y7), cinco quando a duração foi de 20 s (Y1, Y3, Y5, Y6 e

Y7) e cinco quando a duração foi de 60 s (Y1, Y2, Y3, Y5 e Y6). Por outro lado, três sujeitos (Y4, Y8 e Y9) apresentaram uma redução na proporção de respostas, passando a preferir a opção associada ao TO em todas as durações. Além disso, a alocação de respostas dos sujeitos Y8 e Y9 foi proporcional à duração do TO, ou seja, quanto maior o TO, mais respostas foram emitidas na opção R2 (ver dados individuais na Figura 3).

Figura 3

Média individual da proporção da resposta das fases experimentais do Experimento



Nota: A figura 3 apresenta a proporção individual da resposta 1 (R1) em relação à resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da Fase 2 do Experimento 1. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Um achado relevante foi a mudança na alocação da resposta, por inspeção visual, observada em seis dos nove sujeitos, que passaram a preferir a opção não associada ao TO no último bloco da Fase 2. Esse padrão sugere que o TO influenciou o comportamento por mecanismos distintos daqueles tradicionalmente descritos em procedimentos clássicos de punição, os quais se concentram na supressão direta da resposta. De acordo com Dunn (1990), contingências de TOs em esquemas concorrentes podem não reduzir a taxa absoluta de respostas, mas ainda assim alterar a distribuição das respostas entre as alternativas. Assim, o TO pode atuar como uma contingência aversiva que afeta a escolha mesmo quando não há redução na taxa de resposta, visto que, em situações de escolha, a decisão entre alternativas envolve uma comparação relativa entre as opções, levando em consideração fatores como a quantidade de reforço, o esforço ou o custo associado ao movimento entre as opções (Baum, 1982), e a presença do TO pode funcionar como um "custo adicional" associado à opção em que ocorre.

Além disso, sob a perspectiva de outras teorias da escolha, como a DRT (Fantino, 1969), a opção sem TO pode ter sinalizado maior proximidade ao reforço, o que aumentaria seu valor relativo. Assim, mesmo na ausência de diminuição da resposta-alvo, a mudança na alocação sugere que os sujeitos estavam sensíveis aos parâmetros temporais da contingência, favorecendo respostas à alternativa com menor atraso na apresentação do estímulo reforçador.

Apesar do padrão majoritário de escolha pela alternativa sem TO, três sujeitos demonstraram uma preferência sistemática pela opção associada ao TO, sendo que, para alguns, essa preferência aumentou com o acréscimo da duração do TO. Esse resultado, embora contraintuitivo à função aversiva do TO, também pode ser compreendido à luz de teorias que consideram variáveis temporais e contextuais mais amplas. A própria DRT, que propõe que estímulos associados à maior redução do atraso até o reforço tendem a adquirir maior controle sobre o comportamento, pode ser uma base de explicação para os dados observados. O TO, apesar de ser um procedimento que impõe tempo adicional entre a apresentação dos reforçadores, também é caracterizado por ser sinalizado diferencialmente; logo, em contextos em que a entrega do reforço é variável, a sinalização do TO pode também indicar a maior proximidade do reforço e, paradoxalmente, aumentar o valor da alternativa.

Essa interpretação é compatível com o modelo SiGN (Dunn et al., 2024), derivado da DRT, que sugere que estímulos antecedentes (como os sinais do TO) adquirem valor condicional com base na distribuição do reforço, podendo manter o comportamento mesmo na ausência de reforço imediato. Assim, as observações do presente estudo podem ser análogas ao que ocorre nos procedimentos de escolha subótima, em que estímulos que precedem reforços incertos ou atrasados podem aumentar a preferência por alternativas menos vantajosas em termos de taxa de reforço (Smith & Zentall, 2016). Ainda que o TO represente uma interrupção no acesso ao reforço, sua sinalização pode ter funcionado como um marco temporal relevante em um ambiente com reforçamento distribuído de forma variável.

Além disso, Cunningham e Shahan (2019) destacam que eventos sinalizadores de mudança na taxa de reforço (e.g., a introdução do TO) podem modular a sensibilidade ao valor relativo das opções disponíveis. Assim, para esses três sujeitos, o

TO pode ter funcionado como um estímulo discriminativo de transição que sinalizava maior previsibilidade de reforço futuro. Esse efeito pode ter sido amplificado nos TOs mais longos, por oferecerem um intervalo mais claro de separação entre respostas e reforços, favorecendo o controle temporal.

Efeito da Duração do TO na Taxa de Reforço e na Duração da Sessão

As diferentes durações do TO resultaram em variações no tempo total das sessões, o que impactou as taxas de reforço e de TO calculadas com base no time-in (30 min.) e no tempo global (ver Tabela 1).

Tabela 1

Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 1.

TO	Tempo total de sessão	Taxa de reforço R1 (Time-in)	Taxa de reforço R2 (Time-in)	Taxa de reforço global R1	Taxa de reforço global R2	Taxa de TO R2 (Time-in)	Taxa de TO global R2
5s	51,61 (0,42)	3,10 (0,06)	3,08 (0,04)	1,80 (0,02)	1,79 (0,01)	3,06 (0,06)	1,78 (0,02)
20s	74,61 (1,47)	3,08 (0,08)	3,09 (0,07)	1,24 (0,03)	1,24 (0,09)	3,07 (0,12)	1,23 (0,02)
60s	134,59 (4,10)	3,06 (0,10)	3,05 (0,13)	0,68 (0,03)	0,68 (0,01)	3,02 (0,13)	0,67 (0,009)

Nota: A tabela 1 apresenta a média do tempo total de sessão (em minutos); a taxa de reforço no time-in para as opções R1 e R2, a taxa de reforço global para as opções R1 e R2 e a taxa de TO no time-in e taxa de TO global do Bloco 3 da Fase 2 do Experimento 1. Os números entre parênteses representam o desvio padrão de cada média.

A tabela 1 mostra que, mesmo com a introdução do TO e independentemente da duração em vigor, os sujeitos receberam, em média, todos os reforços programados para o período de time-in em ambas as opções. A alteração real na taxa de reforço só foi observada na taxa global, que leva em consideração o tempo total da sessão. Nessa condição, a taxa de reforço, em ambas as opções, foi reduzida em 42% durante o TO de 5 s; em 60% durante o TO de 20 s; e em 78% durante o TO de 60 s. Contudo, a

distribuição dos reforços entre as duas opções se manteve similar em todas as variações do TO. A tabela mostra ainda que os sujeitos receberam todos os TOs programados para o período de time-in, sendo que a taxa de TO durante o tempo total de sessão foi reduzida de forma similar a taxa de reforço para cada duração do TO.

Esses dados são mais um indicativo de que o TO não suprimiu a taxa de resposta a ponto de comprometer o acesso aos reforços disponíveis e, apesar das reduções na taxa global de reforço, a distribuição dos reforços entre as opções de resposta se manteve estável entre as condições. Paralelamente, os dados também demonstraram que os sujeitos passaram por todos os TOs programados. Esses achados indicam que o principal efeito do TO, neste caso, foi a diminuição da densidade temporal de reforçamento, sem alterar de forma consistente a distribuição dos reforços entre as opções.

A análise das regressões lineares demonstrou que, sob diferentes durações do TO, a distribuição relativa da taxa local de reforçamento ($F_{(1, 25)} = 18.566$, $p < .001$, R^2 ajustado = .403, AIC= 17.500, BIC= 21.387) e a global ($F_{(1, 25)} = 18.461$, $p < .001$, R^2 ajustado = .402, AIC= 17.565, BIC= 21.453) foram preditoras consistentes e significativas da alocação da resposta-alvo, além de apresentarem valores de AIC e BIC menores que os demais modelos. Por outro lado, nem a taxa de TO no time-in ($F_{(1, 25)} = 1.416$, $p = 0.245$, R^2 ajustado = 0.016, AIC= 31.007, BIC= 34.895), nem a taxa total de TO ($F_{(1, 25)} = 0.049$, $p = 0.826$, R^2 ajustado = -0.038, AIC= 32.442, BIC= 36.330) explicaram a variabilidade da proporção, sugerindo que, neste experimento, os efeitos do TO sobre a escolha foram mais dependentes da distribuição do reforço no time-in e do tempo total da sessão.

As análises são consistentes com os princípios da Lei da Igualação, proposta por Herrnstein (1961), que estabelece que a razão de respostas entre duas alternativas em

esquemas concorrentes tende a igualar a razão de reforços obtidos nessas mesmas alternativas. Apesar dessa lei não englobar diretamente os efeitos de eventos como o TO, os dados indicam que o comportamento foi mais sensível à distribuição do reforço disponível do que à presença de um evento potencialmente aversivo como o TO. Assim, os resultados sugerem que o TO afetou a alocação de respostas apenas na medida em que modificou as contingências de reforço, e não por sua apresentação isolada.

Por fim, essa análise reforça os achados sobre as taxas de reforço: mesmo com aumento da resposta em todos os TOs, a análise estatística mostrou que a magnitude do aumento foi modulada pela duração do TO. TOs mais longos reduziram mais a taxa global de reforço, impactando, conseqüentemente, a taxa de resposta, enquanto TOs mais curtos tiveram efeito menor, permitindo um aumento maior na taxa de resposta.

Na Fase 3, os sujeitos passaram, em média, por 17 sessões ($SD = 1,39$). Com a descontinuação do TO, as taxas de resposta se mantiveram mais altas que a média da Fase 1 e da Fase 2 (TO de 60 s) nas duas opções de resposta e mais baixas do que a média da Fase 2 nos TOs de 5 s e 20 s, sendo emitidas na Fase 3, em média, 45 resp/min na opção R1 ($SD = 12,88$) e 44 resp/min na opção R2 ($SD = 6,78$). Apesar do aumento na taxa das respostas durante essa fase, a proporção de respostas foi, em média, de 1,03 ($SD = 0,38$). Analisando os dados individuais, observou-se que quatro sujeitos (Y1, Y5, Y6 e Y9) aumentaram a alocação da resposta na opção R1, e os outros 5 sujeitos (Y2, Y3, Y4, Y7 e Y8) aumentaram a alocação da resposta na opção R2, em relação a Fase 1. Em comparação com a Fase 2, quatro sujeitos (Y2, Y3, Y4 e Y5) aumentaram a alocação da resposta na opção R2 para todas as durações do TO e dois sujeitos apresentaram esse mesmo padrão para duas durações do TO (5 s = Y6, Y7; 20 s = Y7; 60 s = Y6). Os sujeitos Y8 e Y9 aumentaram a distribuição da resposta na opção R1

para todas as durações do TO e o sujeito A1 apresentou esse mesmo padrão em relação às durações de 5s e 20s (ver Apêndice A).

A taxa de reforço total foi de 6 ref/min (R1: $M = 2,98$; $SD = 0,08$; / R2: $M = 2,95$; $SD = 0,07$) e o tempo total de sessão foi de aproximadamente 42 min ($SD = 0,32$).

A análise das taxas de resposta da Fase 3, em que o TO foi retirado, sugere um possível efeito residual das contingências da Fase 2, visto que as taxas se mantiveram mais altas do que na Fase 1 (também sem TO). Esses dados podem indicar que a experiência prévia com o TO pode ter produzido alterações duradouras no padrão geral de atividade. Já a proporção de respostas mostrou-se menor do que na Fase 2, sugerindo que a retirada do TO reduziu o contraste entre as opções e favoreceu uma distribuição mais equilibrada das escolhas, ainda que houve uma leve manutenção da preferência adquirida anteriormente, visto que a proporção se manteve um pouco mais alta do que na Fase 1.

Em síntese, os resultados deste experimento indicam que, embora o TO não tenha promovido uma supressão direta da resposta alvo, ele influenciou a alocação da resposta, com parte dos sujeitos preferindo a alternativa sem TO e parte preferindo a alternativa com TO. Observou-se ainda que, mesmo após a retirada do TO, as taxas de resposta permaneceram elevadas, sugerindo um efeito residual das contingências anteriores. No geral, as análises sugerem que o TO pode atuar mais como modulador do padrão de escolha do que como um supressor real do comportamento ao qual foi contingente.

Experimento 2:

Efeito da Duração e Intervalo do TO em Um Esquema Concorrente Simples

Como observado no Experimento 1 do presente trabalho, não foram identificadas reduções consistentes na taxa de resposta, nem uma relação clara entre o

aumento da duração do TO e a alocação de respostas. Isso corrobora outros estudos de esquemas concorrentes, como o de Dunn (1990). Esses achados sugerem que a duração, isoladamente, pode não ser suficiente para afetar as preferências de forma robusta.

Considerando isso, optou-se por investigar uma nova variável: a frequência com que o TO é apresentado. Estudos como o de Thomas (1968) demonstraram que taxas mais elevadas de TO podem gerar maior supressão do comportamento em esquemas concorrentes simples. Diante disso, no segundo experimento, investigou-se o efeito da duração do TO na alocação das respostas ao aumentar a taxa de ocorrência do TO, por meio da alteração do esquema de apresentação do TO de VI 20 s para VI 5 s.

Método

Sujeitos

Foram utilizados 4 ratos albinos machos (*Rattus norvegicus*), linhagem Wistar, que participaram do Experimento 1 (os sujeitos foram escolhidos aleatoriamente). As condições de manutenção foram idênticas às do Experimento 1.

Ambiente Experimental

O experimento foi realizado no mesmo ambiente do Experimento 1.

Materiais e Equipamentos

Os materiais e equipamentos utilizados foram os mesmos descritos no Experimento 1. Contudo, nesse experimento houve o acréscimo de uma caixa de condicionamento operante MED Associates (ENV-0800-VP), modelo standard, com as mesmas dimensões citadas anteriormente, totalizando 4 caixas.

Procedimento Geral

Os procedimentos de treinamento ao bebedouro, modelagem e treino da resposta já haviam sido realizados no Experimento 1. Este experimento contou com apenas uma fase experimental, utilizando a Fase 3 do experimento anterior como linha de base.

Fase 2 (TO)

O procedimento foi semelhante ao Experimento 1, contudo, o TO foi aplicado em um VI 5 s e apenas com duas durações de 5 s e 20 s. O critério de encerramento dessa fase foi de 12 sessões, em cada uma das durações do TO.

Análise de Dados

Os dados foram submetidos à mesma análise do Experimento 1, com a diferença de que os blocos de sessões foram: Bloco 1= sessões 1- 4; Bloco 2= sessões 5 – 8; Bloco 3= sessões 9 – 12.

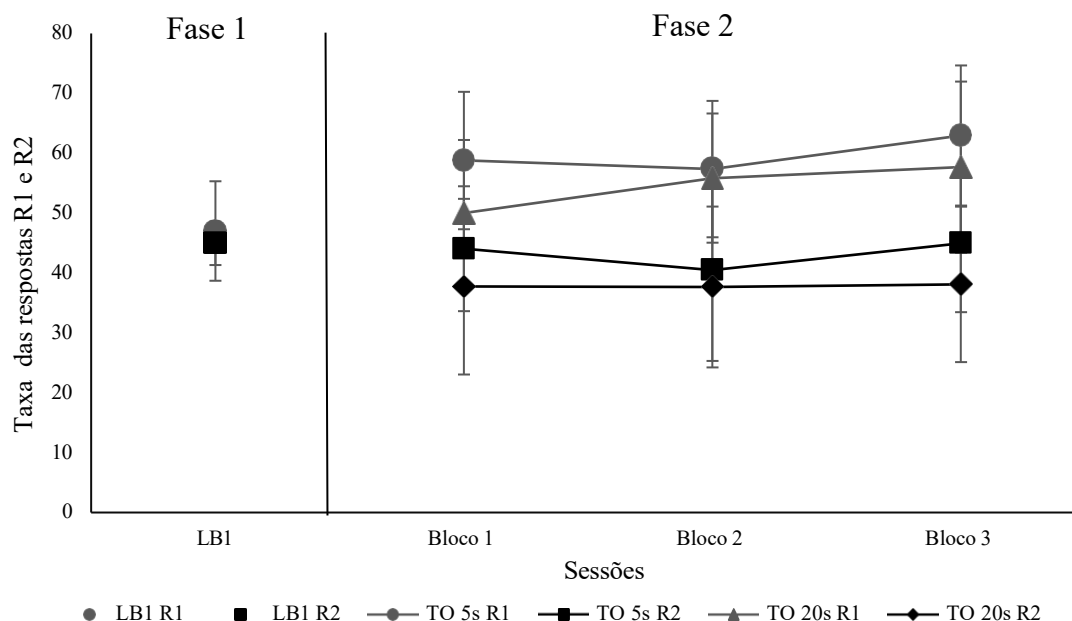
Resultados e Discussão

Os dados referentes à Fase 1 deste experimento correspondem aos da Fase 3 do Experimento 1 para os sujeitos Y1, Y7, Y8 e Y9.

Os participantes passaram, em média, por 18 sessões (SD = 1,09). Ao final da linha de base (Fase 1) (i.e., média das últimas seis sessões), os sujeitos emitiram 47 resp/min na opção R1 (SD = 7,95) e 45 resp/min na opção R2 (SD = 7,46), produzindo, em média, um total de 6 ref/min (R1: M = 2,98; SD = 0,08; R2: M = 2,91; SD = 0,10). A proporção de respostas (Figura 4) foi, em média, 1,05 (SD = 0,14). O tempo médio total da sessão foi de 42 min (SD = 0,44). Para detalhes dos dados individuais, ver a Figura 4 e o apêndice C.

Figura 4

Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 2



Nota: A figura 4 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e dos blocos de sessões da Fase 2 do Experimento 2. O TO foi apresentado contingente apenas à opção R2 da Fase 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Duração do TO na Taxa da Resposta

Nessa fase, durante o Bloco 3, todos os sujeitos apresentaram um incremento na taxa de respostas na opção R1 (sem TO) em comparação com a fase anterior, independentemente da duração do TO. Quando a duração do TO foi de 5 s, a média geral da taxa de respostas na opção R1 aumentou de 47 resp/min na Fase 1 para 63 resp/min (SD = 11,70) na Fase 2 e para 58 resp/min (SD = 14,27) quando o TO foi de 20 s.

Por outro lado, a média geral da taxa de respostas na opção R2 (com TO) permaneceu em 45 resp/min (SD = 11,50) no TO de 5 s e caiu para 38 resp/min (SD = 12,97) no TO de 20 s. Os sujeitos apresentaram variações individuais nessa opção. Durante o TO de 5 s, os sujeitos Y7 e Y9 reduziram a taxa de respostas na opção R2, enquanto os sujeitos Y1 e Y8 apresentaram um aumento nessa mesma opção em

comparação com a Fase 1. Já no TO de 20s, os sujeitos Y1, Y7 e Y9 reduziram a taxa de respostas na opção R2, enquanto o sujeito Y8 apresentou um incremento em relação à linha de base (ver Figura 4). A análise das taxas de resposta com base no GLMM revelou efeito significativo para o fator Duração do TO ($\chi^2_{(1)} = 4.616, p = 0.032$).

Em comparação com o Experimento 1, TOs mais curtos (5 s) aumentaram mais a taxa de respostas do que TOs intermediários (20 s) e longos (60 s). No Experimento 2, em que a frequência do TO foi maior, o aumento médio da taxa de respostas na alternativa punida foi menos evidente, o que reflete maior variabilidade entre os sujeitos. Esses padrões indicam que a duração e a frequência do TO modulam a distribuição da resposta em esquemas concorrentes, com TOs curtos e menos frequentes favorecendo um aumento mais pronunciado da taxa de respostas, enquanto TOs longos ou mais frequentes reduzem esse efeito.

No Experimento 2, mesmo sem ser observada supressão generalizada da resposta, a maior frequência de TO foi capaz de alterar a taxa de resposta de três dos quatro sujeitos, promovendo uma leve redução na taxa de resposta da alternativa com TO (R2). No entanto, o aumento sistemático da taxa de resposta na alternativa sem TO (R1), observado em todos os sujeitos, sugere que o TO também afetou a distribuição relativa das respostas entre as opções. Essa redistribuição comportamental está alinhada à proposta da teoria aditiva de Deluty (1976), segundo a qual a punição aumenta o valor relativo da alternativa não punida. Em contextos de escolha, portanto, o TO pode atuar como um modulador da alocação da resposta, promovendo uma reavaliação das alternativas disponíveis com base em seus valores relativos.

Uma interpretação complementar dos resultados é considerar que, em um procedimento concorrente simples, o TO programado para a opção R2 também afeta indiretamente a taxa de reforçamento da alternativa R1, funcionando como um TO não

contingente nessa opção. Isso ocorre porque, durante o TO, o acesso ao reforço é suspenso globalmente. Assim, embora o TO contingente não tenha reduzido diretamente a probabilidade de respostas em R2, ele promoveu uma desvalorização relativa dessa alternativa em comparação a R1.

Sob essa perspectiva, a situação pode ser interpretada como uma escolha entre reforço contínuo (R1, alta probabilidade de reforço, sem risco de interrupção pelo TO) e reforço probabilístico (R2, acesso ao reforço é incerto devido às interrupções pelo TO), em condições de estímulos mantidas equivalentes. Essa assimetria de acesso ao reforço contribui para aumentar o valor relativo da alternativa sem TO, o que se alinha ao padrão observado nos dados.

Essa análise destaca dois mecanismos que podem contribuir para a distribuição da resposta: (1) o papel da contingência resposta–TO, que acrescenta um custo específico à alternativa em que o TO é contingente; (2) a relevância da taxa global de reforçamento na supressão da resposta, ou seja, quanto maior a redução na taxa global de reforçamento, maior será a supressão da resposta. Nesse sentido, quanto maior a frequência do TO (i.e., menor o VI associado) ou maior a sua duração (mesmo com um impacto menor, e.g., Experimento 1), maior a probabilidade de que a resposta não seja reforçada, o que reduz a taxa global de reforçamento e, conseqüentemente, a taxa de resposta das opções.

A diferença dos resultados observados entre o Experimento 1 e o atual, pode ser compreendida justamente pela mudança na taxa de TO, que se tornou mais frequente. Esse efeito está em consonância com evidências da literatura sobre esquemas de punição, as quais indicam que estímulos aversivos mais frequentes exercem maior poder de supressão (Azrin e Holz, 1966; Hall, 2015). Contudo, os sujeitos haviam sido previamente expostos a diferentes durações de TO no Experimento 1, o que pode ter

implicado também uma possível superexposição ao procedimento de TO. Essa exposição prévia pode ter aumentado a sensibilidade ao TO e resultado em uma maior supressão da resposta alvo (Donaldson & Volmmer 2011). Em conjunto, os dados sugerem que o TO, quando apresentado com alta frequência, altera a preferência entre alternativas de forma robusta, mesmo sem uma redução uniforme na taxa de resposta entre os sujeitos.

Assim como no experimento anterior, o sujeito Y8 manteve taxas elevadas de resposta em ambas as opções, não demonstrando o efeito supressivo do TO, mesmo após a mudança no esquema de apresentação. Esse padrão reforça a variabilidade individual documentada na literatura sobre os efeitos do TO (Toegel et al., 2022; Hall, 2015), indicando que, mesmo diante de contingências aversivas, alguns indivíduos mantêm padrões de resposta pouco modulados por tais eventos.

As razões para esse comportamento contraintuitivo ainda não são totalmente compreendidas, mas a literatura sobre atraso do reforço já identificou padrões semelhantes em contextos de escolha. Estudos mostraram que alguns sujeitos optam consistentemente por alternativas associadas a maiores atrasos e menores taxas de reforço quando essas alternativas proporcionam maior previsibilidade na apresentação do estímulo reforçador (e.g., Cunningham & Shahan, 2018; Dunn et al., 2024). Nesse sentido, é possível que a sinalização diferencial associada ao TO tenha aumentado a previsibilidade temporal da entrega do reforço, contribuindo para a manutenção das respostas do sujeito Y8 na alternativa com TO.

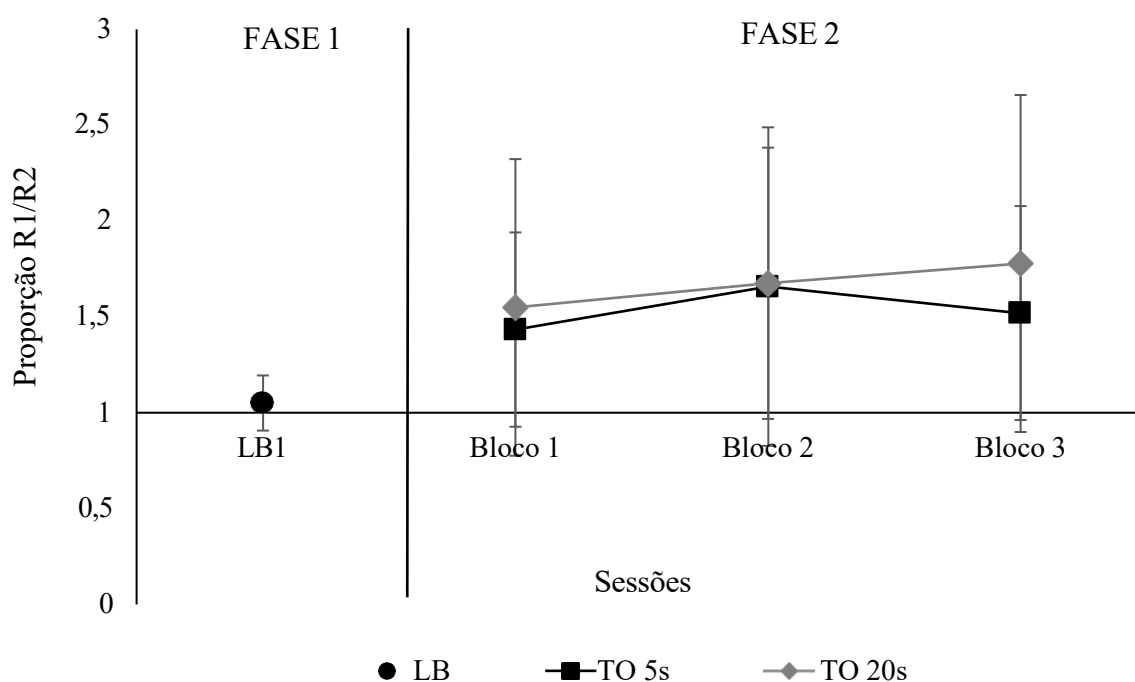
Todos os sujeitos diminuíram a taxa de respostas durante o período de TO (ver Apêndice D), o que demonstra que os efeitos discriminativos do período sem acesso ao estímulo reforçador foram bem estabelecidos pelo procedimento de TO, assim como no experimento anterior.

Efeito da Duração do TO ao Longo dos Blocos de Sessões

Com base na análise descritiva da Fase 2, observa-se uma tendência de aumento progressivo da proporção de respostas nos Blocos 2 e 3, especialmente na condição com TO de 20 s, sugerindo um aumento da preferência por R1 (opção sem TO) com o tempo de exposição. Contudo, a análise da proporção de respostas com base no GLMM não revelou efeito significativo para o fator Blocos de sessões ($\chi^2_{(2)} = 0.520$, $p = 0.771$) (ver figura 5 e os dados completos por bloco nos apêndices C).

Figura 5

Média geral da proporção da resposta das fases experimentais do Experimento 2.



Nota: A figura 5 apresenta a média geral da proporção da resposta 1 (R1) sobre a resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da Fase 2 do Experimento 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Duração do TO na Proporção da Resposta

Ainda na Fase 2, os sujeitos apresentaram uma proporção média de 1,52 para o TO de 5 s (SD= 0,55) e de 1,77 para o TO de 20 s (SD= 0,87), representando um aumento da proporção em direção à opção sem TO em comparação com a fase anterior (Fase 1: Prop = 1,03), mas sem diferença significativa nas proporções em função das durações do TO ($\chi^2_{(2)} = 0.546$, $p = 0.460$).

A análise individual revelou que três de quatro sujeitos (Y1, Y7 e Y9) preferiram a opção sem TO em ambas as durações, alocando mais respostas na opção R1 à medida que a duração do TO aumentava. Já o sujeito Y8 apresentou uma diminuição na proporção de respostas, passando a preferir a opção associada ao TO em todas as durações. Nesse caso, a alocação de respostas foi proporcional à duração do TO, ou seja, mais respostas foram emitidas na opção com TO (R2) à medida que a duração aumentava (ver Figura 6).

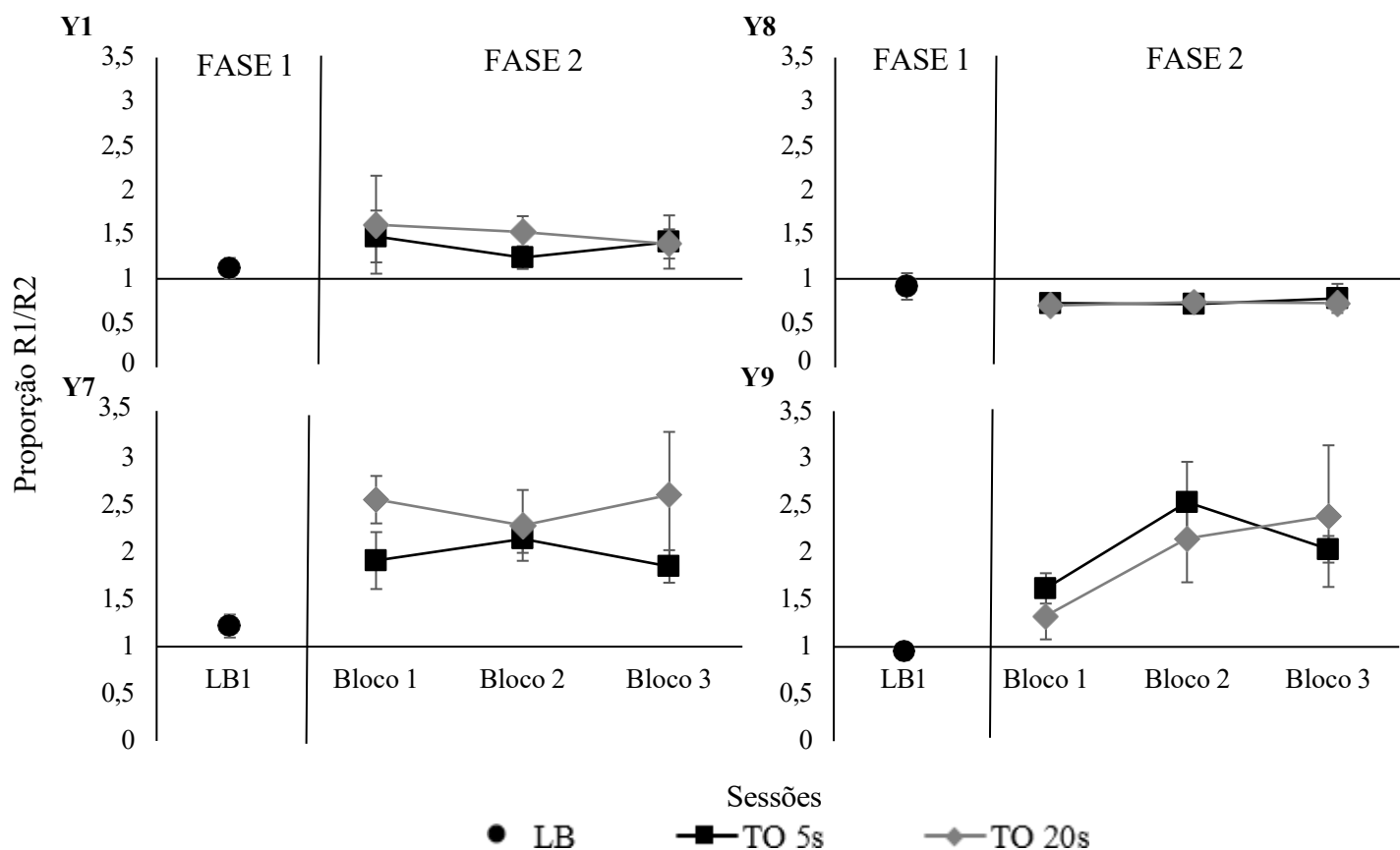
Para três sujeitos, a alocação crescente de respostas na opção sem TO conforme o aumento da duração, sugere um efeito aversivo graduado, em que TOs mais longos tornaram a alternativa associada a eles menos atrativa. Esse padrão seria o esperado sob a perspectiva tanto das teorias da punição (Deluty, 1976; de Villiers, 1980) quanto das teorias do atraso, como a DRT, segundo as quais os estímulos que produzem maior taxa de punição ou maior atraso até o reforço alteram o valor relativo das opções, nesse caso, favorecendo a opção sem TO.

Similar ao observado no experimento anterior, o comportamento do sujeito Y8 contrasta com esse padrão, revelando preferência pela opção com TO, especialmente em durações mais longas. Esse padrão, mais uma vez, mostra-se compatível com os estudos de escolha subótima em que a sinalização temporal é valorizada ao tornar a disponibilização do reforço mais previsível (Cunningham & Shahan, 2019). Esses dados destacam que, embora o TO seja frequentemente classificado como um estímulo

punitivo, seus efeitos sobre a escolha podem depender de múltiplos fatores, incluindo o histórico de reforçamento, a sensibilidade ao tempo e as funções discriminativas adquiridas pelos estímulos envolvidos.

Figura 6

Média individual da proporção da resposta das fases experimentais do Experimento 2



Nota: A figura 6 apresenta a proporção individual da resposta 1 (R1) em relação à resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da Fase 2 do Experimento 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Duração do TO na Taxa de Reforço e na Duração da Sessão

Na Fase 2, observou-se que variações na duração do TO influenciaram o tempo total de sessão, o que, por sua vez, afetou as taxas de reforço e de TO quando consideradas em relação ao time-in (30 min) e ao tempo global. Para mais detalhes, ver Tabela 2.

Tabela 2

Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 2.

TO	Tempo total de sessão	Taxa de reforço K1 (Time-in)	Taxa de reforço K2 (Time-in)	Taxa de reforço global R1	Taxa de reforço global R2	Taxa de TO R2 (Time-in)	Taxa de TO global R2
5s	60,81 (1,76)	3,03 (0,10)	2,89 (0,07)	1,50 (0,09)	1,42 (0,04)	6,30 (0,64)	3,10 (0,22)
20s	105,61 (12,14)	3,06 (0,05)	2,84 (0,19)	0,88 (0,08)	0,81 (0,03)	6,07 (1,11)	1,71 (0,11)

Nota: A tabela 2 apresenta a média do tempo total de sessão (em minutos); a taxa de reforço no time-in para as opções R1 e R2; a taxa de reforço global para as opções R1 e R2; e a taxa de TO no time-in e a taxa de TO global do Bloco 3 da Fase 2 do Experimento 2. Os números entre parênteses representam o desvio padrão de cada média.

A tabela 2 mostra que, mesmo com a mudança da taxa do TO, os sujeitos continuaram recebendo, em média, todos os reforços programados para o período de time-in na opção R1, independentemente da duração do TO. Na opção R2, houve uma pequena redução da taxa de reforço programada durante o time-in de 4% na duração de 5 s e de 5% na duração de 20 s, aproximadamente. A redução na taxa de reforço foi mais evidente na análise da taxa global. Nessa condição, a taxa de reforço foi reduzida em 50% (R1) e 51% (R2) na condição com TO de 5 s e em 71% em ambas as opções na condição com TO de 20 s. A tabela mostra ainda que os sujeitos receberam, em média, metade dos TO programados durante o time-in e que a taxa de TO global também foi reduzida para 51% no TO de 5 s e 72% no TO de 20 s, mantendo um padrão de queda semelhante ao observado para os reforços.

Logo, observa-se que, mesmo com o TO mais frequente, a taxa de reforço, em média, se manteve próxima ao valor esperado para o período de time-in e, mesmo com

as reduções na taxa global de reforço, a distribuição dos reforços entre as opções foi similar. Paralelamente, os dados demonstraram que os sujeitos não passaram por todos os TO programados, o que indica mudança na alocação da resposta para a opção sem TO.

Experimento 3:

Efeito da Duração do TO em Um Esquema de Cadeia Concorrente

Nos dois experimentos anteriores, a manipulação da duração do TO foi conduzida em esquemas concorrentes simples. Neles, observou-se que, para a maioria dos sujeitos, o TO afetou a alocação das respostas entre as alternativas disponíveis; no entanto, não foram observadas diferenças consistentes na alocação entre as durações do TO nem uma redução consistente na taxa de resposta. Para alguns sujeitos, observou-se um aumento da preferência pela opção com TO proporcional ao aumento da duração do TO. Uma possível explicação é que, em esquemas concorrentes simples, os sujeitos ainda podem obter reforços de ambas as alternativas ao longo da sessão, mesmo após escolher uma opção relativamente menos favorecida, como foi observado nos experimentos anteriores em que os sujeitos receberam, em média, todos os reforços programados durante o time-in e também receberam proporções similares de reforço entre as opções ao longo do tempo total da sessão. Logo, apesar do TO ser contingente à apenas uma das opções, ele afeta a taxa global na presença de ambas as opções de resposta.

Para avaliar essa hipótese, foi conduzido um terceiro experimento utilizando um esquema de cadeia concorrente. Nesse tipo de arranjo, uma resposta em uma das opções ativa uma cadeia de dois elos, de modo que o sujeito só pode acessar o reforço daquela alternativa, enquanto a outra opção fica indisponível. Apenas na próxima tentativa poderá escolher novamente entre as opções.

Com base no Experimento 1 de Carlson (1972), o presente experimento adotou um arranjo semelhante, com modificações adaptadas ao objetivo de analisar se, em um esquema de cadeia concorrente, a duração do TO impacta de forma mais consistente a alocação das respostas, revelando diferenças que não se mostraram evidentes nos experimentos anteriores.

Método

Sujeitos

Foram utilizados 8 ratos albinos machos (*Rattus norvegicus*), linhagem Wistar, experimentalmente ingênuos, provenientes do Instituto Evandro Chagas. As condições de manutenção foram idênticas às dos experimentos anteriores.

Ambiente Experimental

O experimento foi realizado na sala 12 do Laboratório de Processos Comportamentais Complexos (LPCC), localizada no prédio do Núcleo de Teoria e Pesquisa do Comportamento 1 (NTPC 1). A sala era similar a sala dos experimentos anteriores e as condições ambientais também foram as mesmas.

Materiais e Equipamentos

Os materiais e equipamentos utilizados foram os mesmos descritos nos experimentos anteriores.

Procedimento Geral

As sessões experimentais ocorreram diariamente no período da tarde, aproximadamente no mesmo horário todos os dias. Inicialmente, os sujeitos passaram pelos procedimentos de treino ao bebedouro, modelagem e treino da resposta, similares aos experimentos anteriores. Esse experimento utilizou um procedimento de cadeias concorrentes e foi composto por quatro fases experimentais: Linha de base (Fase 1), TO (Fase 2), Linha de base 2 (Fase 3) e TO 2 (Fase 4).

Fase 1 (Linha de Base)

Durante todas as fases do experimento, ambas as barras permaneciam disponíveis na câmara operante (esquerda e direita). Cada tentativa iniciava-se com o elo inicial, no qual a luz geral da caixa permanecia acesa e as luzes das barras, apagadas. Nesse momento, o sujeito podia responder livremente em qualquer uma das barras, sendo que cada uma operava sob um esquema de reforço VI 30 s independentes.

Para evitar alternâncias sucessivas entre as barras e promover a escolha consistente, foi programado um COD de 5 segundos. Assim, uma resposta em determinada barra só poderia encerrar o elo inicial caso não tivesse sido precedida, nos 5 segundos anteriores, por uma resposta na barra oposta. Após o cumprimento do esquema VI em uma das barras (i.e. a primeira resposta depois de finalizado o intervalo), era iniciado o elo final correspondente à opção escolhida. Durante o elo final, a luz geral da caixa permanecia acesa, e a luz da barra associada ao elo ativo assumia uma de duas condições: iluminação contínua ou intermitente (ciclo de 0,5 s), a depender da programação específica. A associação entre padrão luminoso (contínuo vs. intermitente para o elo final) e barra (esquerda ou direita para o elo inicial) foi balanceada entre os sujeitos para controle de viés.

Enquanto o elo final estava em vigor, apenas a barra correspondente permanecia ativa e programada para reforço sob um esquema VI 15 s. A barra oposta permanecia fisicamente disponível e todas as respostas nela eram registradas, mas não produziam consequências programadas nem apresentavam estímulos discriminativos.

O elo final era encerrado com a entrega de um reforço, iniciando-se imediatamente uma nova tentativa com um novo elo inicial. Cada sessão experimental continuava até que o sujeito obtivesse 60 reforços, e cada participante permanecia nesta fase por 10 sessões consecutivas.

Fase 2 (TO)

O procedimento foi semelhante ao da Fase 1, com a diferença de que um TO de duração variável (5 s ou 20 s), programado no esquema VI 15 s, foi sobreposto ao elo terminal da opção com maior alocação de respostas na Fase 1. Aleatoriamente, metade dos sujeitos foi exposta primeiro ao TO de 5 s (B2, B4, B5 e B7) e a outra metade primeiro ao TO de 20 s (B1, B3, B6 e B8).

Durante o período de TO, todas as luzes foram apagadas, todos os contadores de reforçamento foram parados, nenhuma barra disponibilizou reforçadores, mas todas as respostas de pressão à barra ocorridas durante esse período foram registradas. Após o TO, as luzes da caixa e da barra escolhida foram ligadas e os contadores do esquema de reforço foram retomados. O tempo de time-in e de TO foi contabilizado separadamente. Ao escolher a opção que contém o TO, a cadeia só era finalizada após a entrega do estímulo reforçador e a sessão se encerrava após a entrega de 60 reforços totais. Essa fase teve duração de 15 sessões.

Fase 3 (Linha de Base)

Seguiu os mesmos critérios da Fase 1.

Fase 4 (TO)

Essa fase seguiu os mesmos critérios da Fase 2; entretanto, os sujeitos expostos anteriormente ao TO de 5 s foram agora expostos ao TO de 20 s, e vice-versa.

Análise de Dados

Os dados foram submetidos a uma análise descritiva similar à dos experimentos anteriores, com a diferença de que a análise focou na fase de escolha (elos iniciais). Além disso, os dados passaram por uma análise estatística semelhante à dos experimentos anteriores, com a adição do efeito fixo “fase” no GLMM. Já a análise descritiva focou nos dados do Bloco 3 das Fases 2 e 4.

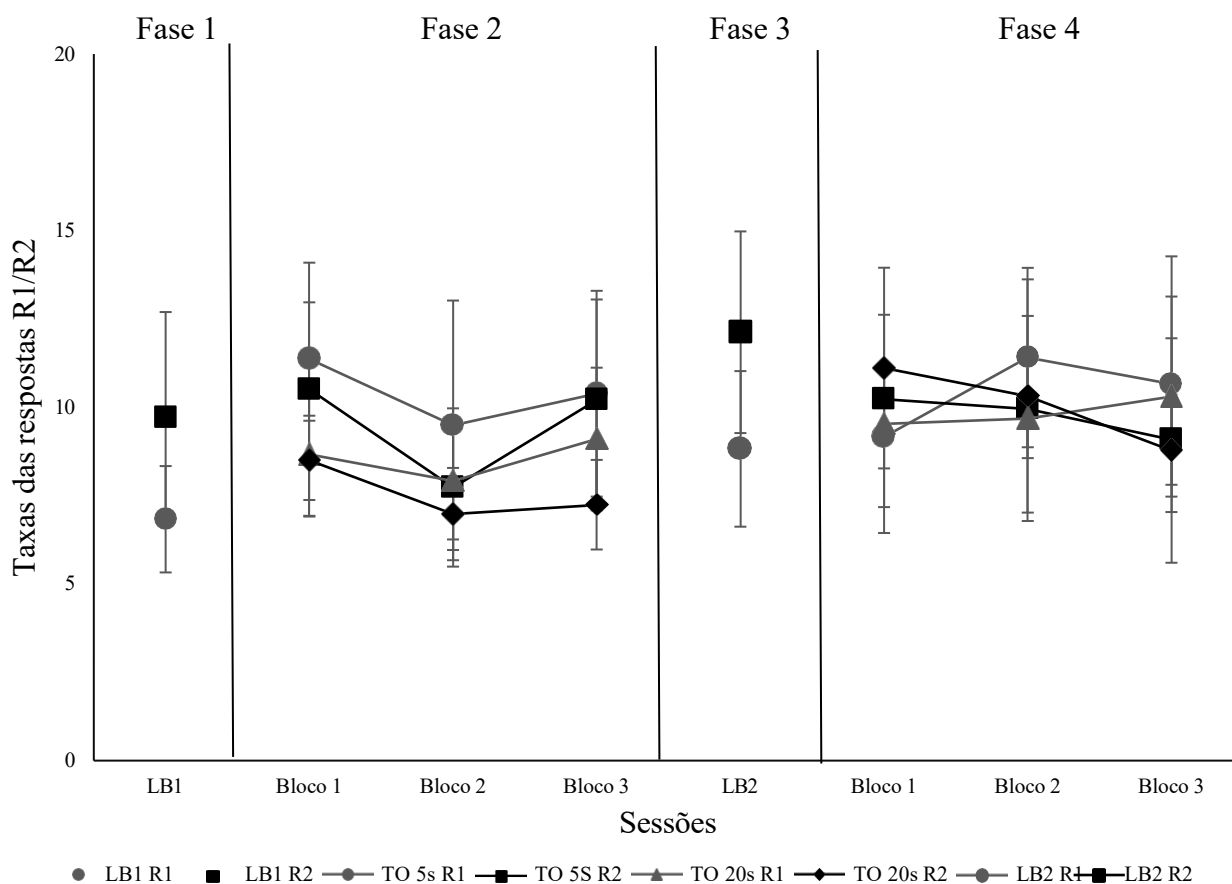
A análise de regressão não foi realizada, visto que esse experimento possui apenas duas durações do TO, o que não permite uma avaliação fidedigna da regressão.

Resultados e Discussão

Na Fase 1, em média (considerando as últimas seis sessões), emitiram 7 resp/min na opção R1 (SD = 1,50) e 10 resp/min na opção R2 (SD = 2,99) no elo inicial. A proporção de respostas foi, em média, 0,74 (SD = 0,12). Os sujeitos receberam o total de 60 reforços, distribuídos entre ambas as opções de resposta, em um tempo médio de sessão de 42 min (SD = 2,83). Para mais detalhes sobre os dados individuais, ver a Figura 7 e o Apêndice E.

Figura 7

Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 3



Nota: A figura 7 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e da Fase 3, e dos blocos de sessões da Fase 2 e da Fase 4 do Experimento 3. O TO foi apresentado contingente apenas à opção R2 da Fase 2 e da Fase 4. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão. Ressalta-se que a opção R2 foi definida com base na preferência do sujeito na Fase 1, o que contribuiu para uma diferença visual entre as opções nessa fase.

Efeito da duração do TO na taxa da resposta (Fases 2 e 4).

Na Fase 2, durante a apresentação do TO de 5 s, os sujeitos apresentaram uma média de 10 resp/min na opção R1 (SD = 2,91) e R2 (SD = 2,84). Todos os sujeitos aumentaram a taxa de resposta na opção R1 (sem TO) em relação à fase anterior. Na opção R2, metade dos sujeitos apresentou um aumento na taxa (B2 e B7), enquanto a outra metade apresentou um decréscimo (B4 e B5), em comparação com a Fase 1.

Os sujeitos que passaram pelo TO de 20s apresentaram uma média de 9 resp/min (SD = 2,01) na opção R1 e de 7 resp/min (SD = 1,26) na opção R2. Desses sujeitos, dois aumentaram a taxa de resposta na opção R1 (B1 e B6) e dois diminuíram a taxa de resposta (B3 e B8). Na opção R2, dois sujeitos apresentaram diminuição na taxa de resposta (B3 e B8), um aumentou a taxa de resposta (B6) e um permaneceu inalterado (B1), em comparação com os dados da fase anterior.

Na Fase 4, os sujeitos que passaram pelo TO de 5 s apresentaram uma média de 11 resp/min (SD = 3,61) para a opção R1 e de 9 resp/min (SD = 1,27) para a opção R2. Dessa forma, três sujeitos (B1, B6 e B8) aumentaram a taxa de resposta na opção R1 (sem TO) e B3 diminuiu nessa mesma opção. Na opção R2, todos os sujeitos diminuíram a taxa de resposta em relação à Fase 3.

Os sujeitos que passaram pelo TO de 20s apresentaram uma média de 10 resp/min (SD = 2,83) na opção R1 e de 9 resp/min (SD = 3,17) na opção R2. Três

sujeitos (B2, B3 e B7) aumentaram a taxa de resposta na opção R1, enquanto o sujeito B4 apresentou diminuição na opção R1. Na opção R2, todos os sujeitos apresentaram diminuição da taxa de resposta, com base nos dados da Fase 3 (ver Figura 7 e Apêndice E).

A análise do GLMM não revelou efeitos significativos para o fator Duração do TO ($\chi^2_{(1)} = 1.56$, $p = .211$), sugerindo que, no bloco final, a variação da duração do TO entre 5 s e 20 s não afetou significativamente a taxa de resposta nas alternativas.

Uma possível explicação para a ausência de efeito significativo da duração do TO neste experimento, em contraste com os dois anteriores, pode estar relacionada às particularidades do procedimento em cadeia concorrente. Nesse caso, apenas o elo inicial foi analisado, o que restringiu a taxa de respostas observáveis em comparação à situação de escolha livre, na qual a alocação podia ser continuamente ajustada pelos sujeitos. Essa configuração pode ter reduzido o poder estatístico para detectar diferenças sutis entre as durações do TO.

Ainda assim, observou-se que a introdução do TO produziu, mesmo sem diferenças significativas, uma redução na taxa de respostas na opção associada ao TO (R2) e, paralelamente, um aumento na taxa de respostas na alternativa sem TO (R1). Esse padrão de resultados é consistente com os efeitos observados por Carlson (1972), que, ao empregar cadeias concorrentes com um TO programado para uma das alternativas, encontrou uma clara supressão das respostas na opção punida e um aumento da resposta na barra oposta. Carlson interpretou esse resultado como evidência do efeito punitivo do TO, já que sua presença reduzia seletivamente a atratividade da alternativa associada ao aversivo, deslocando a resposta para a alternativa não punida. Esses dados reforçam a ideia de que o TO pode atuar como uma contingência punitiva

eficaz em contextos de escolha, modulando a preferência de acordo com os custos aversivos associados a cada opção.

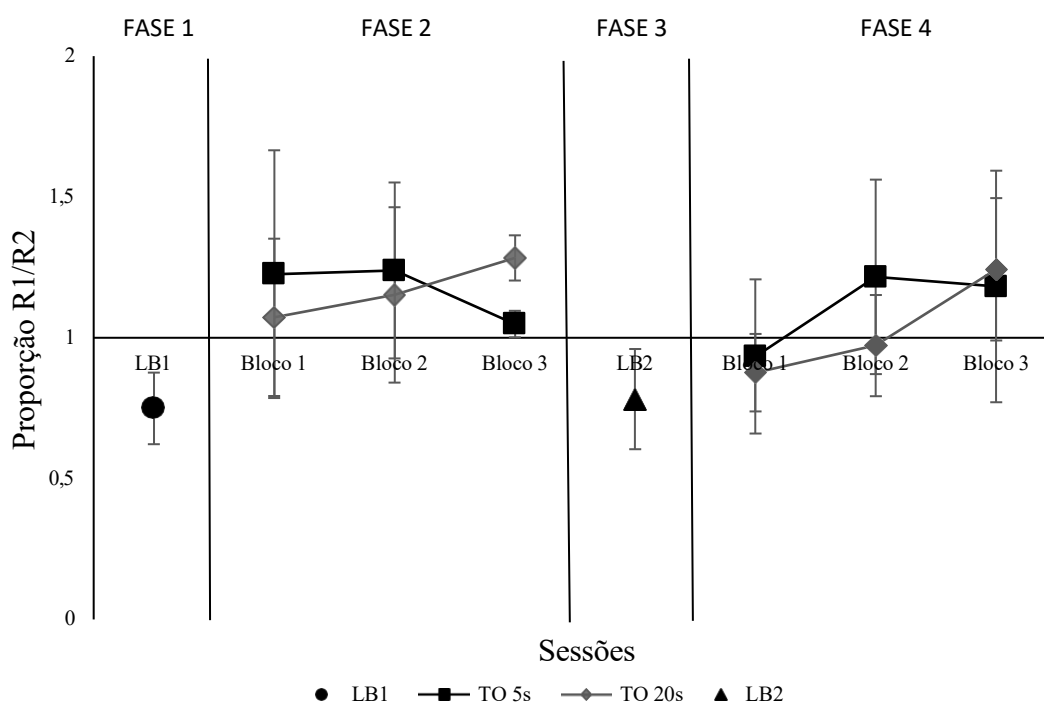
Efeito do TO ao longo das fases e blocos de sessões

A análise visual das fases com TO (Fase 2 e Fase 4) mostrou que houve um aumento na proporção de resposta em comparação às linhas de base (Fase 1 e Fase 3; sem TO). Isto é, houve um aumento relativo na taxa da resposta em direção à opção que não estava associada ao TO. Esse efeito foi mantido ao longo dos blocos de sessões, com pequenas variações. Enquanto a exposição ao TO de 5 s apresentou uma tendência decrescente nos últimos blocos, o TO de 20 s apresentou uma tendência de aumento ao longo dos blocos.

Na análise do GLMM, não foram encontradas diferenças estatisticamente significativas para o fator “fases” ($\chi^2_{(1)} = 1,383, p = 0,240$) nem para o fator “blocos de sessões” ($\chi^2_{(2)} = 3,087, p = 0,214$). Ver Figura 8.

Figura 8

Média geral da proporção da resposta das fases experimentais do Experimento 3



Nota: A figura 8 apresenta a média geral da proporção da resposta 1 (R1) sobre a resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da Fase 2 e Fase 4 do Experimento 3. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Duração do TO na Proporção da Resposta (Fases 2 e 4).

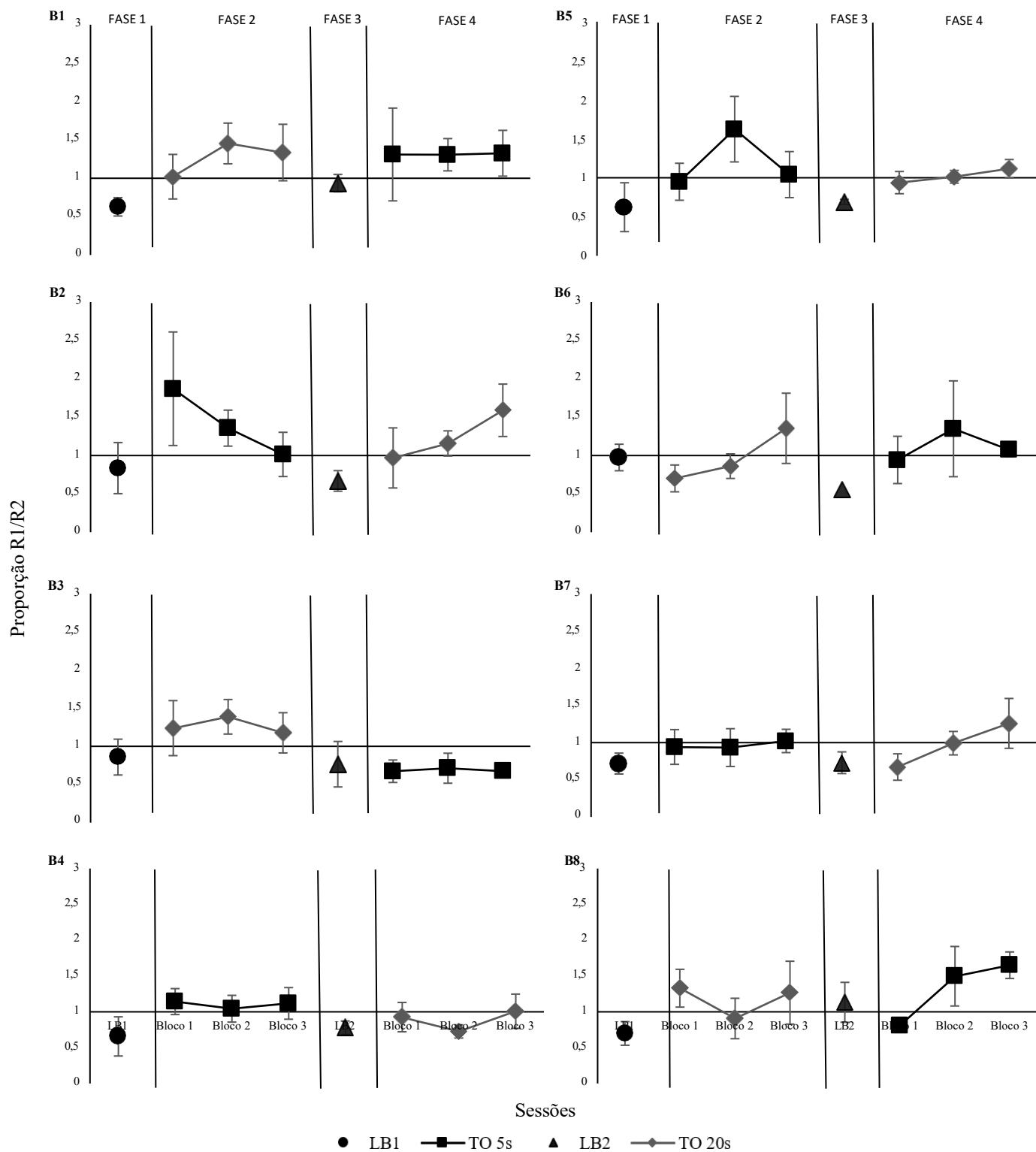
Ainda na Fase 2, os sujeitos expostos ao TO de 5 s apresentaram uma proporção média de 1,04 (SD= 0,04), enquanto os sujeitos expostos ao TO de 20 s apresentaram uma proporção média de 1,28 (SD= 0,08). A análise individual mostrou que a proporção de respostas de todos os sujeitos aumentou, indicando uma preferência pela opção R1 (sem TO), independentemente da duração do TO.

Já na Fase 4, os sujeitos do TO de 5 s apresentaram uma proporção média de 1,18 (SD = 0,41) no bloco 3. Dentre os sujeitos desse grupo, B1, B6 e B8 aumentaram a proporção da resposta em direção à opção sem TO (R1), enquanto o sujeito B3 diminuiu a proporção da resposta em direção à opção com TO (R2), em comparação com a Fase 3. Os sujeitos do TO de 20 s apresentaram uma proporção média de 1,24 (SD = 0,25) e todos os sujeitos aumentaram a proporção da resposta em comparação com a Fase 3. Para mais detalhes, ver a Figura 9.

Ao comparar a proporção da resposta em função das durações do TO, o GLMM não apresentou diferenças significativas entre as durações do TO na proporção da resposta ($\chi^2_{(1)} = 0,306, p = 0,580$).

Figura 9

Média individual da proporção da resposta das fases experimentais do Experimento 3



Nota: A figura 9 apresenta a proporção individual da resposta 1 (R1) sobre a resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da Fase 2 e

da Fase 4 do Experimento 3. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Os resultados do Experimento 3 demonstram que a introdução do TO afetou de maneira sistemática a alocação da resposta, promovendo uma preferência consistente pela opção sem TO (R1) para todos os sujeitos na Fase 2 e para 7 de 8 sujeitos na Fase 4, independentemente da duração em vigor. Na Fase 2, ambos os grupos (TO de 5 s e 20 s) apresentaram aumento na proporção de respostas em direção à alternativa sem TO, sendo esse efeito mais pronunciado no grupo exposto ao TO de 20 s do que no grupo exposto ao TO de 5 s. A análise individual reforça esse padrão, indicando que todos os sujeitos aumentaram a proporção de respostas, revelando uma clara sensibilidade ao TO enquanto contingência aversiva. Esse resultado também está de acordo com efeitos já observados por Carlson (1972), em que foi observada uma alteração consistente na alocação da resposta para a alternativa sem TO.

Na Fase 4, com a reintrodução do TO, observou-se novamente um aumento na proporção de respostas em ambas as durações, sugerindo uma retomada da preferência por R1 em comparação às fases sem TO. Embora os dados individuais apresentem alguma variabilidade, com um sujeito (B3) preferindo consistentemente a opção com TO, o padrão geral confirma que a presença do TO levou à reorganização do comportamento em favor da alternativa não punida.

O resultado do GLMM sugere que, embora o aumento da duração possa ter algum impacto, a simples presença do TO já é suficiente para promover mudança na escolha. Assim, os dados indicam que o TO, independentemente de sua duração, tem efeitos semelhantes sobre a alocação da resposta em cadeias concorrentes, promovendo preferência pela alternativa que evita a interrupção do acesso ao estímulo reforçador.

A ausência de diferenças significativas entre as durações do TO, observada neste e nos experimentos anteriores, está de acordo com achados da literatura (e.g. Naden & Morgan, 1988; McGuffin, 1991; White et al., 1972), que também não encontraram efeitos sistemáticos da variação da duração do TO sobre o comportamento, demonstrando que durações mais longas de TO não necessariamente resultam em maior supressão da resposta. Esses achados convergem para a ideia de que há um limite funcional para a eficácia do TO como estímulo punitivo e que, acima de determinada duração mínima, aumentos adicionais podem não produzir efeitos proporcionais. Essa interpretação reforça a importância de considerar outros parâmetros do TO além da duração, como o esquema de apresentação, ao investigar seus efeitos sobre a escolha e a alocação de respostas.

Efeito da Duração do TO na Taxa de Reforço e na Duração da Sessão (Fases 2 e 4)

Na Fase 2 e Fase 4, observou-se que a variação na duração do TO impactou tanto a duração do time-in, que, diferente dos experimentos anteriores, não tinha um tempo fixo para finalizar a sessão, quanto o tempo global das sessões, influenciando também as taxas de reforço e de TO. Para mais detalhes, ver Tabela 3.

Tabela 3

Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 3.

TO	Tempo de sessão (Time-in)	Tempo total de sessão	Taxa de reforço R1 (Time-in)	Taxa de reforço R2 (Time-in)	Taxa de reforço global R1	Taxa de reforço global R2	Taxa de TO R2 (Time-in)	Taxa de TO global R2
FASE 2								
5c	40,57 (3,96)	47,41 (3,73)	0,76 (0,07)	0,71 (0,11)	0,65 (0,05)	0,60 (0,09)	0,72 (0,09)	0,61 (0,07)
70c	41,63 (1,44)	56,82 (0,86)	0,70 (0,03)	0,74 (0,08)	0,52 (0,03)	0,54 (0,04)	0,76 (0,09)	0,55 (0,05)
FASE 4								
5c	39,23 (2,33)	46,75 (2,61)	0,67 (0,13)	0,86 (0,06)	0,56 (0,10)	0,72 (0,05)	0,89 (0,02)	0,75 (0,01)

20s	41,14 (4,85)	55,91 (5,66)	0,72 (0,10)	0,75 (0,06)	0,53 (0,06)	0,55 (0,03)	0,75 (0,03)	0,55 (0,01)
------------	-----------------	-----------------	----------------	----------------	----------------	----------------	----------------	----------------

Nota: A tabela 3 apresenta a média do tempo de sessão no time-in e do tempo de sessão total (em minutos); a taxa de reforço das opções R1 e R2 no time-in, a taxa de reforço global das opções R1 e R2 e a taxa de TO no time-in e a taxa de TO global do Bloco 3 da Fase 2 e Fase 4 do Experimento 3. Os números entre parênteses representam o desvio padrão de cada média.

Os dados sugerem que o TO, assim como nos experimentos anteriores, aumentou o tempo total de sessão, mas, diferentemente dos Experimentos 1 e 2, o aumento não foi tão expressivo. Houve aumento médio no TO de 5 s e de 20 s na Fase 1 e na Fase 2, respectivamente. Apesar do TO ter apresentado menor diferença na alteração temporal, percebe-se que, ainda assim, ele operou como um modulador da densidade de reforçamento ao longo da sessão. A análise entre as opções com e sem TO (R2 e R1, respectivamente) evidencia essa função. Enquanto as taxas de reforço durante o time-in se mantiveram relativamente próximas entre as opções, as reduções observadas no tempo total foram mais expressivas na alternativa com TO. No TO de 5 s, por exemplo, a taxa de reforço de R2 caiu 15% quando considerada no tempo total, em comparação com uma queda de 14% para R1. Esse padrão se acentua no TO de 20 s, com uma redução de 27% para R2 frente a 26% em R1, na Fase 2.

Contudo, na Fase 4, houve uma redução equivalente na taxa de R1 e R2 (16%) no TO de 5 s, enquanto, na duração de 20 s, o padrão de redução foi semelhante ao da Fase 2, em que R2 (27%) apresentou maior redução do que R1 (26%). Essa diferença pode indicar uma diminuição da aversividade do TO para os sujeitos que passaram pela duração maior na primeira exposição (20 s) e, posteriormente, foram expostos a uma duração menor (5 s), sugerindo que os sujeitos foram sensíveis às mudanças de duração.

A taxa de TO, como esperado, também foi reduzida ao considerar o tempo total da sessão, apresentando uma diminuição de 15% no TO de 5s e de 28% no TO de 20s na Fase 2; e de 16% no TO de 5s e de 27% no TO de 20s na Fase 4. Logo, os dados sugerem que, embora os sujeitos continuem obtendo os reforços programados no *time-in*, a introdução do TO altera significativamente o cenário global da contingência, produzindo um custo de tempo que pode tornar a alternativa com TO relativamente menos vantajosa. Em termos funcionais, isso está alinhado com a ideia de que o TO pode atuar como um estímulo sinalizador de atraso no acesso ao reforço, reduzindo a atratividade da opção onde é aplicado (Fantino, 1969). Assim, o TO pode modificar o valor relativo das alternativas por inseri-las em um cenário de maior atraso e menor densidade geral do reforço.

Na Fase 3, os sujeitos passaram por 10 sessões. Com a descontinuação do TO, foram emitidas, em média, 9 resp/min na opção R1 (SD = 2,20) e 12 resp/min na opção R2 (SD = 2,85) no elo inicial. Em comparação com as fases anteriores, todos os sujeitos aumentaram a taxa de resposta na opção R1, e sete sujeitos também apresentaram taxas mais altas na opção R2 quando comparados com a Fase 1. Apenas o sujeito B8 apresentou uma taxa de resposta abaixo da emitida na Fase 1. Em relação à Fase 2, cinco sujeitos (B1, B2, B5, B6 e B7) emitiram mais respostas na opção R1, enquanto três sujeitos (B3, B4 e B8) emitiram menos respostas nessa mesma opção. Na opção R2, sete sujeitos apresentaram taxas mais altas do que na Fase 2, sendo apenas o sujeito B7 o que apresentou uma taxa mais baixa nessa opção. Para mais detalhes individuais, ver Apêndice E.

A proporção de respostas (Figura 7) foi, em média, 0,78 (SD = 0,17), indicando um aumento da preferência pela opção R1 em comparação com a Fase 1 para cinco sujeitos (B1, B4, B5, B7 e B8) e um aumento da preferência pela opção R2 para três

sujeitos (B2, B3 e B6). Em relação à Fase 2, todos os sujeitos apresentaram uma proporção mais baixa na Fase 3, o que indica um aumento da preferência pela opção R2. Os sujeitos receberam os 60 reforços totais de ambas as opções de resposta em um tempo médio de sessão de 39 min (SD = 1,18).

A retirada do TO na Fase 3 produziu um efeito claro sobre o comportamento dos sujeitos, refletido tanto no aumento das taxas de resposta quanto na redistribuição da alocação entre as opções. Em termos de taxa de resposta, os dados mostram um efeito de recuperação da resposta após a remoção da contingência aversiva para a maioria dos sujeitos em ambas as opções de resposta. Esse padrão reforça a evidência de que o TO apresentou um efeito supressor durante sua presença e que sua retirada permitiu a retomada de níveis mais altos de resposta, em linha com achados anteriores da literatura (e.g., Carlson, 1972). Esse padrão também foi observado nos dados da proporção de respostas, em que houve uma diminuição do valor em comparação com a Fase 2 (embora ainda mais alta do que a Fase 1), sugerindo que a retirada do TO também reduziu a diferença de valor entre as opções, tornando a opção R2 mais atrativa após sua retirada.

Em síntese, os resultados do Experimento 3 indicam que o TO, programado no elo final de cadeias concorrentes, exerceu efeito supressivo sobre a taxa de resposta na alternativa em que foi apresentado e aumentou a alocação de respostas na opção sem TO. Ainda que a diferença entre as durações (5 s e 20 s) não tenha gerado efeitos significativamente distintos na proporção de respostas, a presença do TO foi suficiente para modular o padrão de resposta dos sujeitos.

Em parte, esse resultado pode ser reflexo da estrutura do experimento atual com cadeia concorrente, na qual a escolha feita no elo inicial definiu as consequências recebidas no elo final, sem possibilidade de transição entre opções durante a tentativa.

Esse contexto pode ter amplificado o valor aversivo do TO, uma vez que o sujeito era impedido de acessar o reforço em ambas as alternativas, o que pode ter intensificado o custo de selecionar a opção com TO.

Já a retirada do TO resultou em aumento geral das taxas e em redistribuição das escolhas, evidenciando seu papel como estímulo aversivo nesse arranjo. Esses achados replicam os efeitos observados em estudos clássicos, como o de Carlson (1972), contribuindo para a compreensão do TO em contextos de escolha mais complexos.

Experimento 4:

Efeito da Taxa do TO em Um Esquema Concorrente Simples

Estudos clássicos mostram que a taxa de aplicação do TO é crucial para sua eficácia como punição (Azrin & Holz, 1966). Thomas (1968) demonstrou que TOs mais frequentes (FR 2) geram maior supressão da resposta do que TOs menos frequentes, enquanto a duração do TO não teve efeitos consistentes. Resultados semelhantes foram encontrados em outros estudos com esquemas de razão e de intervalo fixo (Zimmerman & Ferster, 1963; Clark et al., 1973; Lerman et al., 1997).

Em esquemas de intervalo variável, porém, os efeitos do TO são menos consistentes. Ferster (1953) não encontrou supressão significativa, enquanto Dunn (1990) e Hall (2015) relataram que TOs mais frequentes levaram a maior alocação de resposta ou à redução da taxa de resposta. Esses achados indicam que a taxa de apresentação do TO é um fator crítico, embora ainda não esteja claro qual o limiar necessário para sua efetividade. Com base nisso, o presente estudo teve como objetivo analisar o efeito da taxa de apresentação do TO, utilizando esquemas de VIs, em esquemas concorrentes simples.

Método

Sujeitos

Foram utilizados 8 ratos albinos machos (*Rattus norvegicus*), linhagem Wistar, experimentalmente ingênuos, provenientes do Instituto Evandro Chagas. As condições de manutenção foram idênticas às dos experimentos 1 e 2.

Ambiente Experimental

O experimento foi realizado no mesmo ambiente dos experimentos 1 e 2.

Materiais e Equipamentos

Os materiais e equipamentos utilizados foram os mesmos descritos nos experimentos 1 e 2.

Procedimento Geral

Os procedimentos de treinamento ao bebedouro, de modelagem e de treino da resposta foram os mesmos dos experimentos 1 e 2.

Fase 1 (Linha de Base)

O procedimento foi semelhante ao dos experimentos 1 e 2; contudo, o COD foi alterado para 5 s.

Fase 2 (TO)

O procedimento foi semelhante aos experimentos 1 e 2; contudo, o TO foi apresentado com duração fixa de 20 s e o esquema de apresentação foi alternado a cada dia, variando entre VI 3 s, VI 5 s e VI 10 s. O COD do TO também foi alterado para 5 s.

Fase 3 (Linha de Base)

Os critérios de retorno à linha de base foram semelhantes aos do Experimento 1; contudo, o número mínimo de sessões para encerrar a fase foi de 10.

Análise de Dados

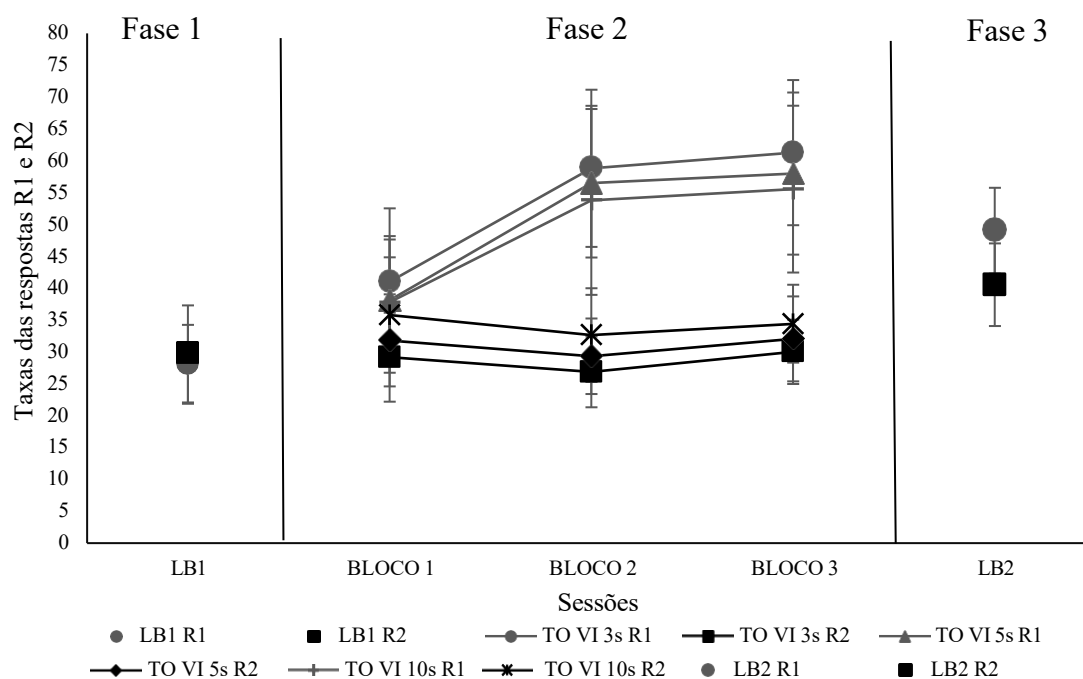
Os dados foram submetidos a uma análise similar à dos experimentos 1 e 2, com a diferença de que foi utilizado, como efeito fixo, o intervalo do TO em vez da duração.

Resultados e Discussão

Na Fase 1, os participantes alcançaram a estabilidade, em média, com 20 sessões (SD = 1,77). Ao final desta fase (i.e., média das últimas seis sessões), os sujeitos emitiram, em média, 29 resp/min na opção R1 (SD = 6,19) e 30 resp/min na opção R2 (SD = 7,61). A taxa média de reforços obtidos foi de 5 ref/min (R1: M= 2,42; SD = 0,10/ R2: M= 2,37; SD = 0,19). A proporção média das respostas (Figura 5) foi de 0,99 (SD = 0,31). O tempo médio total de sessão foi de 40 min (SD = 0,44). Para detalhes sobre os dados individuais, ver Figura 10 e Apêndice F.

Figura 10

Média geral da taxa de resposta das opções R1 e R2 das fases experimentais do Experimento 4



Nota: A figura 10 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e da Fase 3, e dos blocos de sessões da Fase 2 do Experimento 4. O TO foi apresentado contingente apenas à opção R2 da Fase 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Taxa do TO na Taxa da Resposta

Nessa fase, todos os sujeitos apresentaram um incremento na taxa de resposta nas opções R1 e R2 em todas as variações do TO, em comparação com a fase anterior. A média da taxa de R1 passou de 29 resp/min durante a Fase 1 para 61 resp/min (SD= 11,39) no TO VI 3s, 58 resp/min (SD= 12,72) no TO VI 5s e 56 resp/min (SD= 13,09) no TO VI 10s durante o Bloco 3. Na opção R2, a taxa da resposta permaneceu em 30 resp/min (SD= 5,04) no TO VI 3s, e passou para 32 resp/min (SD= 6,66) no TO VI 5s e 34 resp/min (SD= 6,11) no TO VI 10s durante o Bloco 3. Ver dados na Figura 10 e no Apêndice F.

A análise das taxas de resposta com base no GLMM revelou um efeito significativo do fator Opção de resposta ($\chi^2_{(1)} = 7.738$, $p = 0.005$), mas não do fator Intervalo do TO ($\chi^2_{(2)} = 1.497$, $p = 0.473$).

Uma interpretação possível é que esse padrão sugere que a presença ou ausência do TO em si exerceu maior influência sobre a taxa de respostas do que sua frequência. Além disso, embora se observe um leve aumento nas médias marginais à medida que o TO se torna menos frequente, especialmente na Opção 1, esse efeito parece ser sutil e não suficientemente robusto para alcançar significância estatística. Assim, dentro da faixa de valores manipulados (3–10 s), a frequência do TO não parece impactar de maneira expressiva a taxa de respostas.

Os dados da Fase 2 mostram um padrão que, à primeira vista, pode parecer contraditório: houve um aumento geral nas taxas de resposta nas duas opções (R1 e R2) após a introdução do TO, mesmo com a presença do TO em R2. Esse aumento, no entanto, foi mais acentuado na opção R1, sem TO, enquanto a taxa de resposta em R2 aumentou de forma mais modesta. Esse resultado se aproxima de achados anteriores (Experimento 1) que demonstraram que o TO pode alterar a alocação das respostas,

mesmo sem necessariamente produzir supressão direta da taxa de resposta em todas as condições.

Como discutido anteriormente, o TO pode estar atuando como pausas discriminativas que modulam a probabilidade de respostas no período com e sem reforço, análogo ao blecaute no estudo de Ferster (1953), ou ainda, como um componente distinto do período de time-in, o que favoreceria o efeito de contraste positivo, isto é, um aumento de respostas em um componente em função da redução de reforços no outro, análogo ao blecaute do estudo de Lattal (1984). Assim, o TO pode ser entendido não como punição, mas como um marcador temporal que sinaliza alternâncias previsíveis entre períodos de disponibilidade (time-in) e de indisponibilidade de reforço (TO), promovendo uma reorganização da alocação de respostas em vez de apenas suprimir o comportamento.

Durante o período de TO, ao final da Fase 2, observou-se que sete sujeitos apresentaram uma diminuição na taxa de respostas em ambas as opções, indicando que houve discriminação do período em que o reforço não estava disponível. O sujeito A8 foi o único a demonstrar um aumento na taxa de resposta em todas as variações do TO na opção R1 (ver Apêndice G).

Os dados de resposta durante o próprio período de TO demonstraram uma clara supressão da atividade para a maioria dos sujeitos, indicando que esses eventos foram discriminados como sinais de indisponibilidade de reforço. A exceção foi o sujeito A8, que manteve ou até aumentou a taxa de resposta durante o TO.

Efeito da Taxa do TO ao Longo dos Blocos de Sessões

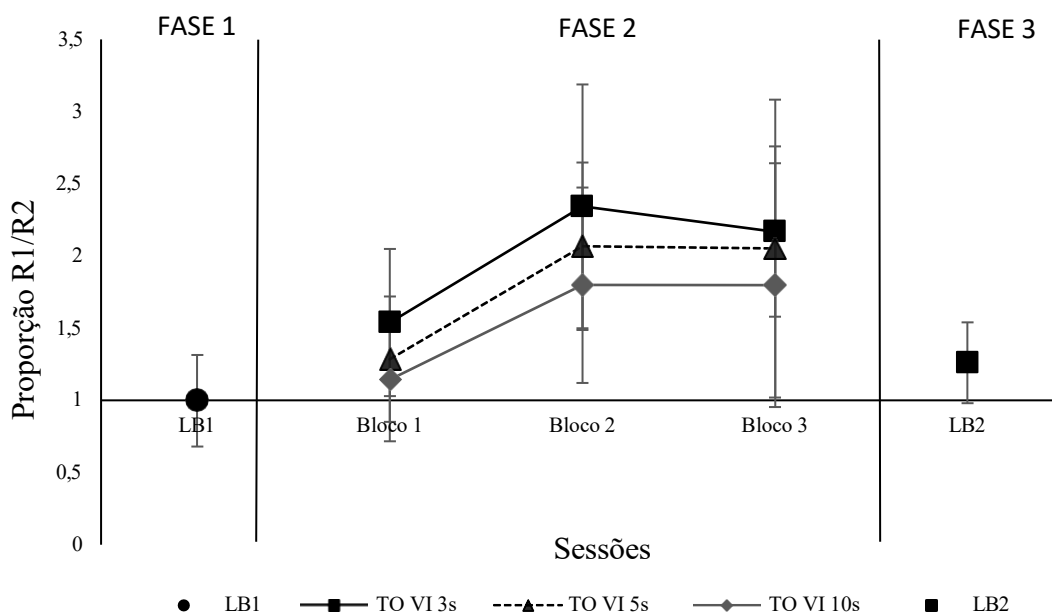
Com base na análise descritiva da Fase 2, observou-se um aumento na proporção de resposta ao longo da fase em todas as condições do TO, seguindo uma tendência progressiva com o VI do TO; ou seja, quanto mais frequente foi o TO, maior foi a

proporção de resposta na opção sem TO (R1). Esse efeito foi mais pronunciado nos blocos 1 e 2, seguido de uma leve desaceleração no bloco 3.

A análise da proporção de resposta com base no GLMM não revelou um efeito significativo para o fator Blocos de sessões ($\chi^2_{(2)} = 4.716, p = 0.095$). Para mais detalhes, ver Figura 11.

Figura 11

Média geral da proporção da resposta das fases experimentais do Experimento 4.



Nota: A figura 11 apresenta a média geral da proporção da resposta 1 (R1) sobre a resposta 2 (R2) das últimas seis sessões da Fase 1 e da Fase 3 e dos blocos de sessões da Fase 2 do Experimento 4. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Efeito da Taxa do TO na Proporção da Resposta

Ainda na Fase 2, os sujeitos apresentaram uma proporção média de respostas no TO VI 3 s de 2,16 (SD= 0,58), no TO VI 5 s, a proporção foi de 2,05 (SD= 1,03) e, no

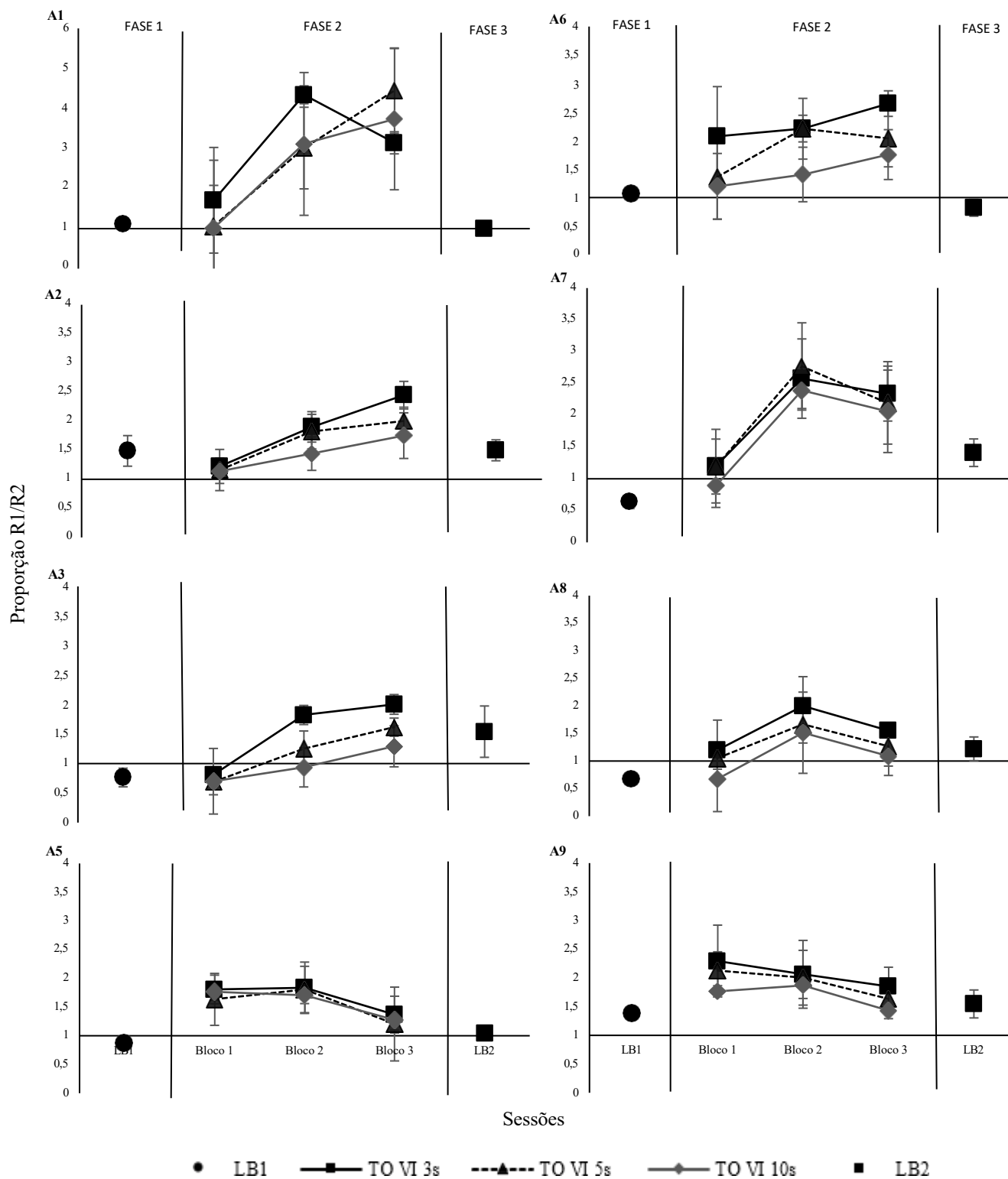
TO VI 10 s, a média da proporção foi de 1,79 (SD= 0,84) no bloco 3. O GLMM indicou diferença significativa na proporção entre os intervalos do TO ($\chi^2_{(2)} = 7.126, p = 0.028$).

Individualmente, observou-se que todos os sujeitos aumentaram a proporção da resposta em relação à Fase 1, em todas as variações do VI do TO durante o Bloco 3.

Para mais detalhes sobre as proporções, ver a Figura 12.

Figura 12

Média individual da proporção da resposta das fases experimentais do Experimento 4



Nota: A figura 12 apresenta a proporção individual da resposta 1 (R1) sobre a resposta 2 (R2) das últimas seis sessões da Fase 1 e Fase 3 e dos blocos de sessões da Fase 2 do Experimento 4. Além disso, o sujeito A1 apresentou uma proporção média maior do que

a dos demais sujeitos, o que exigiu ajustes nos valores do eixo vertical. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Os dados da Fase 2 demonstram um efeito claro da taxa de apresentação do TO sobre a alocação da resposta: quanto menor o intervalo (i.e., maior a frequência de apresentação do TO), maior foi a proporção de respostas emitidas na alternativa sem TO. Essa relação sugere que TO mais frequentes produziram um efeito mais robusto sobre o comportamento de escolha, padrão coerente com estudos que testaram diferentes taxas de TO (e.g. Thomas, 1968; Dunn, 1990; Hall, 2015)

Grace (1996) encontrou resultados similares ao estudo sobre o atraso do reforço, em que atrasos mais frequentes, independentes de serem fixos ou variáveis, aumentaram a sensibilidade da alocação de respostas para a opção com menor atraso na entrega do estímulo reforçador. No presente estudo, os TOs mais frequentes (VI 3 s e VI 5 s) podem ter elevado o atraso efetivo para a disponibilidade do reforço, aumentando a probabilidade de interrupções na opção à qual era contingente e, assim, reduzido seu valor relativo. Assim, a maior proporção de respostas na opção sem TO (R1) nas condições com VIs mais curtos pode refletir não um efeito punitivo do TO, mas uma mudança na estimativa temporal determinada pela frequência com que o acesso ao reforço era interrompido.

Efeito da Taxa do TO na Taxa de reforço e na Duração da Sessão

Durante a Fase 2, a variação na duração do TO afetou o tempo global das sessões, o que se refletiu também nas taxas de reforço e de TO globais quando comparadas com as taxas no time-in fixo (30 min). Mais detalhes podem ser consultados na Tabela 4.

Tabela 4

Média do tempo de sessão, taxa de reforço e taxa de TO no Bloco 3 do Experimento 4.

TO	Tempo total de sessão	Taxa de reforço R1 (Time-in)	Taxa de reforço R2 (Time-in)	Taxa de reforço global R1	Taxa de reforço global R2	Taxa de TO R2 (Time-in)	Taxa de TO global R2
VI 3s	89,97 (11,06)	2,79 (0,04)	2,35 (0,12)	0,95 (0,13)	0,79 (0,06)	4,73 (1,04)	1,56 (0,15)
VI 5s	85,16 (9,79)	2,74 (0,06)	2,36 (0,13)	0,98 (0,14)	0,84 (0,06)	4,28 (0,92)	1,49 (0,16)
VI 10s	74,04 (6,01)	2,67 (0,10)	2,39 (0,14)	1,09 (0,13)	0,97 (0,03)	3,22 (0,56)	1,29 (0,12)

Nota: A tabela 4 apresenta a média do tempo total de sessão (em minutos); a taxa de reforço no time-in, a taxa de reforço total e a taxa de TO no time-in do Bloco 3 da Fase 2 do Experimento 4. Os números entre parênteses representam o desvio padrão de cada média.

Os dados revelam que a introdução do TO impactou significativamente na taxa global de reforçamento obtida (i.e. número de reforços produzidos por responder nessa sessão dividido pelo tempo total da sessão; Time-in + TO), especialmente na opção associada ao TO (R2). A taxa de reforço global sofreu uma redução expressiva quando comparada à taxa durante o time-in: para R1, a queda foi de 60% (VI 3s), 64% (VI 5s) e 59% (VI 10s); já para R2, a redução foi mais acentuada no TO de VI 3s, com diminuições de 66%, apresentando uma redução semelhante à R1 nos outros intervalos. Esses dados indicam que, embora ambas as opções tenham tido a taxa de reforço global reduzida pela introdução do TO, o impacto foi consistentemente maior sobre R2 quando o TO foi mais frequente, o que contribui para explicar a maior alocação de respostas em R1. Além disso, a análise das taxas de TO mostra uma queda considerável entre os períodos de time-in e o tempo total da sessão: de 67% no VI 3s, 65% no VI 5s e 60% no VI 10s.

A análise das regressões lineares demonstrou que, com a mudança do esquema de apresentação do TO todas as VIs analisadas foram um preditor consistente e significativo da proporção da resposta (distribuição relativa da taxa local de

reforçamento ($F_{(3, 20)} = 8.135$, $p < 0.001$, R^2 ajustado = .482, AIC = 48.397, BIC = 54.287); distribuição relativa da taxa total de reforçamento ($F_{(3, 20)} = 8.581$, $p < 0.001$, R^2 ajustado = .497, AIC = 47.684, BIC = 53.574); Taxa do TO no time-in ($F_{(3, 20)} = 4.993$, $p = 0.010$, R^2 ajustado = .342, AIC = 54.122, BIC = 60.013); Taxa do TO no tempo total ($F_{(3, 20)} = 7.272$, $p = 0.002$, R^2 ajustado = .450, AIC = 49.838, BIC = 55.728).

Apesar de todos os modelos terem sido significativos, com uma diferença muito pequena na variância explicada, a distribuição relativa da taxa total de reforçamento foi capaz de explicar 49% da variação da proporção de resposta e apresentou valores de AIC e BIC menores que os dos demais modelos, indicando melhor ajuste aos dados analisados. Esse resultado sugere que os sujeitos estavam sensíveis tanto aos aspectos reforçadores quanto aos punitivos das contingências. No entanto, a proporção de reforço calculada em relação ao tempo total da sessão foi a variável com maior poder explicativo da escolha (R^2 ajustado = 0.497), superando os demais modelos. Esse achado é coerente com modelos quantitativos da escolha que consideram o valor relativo das alternativas com base na taxa efetiva de reforçamento (e.g. Herrnstein, 1961).

Os modelos quantitativos da escolha com punição (e.g. Deluty, 1976; de Villiers, 1980) sugerem uma interação entre as variáveis de reforço e punição, o que explicaria a significância de todas as VIs na explicação da VD nesse experimento. No estudo de de Villiers (1980), a escolha entre alternativas concorrentes pode ser determinada pela diferença entre as taxas de reforço e de punição, o que implica que o efeito punitivo do TO se daria pela redução líquida no valor reforçador de uma opção. Já a teoria de Deluty (1976) sugere uma interação aditiva entre reforço e punição; assim, a taxa de punição sobre uma opção aumenta o valor relativo da alternativa. Aplicando isso aos

dados do TO, pode-se interpretar que, quanto mais o TO impactava a disponibilidade de reforço em R2, a escolha se deslocava para a alternativa menos punida, pelo aumento do seu valor relativo.

Além disso, no Experimento 2, uma interpretação complementar sugeriu que o TO programado para a opção punida (R2) poderia afetar indiretamente a taxa de reforçamento da alternativa não punida (R1), funcionando como um TO não contingente e, assim, influenciando a taxa global de reforço disponível. Essa dinâmica cria uma assimetria entre as alternativas: em R1, o reforço tende a ser relativamente contínuo, já que só é interrompido pelos TOs produzidos em R2, enquanto em R2 o reforço se torna probabilístico, pois a própria resposta aumenta a chance de interrupção pelo TO, especialmente em contextos de altas taxas de TO, como observados nos Experimentos 2 e 4. Essa diferença na previsibilidade de acesso ao reforço configura um dos principais determinantes da distribuição da resposta.

Essa hipótese encontra suporte nos resultados do Experimento 4, em que a análise das regressões lineares demonstrou que a distribuição relativa da taxa total de reforçamento apresentou o melhor ajuste aos dados. Esse achado sugere que os sujeitos estavam sensíveis à variação no reforço efetivamente disponível ao longo da sessão, indicando que a relação entre a contingência do TO e a taxa global de reforço foi o principal fator na desvalorização relativa da opção punida.

De fato, mesmo que não tenham sido observadas diferenças significativas na taxa de resposta entre os diferentes intervalos de TO, a análise detalhada das taxas revelou que a opção sem TO recebeu mais respostas do que a opção com TO. As médias marginais indicaram que, embora as diferenças entre os intervalos de TO fossem pequenas, não eram inexistentes, mostrando uma tendência de maior taxa de resposta quando os TOs eram menos frequentes. Essa variação, mesmo que sutil, afetou a

distribuição relativa das respostas entre as alternativas. Em outras palavras, a frequência do TO modulou a alocação do comportamento, contribuindo para que a proporção de respostas favorecesse consistentemente a opção sem TO. Esse padrão confirma que a alternativa sujeita ao TO sofreu uma desvalorização relativa, uma vez que o reforço associado a ela se tornou menos acessível, impactando a escolha dos sujeitos.

Na Fase 3, com a descontinuação do TO, as taxas de resposta permaneceram mais altas do que a média observada na Fase 1, nas duas opções de resposta. Ao final desta fase (i.e. média das últimas seis sessões), os sujeitos emitiram, em média, 49,03 resp/min na opção R1 (SD = 6,73) e 40,55 resp/min na opção R2 (SD = 6,48). Em comparação com o Bloco 3 da Fase 2, na opção R1, seis dos oito sujeitos (A1, A2, A5, A6, A7 e A9) apresentaram uma diminuição na taxa de resposta em todas as variações de TO. Os sujeitos A3 e A8 apresentaram o mesmo padrão de decréscimo no TO VI 3s, mas apresentaram um aumento na R1 nos TO VI 5s e 10s. Na opção R2, cinco sujeitos (A1, A5, A6, A7 e A8) aumentaram a taxa de resposta em todas as variações do TO; dois sujeitos (A2 e A3) apresentaram esse mesmo padrão para os TO VI 3s e VI 5s, mas diminuíram a taxa durante o TO VI 10s; e o sujeito A9 diminuiu a taxa de resposta nessa opção em todas as variações do TO, em comparação com a Fase 2.

A proporção de respostas foi, em média, de 1,26 (SD = 0,27). Seis sujeitos (A2, A3, A5, A7, A8 e A9) apresentaram aumento na alocação de respostas na opção R1, enquanto os outros dois sujeitos (A1 e A6) aumentaram a alocação na opção R2, em relação à Fase 1. Quando comparada à Fase 2, sete dos oito sujeitos apresentaram diminuição na proporção em comparação com o Bloco 3, independentemente do intervalo do TO. O sujeito A9 apresentou o mesmo padrão em comparação com os TO VI 3s e VI5s, mas houve um aumento da proporção quando comparado ao TO VI 10s.

A taxa de reforço total foi de 5 ref/min (R1: $M = 2,46$; $SD = 0,03$ / R2: $M = 2,32$; $SD = 0,05$), e o tempo total da sessão foi de aproximadamente 40 min ($SD = 0,08$).

A Fase 3 demonstrou que, mesmo com a retirada do TO, as taxas de resposta permaneceram elevadas em relação à linha de base (Fase 1), sugerindo um possível efeito residual da contingência anterior. Contudo, em comparação com o último bloco da Fase 2, observou-se uma tendência geral de diminuição na taxa de resposta na opção R1 e de aumento na taxa de R2, especialmente quando comparadas aos intervalos mais curtos do TO, o que pode refletir um equilíbrio entre os valores das opções após a retirada do TO. Esse equilíbrio também foi observado na proporção da resposta, que apresentou diminuição da preferência por R1 e se aproximou dos níveis da Fase 1. Em suma, à medida que a taxa de TO aumentava, observou-se uma redução significativa na alocação de respostas na alternativa punida. Esse padrão se refletiu em proporções de resposta cada vez maiores a favor de R1, com efeitos mais pronunciados nos intervalos de maior frequência de TO. Os dados sugerem que os sujeitos foram sensíveis não apenas à presença do TO, mas também à sua frequência. Essa sensibilidade à taxa de TO corrobora a hipótese de que, em contextos de escolha, a taxa de reforço pode não ser a única determinante na alocação da resposta (de Villiers, 1980). Esses achados reforçam a necessidade de modelos de escolha que considerem tanto variáveis reforçadoras quanto aversivas para prever com maior precisão a distribuição do comportamento sob contingências concorrentes complexas.

Discussão Geral

O presente estudo propôs-se a analisar os efeitos do TO sobre o comportamento de escolha em procedimentos com esquemas concorrentes. Por meio de quatro experimentos conduzidos com sujeitos não humanos, foram manipulados dois parâmetros fundamentais do TO: sua duração e sua frequência de apresentação, a fim de

compreender como essas variáveis afetam não só a taxa de resposta, mas também sua alocação. A seguir, serão discutidos os principais achados dos experimentos, suas convergências e divergências, bem como possíveis explicações dos efeitos observados do TO.

O primeiro fator analisado ao longo dos experimentos foi o efeito do TO na taxa de resposta e, foi observado que, de forma geral, o TO não produziu supressão direta da taxa de resposta na maioria das condições testadas. Ao contrário, observou-se um aumento consistente na taxa de resposta durante o time-in, especialmente na alternativa sem TO, quando comparada à condição de linha de base. Esse resultado conduz a uma questão central: é possível caracterizar o TO como um evento punitivo se, de forma geral, ele não esteve associado a uma diminuição da taxa de resposta durante o time-in? Essa pergunta desafia a concepção tradicional da punição, que define o evento punitivo como aquele que, quando contingente à resposta, produz sua supressão (Azrin & Holz, 1966). Sob essa definição, os dados aqui obtidos parecem paradoxais. Em todos os experimentos, especialmente nas alternativas sem TO, observou-se um incremento da taxa de resposta durante o time-in, enquanto, nas opções com TO, os efeitos foram mais variáveis, ora neutros, ora associados a aumentos, ora associados a reduções da taxa. Essa ausência de supressão robusta e sistemática levanta dúvidas sobre a eficácia punitiva do TO quando analisada exclusivamente pela métrica da taxa de resposta.

Contudo, esse efeito geral de incremento se manifesta de formas diferentes dependendo do fator manipulado. A duração do TO afetou principalmente a taxa de respostas, de modo que TOs mais curtos resultaram em um maior número de respostas emitidas por minuto. Em contraste, o intervalo do TO não teve efeito significativo sobre a taxa de respostas, mas influenciou a distribuição relativa entre as opções, aumentando a proporção de respostas direcionadas à alternativa sem TO. Dessa forma, o efeito do

TO sobre o comportamento apresenta duas dimensões: a duração regula a quantidade de respostas, enquanto o intervalo regula a alocação relativa entre as alternativas, e ambos contribuem para o aumento observado em relação à linha de base.

Um aspecto relevante para a interpretação dos efeitos do TO sobre a taxa de resposta diz respeito ao contraste comportamental. Conforme definido por Reynolds (1961), o contraste ocorre quando mudanças em uma contingência de reforço ou punição em um componente produzem alterações na taxa de resposta em outro componente, mesmo que este último não tenha sofrido mudanças diretas. Esse efeito foi sugerido por Lattal (1984), que observou aumento da taxa de resposta na condição de reforço, em comparação com a condição de blecaute, apesar de não haver alterações na taxa de reforço, assim como ocorreu no presente estudo entre as condições de TO e time-in.

Rachlin (1973), ao discutir o contraste comportamental em esquemas concorrentes, traz uma perspectiva mais ampliada do fenômeno. Em esquemas concorrentes, observa-se que a taxa de respostas em um componente (por exemplo, A) varia inversamente com a taxa de reforço obtida em outro componente (B), independentemente de alterações diretas nas contingências de A (Catania, 1963; Rachlin & Baum, 1969). Esse fenômeno não se explica apenas pela redistribuição do tempo de resposta, mas sim pelo valor relativo dos componentes, de modo que qualquer aumento de reforço em B reduz o valor relativo de A, gerando diminuição da taxa de respostas em A, e vice-versa.

A distribuição de respostas em esquemas concorrentes segue padrões temporais específicos: os organismos respondem em taxas relativamente constantes, alternando períodos de resposta e de pausa, de forma que o aumento da duração dos períodos de resposta em um componente está associado à diminuição das respostas no outro

(Cotton, 1953; Gilbert, 1958). Dessa forma, o contraste comportamental reflete uma adaptação do comportamento à diferença relativa de valor entre opções disponíveis, e a distribuição temporal das respostas (períodos de atividade e inatividade) funciona como o mecanismo pelo qual essa adaptação se manifesta, evidenciando o efeito do valor relativo dos reforços sobre a alocação de comportamento (Rachlin, 1973). Essa interpretação apoia os dados do presente estudo, em que houve um aumento consistente no número de respostas na alternativa sem TO, mesmo com taxa de reforço constante.

Contudo, também foi observado um aumento na taxa de respostas na opção com TO. Uma possível explicação para esse fenômeno é fornecida pelo Experimento 2 de Hall (2015), no qual diferentes esquemas de VR e VI foram avaliados em esquemas múltiplos. No estudo de Hall (2015), foram observados efeitos semelhantes aos relatados no presente estudo, especialmente em sujeitos submetidos a frequências mais baixas de TO (VI 20, VI 45, VR 10, VR 15 e VR 20). Nesse contexto, alguns sujeitos não apresentaram alteração na taxa de respostas em relação à linha de base, enquanto outros exibiram aumento da resposta. Hall atribui essa variação à influência das pausas pós TO e pós reforço, pois ambas foram moduladas pela contingência e frequência do TO, porém de maneiras distintas: a pausa pós TO aumentou diretamente com a taxa de apresentação do TO, especialmente quando este era contingente à resposta. Já a pausa pós reforço foi afetada de forma indireta; quando o TO era frequente, os sujeitos aumentavam a latência para emitir novas respostas após receberem um reforço. Dessa forma, o TO não apenas prolongou a pausa imediatamente após sua ocorrência, mas também estendeu a pausa natural que segue a entrega de reforço, contribuindo de maneira combinada para a redução geral da taxa de respostas. No entanto, quando o TO era aplicado em baixa frequência ou de forma independente da resposta, esse efeito

combinado foi fraco, permitindo que a taxa global de respostas se mantivesse ou, em alguns casos, até aumentasse.

Embora o presente experimento tenha utilizado esquemas concorrentes simples e encadeados, Davson e MacCarthy (1988) sugerem que um esquema múltiplo pode ser compreendido pelos mesmos mecanismos que uma situação concorrente simples ou simultânea. Nesse sentido, a interpretação proposta por Hall poderia ser estendida aos dados observados neste estudo. No entanto, as pausas pós TO e pós reforço não foram incluídas na análise, o que limita a aplicação dessa explicação. Futuros trabalhos poderiam investigar de forma mais detalhada o papel dessas pausas em esquemas concorrentes, a fim de compreender melhor sua influência na alocação de respostas.

Toegel et al. (2022) propõem uma interpretação alternativa ao demonstrar que o TO, diferentemente de estímulos aversivos como o choque, não produz efeitos punitivos diretos e uniformes, mas depende da taxa de reforço disponível na contingência. No Experimento 4 do estudo de Toegel et al. (2022), o TO de 30 s em VR 5 foi analisado em esquemas múltiplos, com taxas de reforço em VI manipuladas para variar de 0,5 a 6 reforços/min e, nesse arranjo, observou-se uma relação inversa consistente entre a taxa de reforço e a supressão: o TO foi mais eficaz como punidor sob taxas de reforço mais altas.

Apesar disso, os autores ressaltam que essa relação entre TO e taxa de reforçamento não é linear. Eles argumentam que a taxa de resposta observada representa o resultado líquido da combinação entre reforço e punição, uma vez que ambos são contingentes à mesma resposta. Assim, o aumento da taxa de reforço pode intensificar simultaneamente a função aversiva do TO e a força da resposta reforçada. Como consequência, a interação entre a taxa de reforço e a eficácia punitiva do TO pode

resultar em diferentes padrões comportamentais: aumento, diminuição ou nenhuma alteração na taxa de resposta (Toegel et al., 2022).

Como o presente estudo analisou apenas duas taxas de reforçamento (VI 15s e VI 20s), não é possível afirmar que a hipótese de Toegel et al. (2022) se aplique aos dados aqui observados. Pesquisas futuras podem expandir essa investigação em contextos com densidades de reforçamento diferentes, a fim de verificar se existe um limiar de reforçamento a partir do qual o TO se torna efetivamente um supressor da taxa de resposta.

Nesse sentido, a supressão de uma resposta punida não parece ser o único, e possivelmente nem o melhor, indicador dos efeitos aversivos de um estímulo como o TO. Com base nisso, o presente estudo também analisou a alocação de respostas em esquemas concorrentes, simples e encadeados. A distribuição do comportamento entre as opções de resposta pode ser interpretada como um produto da interação entre variáveis reforçadoras e punitivas (Fontes, 2022), o que representa uma métrica alternativa e mais sensível para avaliar os efeitos aversivos do TO.

O primeiro aspecto observado em relação à alocação da resposta foi o efeito das exposições ao TO ao longo dos blocos de sessões. Observaram-se, nos quatro experimentos, padrões temporais distintos, mas convergentes quanto à influência da exposição contínua ao TO sobre a distribuição da resposta. De modo geral, houve uma tendência crescente da escolha pela alternativa sem TO nos experimentos 2, 3 e 4.

No Experimento 1, em contrapartida, observou-se um leve aumento na escolha pela opção sem TO no Bloco 1, seguido por uma redução gradual na proporção dessa resposta ao longo das exposições, aproximando-se dos níveis observados na linha de base. Essa diminuição ao longo dos blocos poderia ser interpretada à luz da hipótese de habituação, entendida como uma redução da responsividade a um estímulo após

exposição repetida (Fontes & Shahan, 2021). No entanto, os estudos destacam que a habituação é mais provável quando o estímulo é constante, de baixa intensidade e apresentado em alta frequência (e.g., Azrin, 1960; McSweeney et al., 1996). No presente experimento, embora o TO tenha sido apresentado sob um esquema VI, ele foi programado com três diferentes durações (5 s, 20 s e 60 s) alternadas entre as sessões, de modo que nenhuma duração específica se repetia em dias consecutivos. Essa variação temporal deveria, em princípio, dificultar o desenvolvimento de habituação tradicional. Ainda assim, a tendência observada ao longo dos blocos foi de aumento gradual da escolha pela opção com TO, o que poderia indicar um processo de adaptação mais complexo.

Nos demais experimentos, por outro lado, as contingências foram mais consistentes: a frequência do TO foi mais alta (Experimento 2: VI 5 s, durações de 5 s e 20 s; Experimento 3: VI 3 s, 5 s e 10 s, com menor variação entre os valores); a duração permaneceu fixa por vários dias consecutivos (Experimento 3: durações de 5 s e 20 s por 15 dias cada; Experimento 4: duração de 20 s); ou o tipo de esquema utilizado favoreceu a regularidade (Experimento 3: cadeias concorrentes). Tais condições deveriam facilitar a habituação, o que, entretanto, não foi observado.

Esses resultados sugerem uma hipótese alternativa relacionada às funções sinalizadoras do TO. No Experimento 1, considerando as grandes variações nas durações do TO (5 s, 20 s e 60 s) e a incerteza quanto à entrega de reforço sob VI, o TO pode ter deixado de ser percebido como estímulo estritamente punitivo, passando a atuar como marcador temporal. Conforme discutido no Capítulo 1 a respeito do Experimento 1 de Toegel et al. (2022), essa situação também pode ser conceituada como um esquema encadeado: a alternativa com TO funcionaria como o elo inicial da

cadeia, enquanto o estímulo de time-in (por exemplo, a luz que sinaliza acesso ao reforço) funcionaria como o elo terminal, indicando a entrega do reforço primário.

Assim, enquanto o estímulo durante o TO indica a ausência de reforço, o estímulo de time-in pode funcionar como reforçador condicionado, diminuindo o atraso percebido até a entrega do reforço. Já a alternativa sem estímulo discriminativo não oferece pistas sobre quando o reforço ocorrerá, deixando o sujeito sem referência temporal, tornando essa opção menos previsível e potencialmente menos atrativa). Esse efeito seria especialmente relevante em contextos de incerteza temporal, como os proporcionados por esquemas VI combinados a TO de durações variadas.

Essa perspectiva sobre o TO pode ser respaldada por teorias do atraso do reforço. Segundo a Teoria da Informação, o valor de um estímulo está relacionado à sua capacidade de orientar o comportamento em direção aos reforços distribuídos no tempo, de modo que a sinalização de reforço (S+) no intervalo entre o tempo total e a apresentação do reforçador (incluindo o tempo da presença S-) pode reduzir a incerteza temporal e, assim, influenciar a taxa e a alocação de respostas.

De forma complementar, o modelo SiGN indica que estímulos que sinalizam a ausência de reforço (S-) têm efeito supressivo limitado (Dunn et al., 2024; McDevitt et al., 2016). Contudo, mesmo quando um S- exerce pouco efeito supressivo em relação a um estímulo de reforço, ele pode ser preferido em contextos em que não há informações claras sobre a ocorrência do reforço (Sears et al., 2022).

Em contraste, os demais experimentos apresentam parâmetros mais estáveis, como mencionado anteriormente, o que pode tornar o contexto mais previsível, favorecendo a função aversiva do TO. Além disso, é importante considerar algumas diferenças metodológicas relevantes entre os experimentos, que podem ter influenciado os efeitos observados. No Experimento 2, os sujeitos já haviam sido expostos ao TO no

Experimento 1, o que levanta a possibilidade de efeitos de reexposição, como sensibilização, ainda que sob parâmetros distintos. Já no Experimento 3, o delineamento adotado introduziu uma diferença estrutural fundamental: foram utilizadas cadeias concorrentes, nas quais a escolha do sujeito era seguida por um período de comprometimento com a alternativa escolhida, impedindo o acesso ao reforço da outra opção durante aquele ciclo. Isso difere dos Experimentos 1 e 2, nos quais, apesar da presença de COD, era possível acessar reforços em ambas as alternativas. Esses fatores devem ser considerados ao comparar os dados entre os experimentos e reforçam a importância de delineamentos mais sistemáticos em estudos futuros, especialmente para isolar os efeitos de cada parâmetro de forma independente.

Também é necessário destacar uma limitação comum a todos os experimentos: a exposição ao TO foi restrita a poucos blocos de sessões em cada condição, o que pode ter afetado a observação de efeitos mais robustos, como reversões de preferência, sensibilização ou adaptação tardia. Estudos futuros com exposição mais prolongada ao TO poderiam lançar luz sobre a estabilidade desses efeitos ao longo do tempo e revelar transições comportamentais mais complexas.

O segundo aspecto observado em relação à alocação da resposta foi, em um nível mais pontual, ao final da exposição do TO. Os resultados dos quatro experimentos aqui relatados revelam que a apresentação do TO sobreposto a uma das opções em esquemas concorrentes influenciou sistematicamente a alocação da resposta, ainda que houvesse variações entre indivíduos e entre os parâmetros manipulados.

De forma geral, os dados indicaram que, com exceção de alguns sujeitos, houve um deslocamento da alocação da resposta em direção à opção sem TO (R1) em todas as manipulações. Esse padrão se manteve independentemente da duração do TO (Experimentos 1, 2 e 3), embora tenha sido mais pronunciado com o aumento da

frequência do TO (Experimento 4). No Experimento 3, o uso de cadeias concorrentes evidenciou também um efeito da presença do TO nas fases terminais, com maior proporção de respostas na cadeia associada à alternativa sem TO. Esses achados podem ser consistentes com uma função punitiva do TO e apontam para a necessidade de modelos teóricos mais sensíveis à distribuição comportamental.

Uma interpretação dos efeitos observados pode ser feita à luz dos modelos quantitativos de escolha com punição propostos por Deluty (1976) e de Villiers (1980), ambos derivados de adaptações da Lei da Igualação (Herrnstein, 1961). Esses modelos oferecem formas distintas de descrever a interação entre reforçamento e punição em contextos concorrentes, e os dados do presente estudo podem se ajustar de maneira diferenciada a cada um, dependendo do parâmetro manipulado.

O modelo de Deluty (1976) propõe que a punição não afeta diretamente o valor da alternativa punida, mas sim aumenta o valor relativo da alternativa não punida. Nesse modelo, quanto maior a taxa de punição associada a uma opção, maior o deslocamento da resposta para a alternativa concorrente. Os Experimentos 2, 3 e 4, nos quais a proporção da resposta aumentou em direção à alternativa sem TO (R1), mesmo sem mudanças na taxa de reforço programada, são compatíveis com essa proposta. O experimento 1 apresenta um padrão mais ambíguo e sutil do que os demais, com três sujeitos preferindo consistentemente a opção com TO, mas, de modo geral, a maioria dos sujeitos também apresentou um aumento da alocação na opção sem TO. A redistribuição do comportamento ocorreu sem supressão absoluta da resposta punida, indicando que o TO pode ter funcionado como um fator de valorização relativa da alternativa não punida.

Já o modelo de de Villiers (1980) assume que punição e reforço operam em um sistema de subtração direta, ou seja, o valor da alternativa punida é reduzido

proporcionalmente à sua taxa de punição. Nesse caso, pequenas variações na taxa de punição podem gerar grandes deslocamentos na alocação da resposta, mesmo sem alterações na taxa de reforço, resultando em padrões de *overmatching*. Essa previsão pode ser observada de maneira mais clara no Experimento 4, no qual o aumento da frequência de apresentação do TO (VI 3s, 5s e 10s) produziu efeitos graduais, mas robustos, na alocação da resposta, sugerindo maior sensibilidade à taxa de punição. Esses achados reforçam a ideia de que, nesse caso, o TO afetou diretamente o valor da alternativa punida, tornando-a progressivamente menos atrativa à medida que sua frequência aumentava.

Apesar da utilidade conceitual desses modelos, é importante reconhecer suas limitações. Ambos foram formulados com base em contingências de punição, utilizando choques elétricos, e assumem que a punição altera diretamente ou indiretamente o valor das alternativas. No entanto, nem sempre é trivial determinar se um evento como o TO funciona como um estímulo punitivo clássico. Além disso, os dois modelos não conseguem fornecer previsões quantitativas precisas sobre os efeitos da punição quando comparados a modelos de igualação que não analisam a punição diretamente, como a GLM (Baum, 1974; Critchfield et al., 2003; Klapes et al., 2018; Fontes & Shahan, 2021).

Estudo recente, como o de Fontes (2022), argumenta que não é necessário incluir a punição de forma explícita para compreender seus efeitos sobre a escolha, basta considerar que ela altera os parâmetros funcionais da escolha, como a sensibilidade relativa ao reforço ou o viés, apresentados na GLM. Embora não tenhamos avaliado formalmente a adequação dos dados do presente estudo à GLM, observou-se que a principal variável responsável pela variação na distribuição das respostas foi a distribuição da taxa de reforços (no Experimento 1, a distribuição relativa de reforço

local e total; no Experimento 4, a distribuição relativa de reforço total), corroborando, em certa medida, as previsões da GLM.

Essa constatação ainda reforça a ideia de que nem todo estímulo que altera a alocação da resposta deve ser interpretado sob a lógica tradicional da punição. No caso do TO, essa discussão é especialmente pertinente. Diferentemente de estímulos classicamente aversivos, como o choque elétrico, o TO não é um estímulo físico direto, mas sim um evento programado que impõe uma interrupção temporária na disponibilidade de reforço, contingente a uma resposta. Assim, em vez de tratá-lo exclusivamente como um estímulo punitivo, é possível compreendê-lo como um evento que altera a estrutura temporal da contingência, prolongando o intervalo até o próximo reforço disponível. Nessa perspectiva, teorias que tratam do valor do reforço em função do tempo, como a DRT e a HVA, podem oferecer ferramentas mais adequadas para explicar os efeitos do TO sobre a distribuição da resposta em contingências concorrentes.

Segundo a DRT, o valor de um estímulo condicionado depende de quanto ele reduz o atraso médio até o reforço. No caso dos experimentos com TO, como esse procedimento acrescenta um período em que o reforço não pode ser obtido, ele prolonga o tempo total até o reforço e, conseqüentemente, reduz o valor da alternativa associada. Assim, mesmo quando as taxas de reforço programadas eram iguais entre as opções, observou-se maior preferência pela alternativa sem TO (R1), que oferecia um caminho mais curto e previsível ao reforço, sem os atrasos adicionais impostos pelo TO.

Na perspectiva da HVA, o valor de uma alternativa é descrito como resultado do desconto hiperbólico em função do atraso. Nesse contexto, o TO, ao introduzir um intervalo sem acesso ao reforço, amplia o atraso efetivo entre a resposta e a próxima oportunidade de reforçamento, diminuindo o valor da opção; segundo, quando

sinalizado, o TO funciona como um indicativo de que não haverá reforço naquele ciclo, o que reduz a probabilidade percebida de reforço. Em ambos os casos, o efeito é a diminuição do valor subjetivo atribuído à alternativa associada ao TO

Embora a DRT e a HVA prevejam, no geral, que a escolha será direcionada para a alternativa com o menor atraso até o reforço, os dados do Experimento 1, nos quais três sujeitos (Y4, Y8 e Y9) preferiram consistentemente a opção com TO, aumentando a preferência à medida que aumentava a duração do TO, indicam que essa previsão pode não se aplicar de maneira uniforme. Uma possível explicação, com base em desdobramentos contemporâneos da DRT, como o modelo SiGN, e na Teoria da Informação, como colocado anteriormente na discussão dos blocos de sessões, é que o TO pode ter funcionado como um estímulo sinalizador de estrutura temporal, em vez de ser percebido apenas como um atraso adicional.

É importante ressaltar que as interpretações apresentadas sobre as teorias do atraso, constituem extrapolações conceituais aplicadas ao presente estudo. Essas teorias foram originalmente desenvolvidas para explicar o comportamento de escolha em procedimentos com cadeias concorrentes, nos quais os estímulos do elo inicial e do terminal têm funções bem definidas como preditores do reforço. No presente trabalho, apenas o Experimento 3 utilizou cadeias concorrentes em sua estrutura, o que limita a aplicação direta dessas abordagens aos demais experimentos.

Além disso, o estudo não foi desenhado com o objetivo de testar as previsões quantitativas específicas desses modelos teóricos, nem foram realizadas manipulações sistemáticas que permitissem avaliar qual modelo explicaria melhor os dados. Diante disso, recomenda-se que futuras pesquisas explorem os efeitos do TO em procedimentos de cadeia concorrente, nos quais se possa isolar com mais precisão os efeitos da sinalização, da previsibilidade e do atraso no reforço. Tais arranjos

experimentais possibilitariam não apenas uma adequação mais rigorosa dos modelos de escolha com base no atraso e na informação, mas também uma análise mais refinada do TO enquanto evento que altera a estrutura temporal da contingência de reforço, ampliando a compreensão de seu impacto sobre a alocação do comportamento.

Outro aspecto relevante observado no presente estudo foi a ausência de diferenças significativas entre as diferentes durações do TO nos Experimentos 1, 2 e 3, um padrão que contrasta com os efeitos claros observados nas manipulações da taxa de apresentação do TO. Os dados do GLMM indicam que, embora a presença do TO influencie a alocação de respostas, aumentar sua duração afetou principalmente a magnitude da taxa de resposta, de modo que durações maiores produziram incrementos menores, sem alterar significativamente a distribuição relativa entre as opções. Por outro lado, mudanças na frequência do TO geraram efeitos consistentes sobre a escolha, aumentando a preferência pela alternativa sem TO à medida que a frequência do TO aumentava.

Além disso, verificou-se que o TO impactou de maneira expressiva a duração total das sessões, sobretudo quando comparado ao tempo efetivo de time-in. Tanto as manipulações de duração quanto as de frequência do TO promoveram aumentos relevantes no tempo total, o que, por sua vez, alterou diretamente as taxas globais de reforço e de TO. As variações percentuais observadas nas taxas de TO e no reforço ao longo dos quatro experimentos, aliadas aos resultados das regressões lineares dos Experimentos 1 e 4, ajudam a compreender por que apenas as manipulações de frequência produziram efeitos significativos sobre a alocação da resposta.

Especificamente, no Experimento 4, em que variou a frequência de apresentação do TO por meio de diferentes VIs, foram observadas diferenças superiores a 30% nas taxas de TO entre as condições, indicando um aumento substancial na interrupção do

acesso ao reforço à medida que o TO se tornava mais frequente. Já nos Experimentos 1 a 3, em que apenas a duração do TO foi manipulada, as variações na taxa de TO foram consideravelmente menores, geralmente abaixo de 10%, sugerindo um impacto funcional mais discreto. Isso inclui o Experimento 2, no qual o TO foi apresentado com uma frequência mais alta (VI 5 s), mas, ainda assim, a variação entre durações de 5 s e 20 s não foi suficiente para produzir diferenças significativas na escolha.

Essa assimetria também se reflete nos resultados das regressões: no Experimento 4, tanto a taxa de TO quanto a distribuição relativa do reforço foram preditores significativos da escolha, explicando parcela substancial da variabilidade na alocação de respostas. Em contraste, no Experimento 1, apenas a distribuição relativa do reforço (no time-in e no tempo total) foi significativa, enquanto as taxas de TO não contribuíram para prever a escolha dos sujeitos. Esses achados indicam que manipulações que afetam diretamente a frequência com que o acesso ao reforço é interrompido tendem a produzir efeitos comportamentais mais pronunciados na alocação da resposta.

Por outro lado, a manipulação da duração do TO mostrou-se determinante para a taxa de respostas. TOs mais curtos (5 s) resultaram em maiores taxas de resposta em comparação com TOs mais longos (20 s e 60 s), sugerindo que a duração do TO modula diretamente a intensidade da resposta, sendo TOs mais curtos associados a um maior número de respostas emitidas durante o time-in. A comparação entre Experimentos 1 e 2 evidencia que a frequência do TO influencia a magnitude desse efeito: quanto mais frequente o TO, como no Experimento 2, menor o aumento na taxa de respostas na alternativa punida, além de maior variabilidade entre sujeitos. Assim, duração e frequência do TO interagem para regular a distribuição das respostas em esquemas concorrentes, favorecendo maior emissão de respostas em TOs curtos e menos frequentes e menor emissão em TOs longos ou mais frequentes.

O mecanismo de atuação do TO pode explicar esse padrão de resultados. Além de afetar diretamente a taxa de respostas, o TO altera indiretamente o valor relativo das alternativas. Em um procedimento concorrente simples, o TO aplicado à R2 reduz temporariamente o acesso ao reforço, funcionando como um TO “não contingente” para R1, uma vez que, durante o período de TO, o reforço não está disponível globalmente. Isso cria assimetria: R1 mantém acesso relativamente contínuo ao reforço, enquanto R2 apresenta acesso probabilístico, tornando a alternativa sem TO mais atrativa e desviando a escolha dos sujeitos. Portanto, a distribuição de respostas parece depender de dois mecanismos principais: (1) a contingência resposta–TO, que impõe um custo específico à alternativa punida; e (2) a taxa global de reforçamento, de modo que reduções acentuadas nessa taxa levam a menor emissão de respostas e maior alocação na alternativa impune. Nesse contexto, a frequência do TO mostrou maior impacto sobre o segundo mecanismo do que a duração, o que contribui para as diferenças observadas.

Esses achados também são consistentes com estudos anteriores que investigaram os efeitos da duração do TO e não encontraram efeitos diferenciais robustos associados a durações mais longas (e.g. Naden & Morgan, 1988; McGuffin, 1991; White et al., 1971). Tais resultados reforçam a hipótese de que a duração do TO exerce um efeito funcional limitado, mesmo quando combinada a taxas mais altas de apresentação, como no Experimento 2. Em contrapartida, o Experimento 4, ao manipular a frequência do TO em esquemas concorrentes, produziu alterações significativas na alocação na resposta, o que também está de acordo com a literatura que manipulou esse parâmetro (e.g. Thomas, 1968; Hall, 2015). Esse último achado revela uma possível interação entre características do TO e variáveis reforçadoras, uma variável que pode representar um aspecto crítico do controle exercido pelo TO sobre o comportamento, como

sugerido por Toegel et al. (2022), que argumentam que os efeitos do TO dependem, em grande parte, do nível de reforçamento disponível no contexto.

Uma explicação alternativa para a ausência de diferenças entre as durações do TO pode ser encontrada nas previsões do modelo HVA, que descreve o desconto hiperbólico do valor do reforço em função do atraso. Quando o TO é introduzido, o atraso total até o reforço aumenta; entretanto, como a função de desconto é hiperbólica, acréscimos sucessivos no atraso produzem reduções cada vez menores no valor do reforço. Assim, o valor subjetivo global não se altera de forma marcante. Esse padrão é consistente com achados experimentais que mostram que durações mais longas de TO não resultam em supressão adicional proporcional (e.g. White et al., 1972; Naden & Morgan, 1988; McGuffin, 1991).

Diante dessa assimetria nos efeitos do TO, torna-se evidente que manipulações isoladas de duração ou frequência fornecem apenas uma visão parcial do controle exercido pelo procedimento sobre o comportamento de escolha. Os resultados sugerem que, para compreender plenamente como o TO influencia a alocação de respostas, é necessário investigar arranjos experimentais que manipulem simultaneamente múltiplas dimensões do TO (como duração, taxa e sinalização) em interação com características do reforço (como taxa e sinalização). Essa abordagem permitiria esclarecer por que, em alguns contextos, o TO produz apenas mudanças na alocação relativa de respostas, enquanto, em outros, também provoca alterações na frequência da resposta-alvo.

Além disso, a complexidade desses efeitos evidencia a utilidade de modelos quantitativos de escolha mais abrangentes, capazes de integrar simultaneamente variáveis relacionadas ao reforço e à punição ou ao atraso, oferecendo uma estrutura conceitual robusta para interpretar os efeitos combinados de diferentes parâmetros do TO sobre o comportamento de escolha.

Conclusão

Os resultados do presente estudo demonstram que o TO, quando programado em contingências concorrentes, influencia sistematicamente a alocação da resposta, ainda que de forma heterogênea entre indivíduos e condições experimentais. De maneira geral, a duração do TO se mostrou consistente na modulação da taxa de respostas, enquanto a frequência do TO revelou-se determinante para o deslocamento da escolha em direção à alternativa sem TO. Tal padrão foi evidenciado tanto nas análises descritivas quanto nas regressões lineares, que apontaram a taxa de TO e a proporção de reforço como preditores mais consistentes da escolha do que a duração do TO isoladamente.

Esses achados indicam que o TO pode operar como um estímulo punitivo, mas sua eficácia depende criticamente do contexto em que é inserido, particularmente da frequência com que interrompe o acesso ao reforço. Além disso, os resultados sugerem que o TO não atua exclusivamente como um evento aversivo tradicional, podendo também exercer funções discriminativas ou moduladoras da estrutura temporal das contingências. Isso se alinha a interpretações baseadas em modelos contemporâneos de escolha, como a DRT, a HVA, a Teria da Informação e a SiGN, que consideram o papel da previsibilidade e do atraso no reforço na determinação da preferência.

Embora amplamente utilizado em contextos aplicados como estratégia de punição, com o objetivo de suprimir comportamentos indesejados, os resultados aqui obtidos demonstram que os efeitos do TO podem ser mais complexos e até contraintuitivos em ambientes complexos com contingências concorrentes. Observou-se, por exemplo, que o TO nem sempre resultou na supressão da resposta-alvo, podendo, inclusive, estar associado ao aumento da taxa de respostas, tanto na alternativa punida quanto na alternativa não punida. Esses achados indicam que a função punitiva

do TO não pode ser assumida a priori, especialmente em contextos mais ricos e variáveis.

Diante disso, torna-se fundamental que futuras investigações considerem arranjos experimentais que capturem essas dinâmicas de forma mais realista, ampliando a compreensão dos efeitos do TO e subsidiando práticas clínicas mais eficazes e fundamentadas. Repensar o TO para além de uma concepção unidimensional de punição e promover uma análise mais integrada entre variáveis reforçadoras e punitivas pode contribuir significativamente para o avanço teórico e prático na Análise do Comportamento.

Referências

- Azrin, N. H. (1960). Effects of punishment intensity during variable-interval reinforcement. *Journal of the Experimental Analysis of Behavior*, 3(2), 123–142. <https://doi.org/10.1901/jeab.1960.3-123>
- Azrin, N. H., & Holz, W. C. (1966). Punishment. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 380–447). Appleton-Century-Crofts.
- Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neurosciences*, 32(2), 73–78. <https://doi.org/10.1016/j.tins.2009.02.001>
- Bateson, M., & Kacelnik, A. (1995). Preferences for fixed and variable food sources: Variability in amount and delay. *Journal of the Experimental Analysis of Behavior*, 63(3), 313–329. <https://doi.org/10.1901/jeab.1995.63-313>
- Baum W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22(1), 231–242. <https://doi.org/10.1901/jeab.1974.22-231>
- Baum W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, 32(2), 269–281. <https://doi.org/10.1901/jeab.1979.32-269>
- Baum, W. M. (1982). Choice, changeover, and travel. *Journal of the Experimental Analysis of Behavior*, 38(1), 35–49. <https://doi.org/10.1901/jeab.1982.38-35>
- Baum, W. M. (2010). Dynamics of choice: A tutorial. *Journal of the Experimental Analysis of Behavior*, 94(2), 161–174. <https://doi.org/10.1901/jeab.2010.94-161>

- Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, *12*, 861-874.
<https://doi.org/10.1901/jeab.1969.12-861>
- Bono, R., Alarcón, R., & Blanca, M. J. (2021). Report quality of generalized linear mixed models in psychology: A systematic review. *Frontiers in psychology*, *12*, 666182. <https://doi.org/10.3389/fpsyg.2021.666182>
- Boyle, M. A., Hoffmann, A. N., & Lambert, J. M. (2018). Behavioral contrast: Research and areas for investigation. *Journal of Applied Behavior Analysis*, *51*(3), 702–718. <https://doi.org/10.1002/jaba.469>
- Branch, M. N. (1977). Signalled and unsignalled percentage reinforcement of performance under a chained schedule. *Journal of the Experimental Analysis of Behavior*, *27*(1), 71–83. <https://doi.org/10.1901/jeab.1977.27-71>
- Carlson, J. G. (1972). Timeout punishment: Rate of reinforcement and delay of timeout. *Learning and Motivation*, *3*(1), 31-43. [https://doi.org/10.1016/0023-9690\(72\)90045-8](https://doi.org/10.1016/0023-9690(72)90045-8)
- Catania, A. C. (1963). Concurrent performances: Reinforcement interaction and response independence. *Journal of the Experimental Analysis of Behavior*, *6*(2), 253–263. <https://doi.org/10.1901/jeab.1963.6-253>
- Catania, A. C. (1999). *Aprendizagem: Comportamento, linguagem e cognição* (4ª ed). Artmed. (Trabalho original publicado em 1998).
- Chow, J. J., Smith, A. P., Wilson, A. G., Zentall, T. R., & Beckmann, J. S. (2017). Suboptimal choice in rats: Incentive salience attribution promotes maladaptive decision-making. *Behavioural Brain Research*, *320*, 244–254.
<https://doi.org/10.1016/j.bbr.2016.12.013>

- Cicerone, R. A. (1976). Preference for mixed versus constant delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 25(2), 257–261.
<https://doi.org/10.1901/jeab.1976.25-257>
- Clark, H. B., Rowbury, T., Baer, A. M., & Baer, D. M. (1973). Timeout as a punishing stimulus in continuous and intermittent schedules. *Journal of Applied Behavior Analysis*, 6(3), 443–455. <https://doi.org/10.1901/jaba.1973.6-443>
- Corralejo, S. M., Jensen, S. A., Greathouse, A. D., & Ward, L. E. (2018). Parameters of time-out: Research update and comparison to parenting programs, books, and online recommendations. *Behavior Therapy*, 49(1), 99–112.
<https://doi.org/10.1016/j.beth.2017.09.005>
- Cotton, J. W. (1953). Running time as a function of amount of food deprivation. *Journal of Experimental Psychology*, 46(3), 188–198.
<https://doi.org/10.1037/h0061014>
- Critchfield, T. S., Paletz, E. M., MacAleese, K. R., & Newland, M. C. (2003). Punishment in human choice: Direct or competitive suppression? *Journal of the Experimental Analysis of Behavior*, 80(1), 1–27.
<https://doi.org/10.1901/jeab.2003.80-1>
- Cunningham, P. J., & Shahan, T. A. (2018). Suboptimal choice, reward-predictive signals, and temporal information. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44(1), 1. <https://doi.org/10.1037/xan0000160>
- Cunningham, P. J., & Shahan, T. A. (2019). Rats engage in suboptimal choice when the delay to food is sufficiently long. *Journal of Experimental Psychology: Animal Learning and Cognition*, 45(3), 301–310. <https://doi.org/10.1037/xan0000211>
- Davison, M., & McCarthy, D. (1988). *The matching law: A research review* (1^a ed.). Routledge. <https://doi.org/10.4324/9781315638911>

- de Villiers, P. A. (1980). Toward a quantitative theory of punishment. *Journal of the Experimental Analysis of Behavior*, 33(1), 15–25.
<https://doi.org/10.1901/jeab.1980.33-15>
- DeFulio, A., & Hackenberg, T. D. (2007). Discriminated timeout avoidance in pigeons: The roles of added stimuli. *Journal of the Experimental Analysis of Behavior*, 88(1), 51–71. <https://doi.org/10.1901/jeab.2007.59-06>
- Deluty, M. Z. (1976). Choice and the rate of punishment in concurrent schedules. *Journal of the Experimental Analysis of Behavior*, 25(1), 75–80. <https://doi.org/10.1901/jeab.1976.25-75>
- Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, 6(4), 693–728. <https://doi.org/10.1017/S0140525X00017969>
- Donaldson, J. M., & Vollmer, T. R. (2011). An evaluation and comparison of time-out procedures with and without release contingencies. *Journal of Applied Behavior Analysis*, 44(4), 693-705. <https://doi.org/10.1901/jaba.2011.44-693>
- Dunn, R. (1990). Timeout from concurrent schedules. *Journal of the Experimental Analysis of Behavior*, 53(1), 163-174. <https://doi.org/10.1901/jeab.1990.53-163>
- Dunn, R. M., Pisklak, J. M., McDevitt, M. A., & Spetch, M. L. (2024). Suboptimal choice: A review and quantification of the signal for good news (SiGN) model. *Psychological Review*, 131(1), 58–78. <https://doi.org/10.1037/rev0000416>
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior*, 53, 201–218. <https://doi.org/10.1901/jeab.1990.53-201>
- Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, 12, 723-730. <https://doi.org/10.1901/jeab.1969.12-723>

- Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313-339). Prentice-Hall.
- Fantino, E., & Romanowich, P. (2007). The effect of conditioned reinforcement rate on choice: A review. *Journal of the Experimental Analysis of Behavior*, *87*(3), 409–421. <https://doi.org/10.1901/jeab.2007.44-06>
- Fantino, E., Freed, D., Preston, R. A., & Williams, W. A. (1991). Choice and conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, *55*(2), 177–188. <https://doi.org/10.1901/jeab.1991.55-177>
- Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior*, *60*(1), 159–169. <https://doi.org/10.1901/jeab.1993.60-159>
- Ferster, C. B. (1953). Sustained behavior under delayed reinforcement. *Journal of Experimental Psychology*, *45*(4), 218–224. <https://doi.org/10.1037/h0062158>
- Ferster, C. B., & Appel, J. B. (1961). Punishment of S Δ responding in matching to sample by time out from positive reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*(1), 45-56. <https://doi.org/10.1901/jeab.1961.4-45>.
- Flory, R. K., & Lickfett, G. G. (1974). Effects of lick-contingent timeout on schedule-induced polydipsia. *Journal of the Experimental Analysis of Behavior*, *21*(1), 45–55. <https://doi.org/10.1901/jeab.1974.21-45>
- Fontes, R. M. (2022). *Punishment and choice* [Doctoral dissertation, Utah State University].
- Fontes, R. M., & Shahan, T. A. (2021). Punishment and its putative fallout: A reappraisal. *Journal of the Experimental Analysis of Behavior*, *115*(1), 185–203. <https://doi.org/10.1002/jeab.653>

- Foxx, R. M., & Shapiro, S. T. (1978). The timeout ribbon: A nonexclusionary timeout procedure. *Journal of Applied Behavior Analysis, 11*(1), 125–136.
<https://doi.org/10.1901/jaba.1978.11-125>
- Gilbert, T. F. (1958). Fundamental dimensional properties of the operant. *Psychological Review, 65*(5), 272–282. <https://doi.org/10.1037/h0044071>
- Grace, R. C. (1996). Choice between fixed and variable delays to reinforcement in the adjusting-delay procedure and concurrent chains. *Journal of Experimental Psychology: Animal Behavior Processes, 22*(3), 362–383.
<https://doi.org/10.1037/0097-7403.22.3.362>.
- Hackenberg, T. D., & Defulio, A. (2007). Timeout from reinforcement: Restoring a balance between analysis and application. *Revista Mexicana de Análisis de la Conducta, 33*(SpecIssue), 37–44.
- Hall, E. G. (2015). *Variable-interval and variable-ratio schedules of punishment by timeout from positive reinforcement*. (Doctoral dissertation. West Virginia University).
- Harris, K. R. (1984). Definitional, Parametric, and Procedural Considerations in Timeout Interventions and Research. *Exceptional Children, 51*(4), 279-288. <https://doi.org/10.1177/001440298405100403>
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*(3), 267–272. <https://doi.org/10.1901/jeab.1961.4-267>
- Herrnstein, R. J. (1964). Aperiodicity as a factor in choice. *Journal of the Experimental Analysis of Behavior, 7*(2), 179–182. <https://doi.org/10.1901/jeab.1964.7-179>
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266. <https://doi.org/10.1901/jeab.1970.13-243>

- Holtyn, A. F. (2012). *Avoidance of timeout from response-independent schedules of food or sucrose water with and without a limited hold on consumption*. (Doctoral dissertation, West Virginia University).
- Hunziker, M. H. L. (2011). Afinal, o que é controle aversivo? *Acta Comportamentalia*, 19(4), 9-19. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0188-81452011000400006&lng=pt&tlng=pt.
- Hunziker, M. H. L. (2017). Dogmas sobre o controle aversivo. *Acta Comportamentalia*, 25(1), 85-100. Recuperado de <https://www.revistas.unam.mx/index.php/acom/article/view/58803>
- Ishii, T., & Sakagami, T. (2007). Reinforcement omission in concurrent fixed-interval and random-interval schedules. *Behavioural Processes*, 74(3), 334–341. <https://doi.org/10.1016/j.beproc.2006.12.004>
- Jackson, J. L., & Calhoun, K. S. (1977). Effects of two variable-ratio schedules of timeout: Changes in target and non-target behaviors. *Journal of Behavior Therapy and Experimental Psychiatry*, 8(2), 195-199.
- Kaufman, A., & Baron, A. (1968). Suppression of behavior by timeout punishment when suppression results in loss of positive reinforcement. *Journal of the Experimental Analysis of Behavior*, 11(5), 595-607. <https://doi.org/10.1901/jeab.1968.11-595>
- Kelley, M. E., Lerman, D. C., Fisher, W. W., Roane, H. S., & Zangrillo, A. N. (2011). Reinforcement delay fading during differential reinforcement of communication: The effects of signals on response maintenance. *Journal of the Experimental Analysis of Behavior*, 96(1), 107–122. <https://doi.org/10.1901/jeab.2011.96-107>

- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, *21*, 463–473.
<https://doi.org/10.1901/jeab.1974.21-463>
- Killeen, P. R. (2014). A theory of behavioral contrast. *Journal of the Experimental Analysis of Behavior*, *102*(3), 363–390. <https://doi.org/10.1002/jeab.107>
- Killeen, P. R., & Fantino, E. (1990). Unification of models for choice between delayed reinforcers. *Journal of the Experimental Analysis of Behavior*, *53*, 189–200.
<https://doi.org/10.1901/jeab.1990.53-189>
- Klapes, B., Riley, S., & McDowell, J. J. (2018). Toward a contemporary quantitative model of punishment. *Journal of the Experimental Analysis of Behavior*, *109*(2), 336–348. <https://doi.org/10.1002/jeab.317>
- Lalli, J. S., Mauro, B. C., & Mace, F. C. (2000). Preference for unreliable reinforcement in children with mental retardation: The role of conditioned reinforcement. *Journal of Applied Behavior Analysis*, *33*(4), 533–544.
<https://doi.org/10.1901/jaba.2000.33-533>
- Lattal K. A. (1984). Signal functions in delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, *42*(2), 239–253.
<https://doi.org/10.1901/jeab.1984.42-239>
- Lattal, K. A. (2010). Delayed reinforcement of operant behavior. *Journal of the Experimental Analysis of Behavior*, *93*(1), 129–139.
<https://doi.org/10.1901/jeab.2010.93-129>
- Leitenberg, H. (1965). Is time-out from positive reinforcement an aversive event? A review of the experimental evidence. *Psychological Bulletin*, *64*, 428–441.
<https://doi.org/10.1037/h0022657>

- Lerman, D. C., Iwata, B. A., Shore, B. A., & DeLeon, I. G. (1997). Effects of intermittent punishment on self-injurious behavior: An evaluation of schedule thinning. *Journal of Applied Behavior Analysis*, *30*(2), 187-201
- Lydon, S., Healy, O., Morán, L., & Foody, C. (2015). A quantitative examination of punishment research. *Research in Developmental Disabilities*, *36C*, 470-484. <https://doi.org/10.1016/j.ridd.2014.10.036>
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*(4), 426–436. <https://doi.org/10.1037/0097-7403.10.4.426>
- Mazur, J. E. (1986). Choice between single and multiple delayed reinforcers. *Journal of the Experimental Analysis of Behavior*, *46*(1), 67–77. <https://doi.org/10.1901/jeab.1986.46-67>
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *The effect of delay and of intervening events on reinforcement value* (pp. 55–73). Lawrence Erlbaum Associates, Inc.
- Mazur, J. E. (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior*, *51*(1), 87–99. <https://doi.org/10.1901/jeab.1989.51-87>
- Mazur, J. E. (1991). Choice with probabilistic reinforcement: Effects of delay and conditioned reinforcers. *Journal of the Experimental Analysis of Behavior*, *55*(1), 63–77. <https://doi.org/10.1901/jeab.1991.55-63>
- Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, *108*(1), 96–112.
- Mazur, J. E. (2006). *Learning and behavior* (6th ed.). Prentice Hall.

- Mazur, J. E., Snyderman, M., & Coe, D. (1985). Influences of delay and rate of reinforcement on discrete-trial choice. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(4), 565–575.
- McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, *105*(1), 23–40. <https://doi.org/10.1002/jeab.192>
- McDevitt, M. A., Pisklak, J. M., Dunn, R. M., & Spetch, M. L. (2024). Temporal context effects on suboptimal choice. *Psychonomic Bulletin & Review*, *31*(6), 2737-2745. <https://doi.org/10.3758/s13423-024-02519-y>
- McDiarmid, C. G., & Rilling, M. E. (1965). Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychonomic Science*, *2*(11), 195–196. <https://doi.org/10.3758/BF03343402>
- McDowell, J. J., Calvin, O. L., & Klapes, B. (2016). A survey of residual analysis and a new test of residual trend. *Journal of the Experimental Analysis of Behavior*, *105*(3), 445–458. <https://doi.org/10.1002/jeab.208>
- McDowell, J.J. (1989). Two modern developments in matching theory. *The Behavior Analyst*, *12*, 153-166. <https://doi.org/10.1007/BF03392492>
- McGuffin, P. W. (1991). The effect of timeout duration on frequency of aggression in hospitalized children with conduct disorders. *Behavioral Interventions*, *6*(4), 279-288.
- McMillan, D. E. (1967). A comparison of the punishing effects of response-produced shock and response-produced time out. *Journal of the Experimental Analysis of Behavior*, *10*(5), 439–449. <https://doi.org/10.1901/jeab.1967.10-439>

- McSweeney, F. K., Hinson, J. M., & Cannon, C. B. (1996). Sensitization–habituation may occur during operant conditioning. *Psychological Bulletin*, *120*(2), 256–271. <https://doi.org/10.1037/0033-2909.120.2.256>
- Molet, M., Miller, H. C., Laude, J. R., Kirk, C., Manning, B., & Zentall, T. R. (2012). Decision making by humans in a behavioral task: Do humans, like pigeons, show suboptimal choice? *Learning & Behavior*, *40*(4), 439–447. <https://doi.org/10.3758/s13420-012-0065-7>
- Nader, M. A., & Morgan, D. (2001). Effects of negative punishment contingencies on cocaine self-administration by rhesus monkeys. *Behavioural Pharmacology*, *12*(2), 91-99.
- O'Daly, M., & Fantino, E. (2003). Delay reduction theory: Choice, value, and conditioned reinforcement. *The Behavior Analyst Today*, *4*(2), 141–150. <https://doi.org/10.1037/h0100116>
- Ono, K., Koshitsuka, Y., & Kubo, N. (2021). The effects of warned loss magnitude and timeout duration on human avoidance behavior. *Journal of the Experimental Analysis of Behavior*, *115*(2), 481–494. <https://doi.org/10.1002/jeab.658>
- Parker, B.K., Barker, D.L. & Topping, J. (1971). The differential effects of timeout and shock on DRL responding in the CER paradigm. *Psychonomic Science*, *22*, 133-135. <https://doi.org/10.3758/BF03332533>
- Pierce, W. D., & Cheney, C. D. (2013). *Behavior analysis and learning* (5th ed.). Psychology Press. <https://doi.org/10.4324/9780203441817>
- Pietras, C. J., & Hackenberg, T. D. (2000). Timeout postponement without increased reinforcement frequency. *Journal of the Experimental Analysis of Behavior*, *74*, 147-164. <https://doi.org/10.1901/jeab.2000.74-147>

- Rachlin, H. (1973). Contrast and matching. *Psychological Review*, 80(3), 217–234.
<https://doi.org/10.1037/h0034325>
- Rachlin, H., & Baum, W. M. (1969). Response rate as a function of amount of reinforcement for a signalled concurrent response. *Journal of the Experimental Analysis of Behavior*, 12(1), 11–16. <https://doi.org/10.1901/jeab.1969.12-11>
- Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice. *Psychological Review*, 93(1), 33–58.
<https://doi.org/10.1037/0033-295X.93.1.33>
- Reed, D. D., & Kaplan, B. A. (2011). The matching law: a tutorial for practitioners. *Behavior Analysis in Practice*, 4(2), 15-24.
<https://doi.org/10.1007/BF03391780>
- Reynolds, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior*, 4(1), 57–71. <https://doi.org/10.1901/jeab.1961.4-57>
- Richards, R. W. (1981). A comparison of signaled and unsignaled delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 35, 145–152.
<https://doi.org/10.1901/jeab.1981.35-145>
- Sears, B., Dunn, R. M., Pisklak, J. M., Spetch, M. L., & McDevitt, M. A. (2022). Good news is better than bad news, but bad news is not worse than no news. *Learning & Behavior*, 50, 482–493. <https://doi.org/10.3758/s13420-021-00489-y>
- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, 93(2), 269–289.
<https://doi.org/10.1901/jeab.2010.93-269>
- Shahan, T. A., & Cunningham, P. (2015). Conditioned reinforcement and information theory reconsidered. *Journal of the Experimental Analysis of Behavior*, 103(2), 405–418. <https://doi.org/10.1002/jeab.142>

- Sidman, M. (1989). *Coercion and its fallout*. Boston: Authors Cooperative.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*(2), 212–220. <https://doi.org/10.1037/xan0000092>
- Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior*, *15*(1), 27–38. <https://doi.org/10.1901/jeab.1971.15-27>
- Teixeira, L. M. T. (2023). *Contraste comportamental positivo é função da redução da taxa de reforços ou da taxa de respostas?* [Dissertação de mestrado, Universidade Estadual de Londrina].
- Thomas, J. R. (1965). Discriminated time-out avoidance in pigeons. *Journal of the Experimental Analysis of Behavior*, *8*(5), 329–338. <https://doi.org/10.1901/jeab.1965.8-329>
- Thomas, J. R. (1968). Fixed-ratio punishment by timeout of concurrent variable-interval behavior. *Journal of the Experimental Analysis of Behavior*, *11*(5), 609–616. <https://doi.org/10.1901/jeab.1968.11-609>
- Toegel, C., Holtyn, A. F., Toegel, F., & Perone, M. (2022). The aversiveness of timeout from response-dependent and response-independent food deliveries as a function of delivery rate. *Journal of the experimental analysis of behavior*, *117*(2), 201–239. <https://doi.org/10.1002/jeab.742>
- van Haaren, F., & Anderson, K. G. (1998). Avoidance of time-out from response-independent food presentation: Effects of chlordiazepoxide and buspirone.

Pharmacology Biochemistry and Behavior, 61, 207-214.

[https://doi.org/10.1016/s0091-3057\(98\)00101-4](https://doi.org/10.1016/s0091-3057(98)00101-4)

White, G. D., Nielsen, G., & Johnson, S. M. (1972). Timeout duration and the suppression of deviant behavior in children. *Journal of Applied Behavior Analysis*, 5(2), 111–120. <https://doi.org/10.1901/jaba.1972.5-111>

Williams, B. A. (1994). Conditioned reinforcement: Experimental and theoretical issues. *The Behavior Analyst*, 17(2), 261–285.

<https://doi.org/10.1007/BF03392675>

Wolf, T. L., McLaughlin, T. F., & Williams, R. L. (2006). Time-out interventions and strategies: A brief review and recommendations. *International Journal of Special Education*, 21(3), 22–29.

Zentall, T. R. (2014). Suboptimal choice by pigeons: An analog of human gambling behavior. *Behavioural processes*, 103, 156-164.

<https://doi.org/10.1016/j.beproc.2013.11.004>

Zimmerman, J., & Ferster, C. B. (1963). Intermittent punishment of SΔ responding in matching to sample. *Journal of the Experimental Analysis of Behavior*, 6(3), 349–356. <https://doi.org/10.1901/jeab.1963.6-349>

Apêndice A

Taxa e Proporção da Resposta Durante as Fases Experimentais do Experimento 1

		Sujeitos								
		Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
		FASE 1 (Linha de base)								
	Ses	26	25	28	26	27	25	25	25	25
LB1	R1	29,48 (1,24)	35,38 (2,19)	16,91 (3,62)	26,87 (4,22)	40,37 (7,21)	24,58 (1,95)	42,02 (4,28)	24,86 (4,25)	26,81 (0,85)
	R2	30,75 (2,97)	26,72 (2,84)	20,91 (4,19)	22,39 (1,13)	45,14 (8,32)	24,68 (1,61)	32,78 (4,19)	26,63 (3,42)	34,79 (4,14)
	Prop	0,96 (0,08)	1,34 (0,18)	0,81 (0,09)	1,20 (0,19)	0,89 (0,07)	0,99 (0,08)	1,29 (0,16)	0,93 (0,13)	0,78 (0,11)
		FASE 2 (TO)								
		TO 5s								
Bloco 1	R1	52,61 (10,66)	35,74 (3,39)	30,30 (11,00)	51,78 (13,72)	63,42 (9,08)	39,19 (9,80)	56,50 (6,74)	36,13 (6,47)	35,42 (3,66)
	~	50,14 (4,65)	30,34 (2,29)	28,67 (10,23)	29,30 (5,51)	54,50 (9,56)	35,94 (8,39)	42,29 (7,48)	47,33 (8,88)	48,78 (12,83)
	Prop	1,04 (0,16)	1,18 (0,13)	1,05 (0,20)	1,76 (0,41)	1,19 (0,25)	1,08 (0,09)	1,34 (0,12)	0,78 (0,22)	0,76 (0,18)
Bloco 2	R1	53,73 (4,06)	33,72 (4,18)	39,46 (7,41)	48,72 (9,79)	63,38 (5,26)	56,25 (11,70)	73,48 (11,68)	39,83 (7,14)	33,06 (4,21)
	~	47,35 (5,54)	31,75 (5,67)	45,48 (4,50)	36,01 (1,89)	65,76 (2,64)	33,02 (2,57)	52,48 (5,63)	46,10 (5,21)	70,75 (3,08)
	Prop	1,13 (0,05)	1,08 (0,22)	0,86 (0,09)	1,35 (0,29)	0,96 (0,10)	1,69 (0,27)	1,42 (0,32)	0,86 (0,12)	0,46 (0,06)
Bloco 3	R1	52,42 (1,09)	39,66 (2,27)	38,22 (2,83)	39,72 (3,47)	65,18 (5,61)	71,26 (3,92)	64,32 (6,72)	45,56 (8,88)	42,98 (2,67)
	~	47,89 (4,09)	28,82 (2,22)	50,88 (3,36)	41,25 (2,34)	64,82 (6,34)	35,46 (3,21)	52,28 (6,88)	50,99 (6,12)	57,92 (8,13)
	Prop	1,10 (0,09)	1,38 (0,16)	0,75 (0,07)	0,96 (0,10)	1,01 (0,13)	2,02 (0,25)	1,24 (0,23)	0,88 (0,09)	0,75 (0,15)
		TO 20s								
Bloco 1	~	39,26 (7,23)	37,94 (7,23)	34,03 (5,62)	48,94 (10,28)	58,18 (5,97)	37,70 (5,58)	57,74 (8,95)	33,96 (3,52)	33,83 (5,27)

	R2	34,65 (3,85)	24,70 (3,85)	29,86 (5,56)	28,19 (3,68)	48,08 (6,69)	37,74 (7,69)	37,12 (8,67)	44,14 (5,55)	37,93 (9,90)
	Prop	1,12 (0,43)	1,56 (0,42)	1,14 (0,06)	1,76 (0,51)	1,22 (0,14)	1,04 (0,31)	1,70 (0,82)	0,77 (0,11)	0,96 (0,34)
Bloco 2	R1	48,21 (3,90)	29,37 (3,90)	36,52 (3,78)	36,90 (3,55)	61,85 (3,40)	46,86 (9,10)	59,64 (8,92)	41,54 (3,56)	24,18 (1,86)
	R2	45,09 (3,30)	26,08 (3,20)	41,56 (5,15)	33,48 (3,10)	60,52 (4,43)	31,65 (3,30)	50,54 (4,35)	50,57 (3,96)	63,70 (9,32)
	Prop	1,09 (0,30)	1,14 (0,25)	0,88 (0,14)	1,10 (0,08)	1,02 (0,07)	1,14 (0,26)	1,19 (0,24)	0,82 (0,10)	0,38 (0,08)
Bloco 3	R1	44,14 (7,48)	28,64 (7,48)	38,94 (2,06)	35,84 (5,42)	62,51 (5,15)	59,61 (6,12)	67,42 (6,03)	44,82 (2,71)	31,56 (4,05)
	R2	41,25 (5,07)	21,93 (5,07)	44,13 (2,53)	41,57 (3,55)	58,74 (3,99)	32,82 (2,52)	50,27 (3,01)	63,66 (5,14)	65,36 (5,50)
	Prop	1,07 (0,11)	1,30 (0,19)	0,88 (0,04)	0,87 (0,18)	1,06 (0,05)	1,81 (0,16)	1,34 (0,11)	0,70 (0,08)	0,48 (0,09)

TO 60s

Bloco 1	R1	37,99 (10,60)	33,25 (5,67)	20,26 (4,29)	46,47 (8,17)	47,08 (9,45)	35,59 (8,15)	48,98 (8,50)	35,66 (1,52)	23,64 (4,98)
	R2	34,58 (3,10)	22,35 (3,36)	21,03 (4,17)	24,48 (3,37)	40,51 (4,45)	30,15 (4,46)	28,30 (6,03)	38,60 (2,87)	35,62 (4,33)
	Prop	1,10 (0,35)	1,51 (0,35)	0,96 (0,14)	1,95 (0,58)	1,15 (0,16)	1,17 (0,15)	1,80 (0,58)	0,92 (0,07)	0,67 (0,18)
Bloco 2	R1	46,81 (4,97)	24,97 (1,28)	30,58 (6,68)	38,70 (5,60)	51,27 (6,93)	39,53 (9,50)	49,24 (6,26)	36,08 (8,41)	20,40 (2,23)
	R2	38,62 (4,19)	22,26 (4,43)	32,96 (6,44)	29,30 (2,04)	50,26 (4,73)	31,36 (1,98)	42,96 (0,84)	41,24 (11,69)	58,62 (5,36)
	Prop	1,23 (0,24)	1,16 (0,25)	0,93 (0,19)	1,32 (0,24)	1,02 (0,12)	1,25 (0,28)	1,14 (0,13)	0,88 (0,07)	0,35 (0,05)
Bloco 3	R1	38,15 (3,20)	30,75 (3,63)	33,78 (1,44)	32,37 (6,79)	53,92 (2,63)	56,16 (5,13)	50,84 (4,73)	39,68 (3,96)	23,91 (4,79)
	R2	32,76 (7,07)	20,36 (3,42)	40,22 (3,68)	36,96 (3,91)	50,60 (3,85)	29,17 (3,24)	44,36 (4,13)	63,61 (6,80)	56,17 (6,93)
	Prop	1,20 (0,26)	1,52 (0,11)	0,84 (0,08)	0,88 (0,20)	1,06 (0,07)	1,95 (0,37)	1,15 (0,18)	0,62 (0,06)	0,43 (0,10)

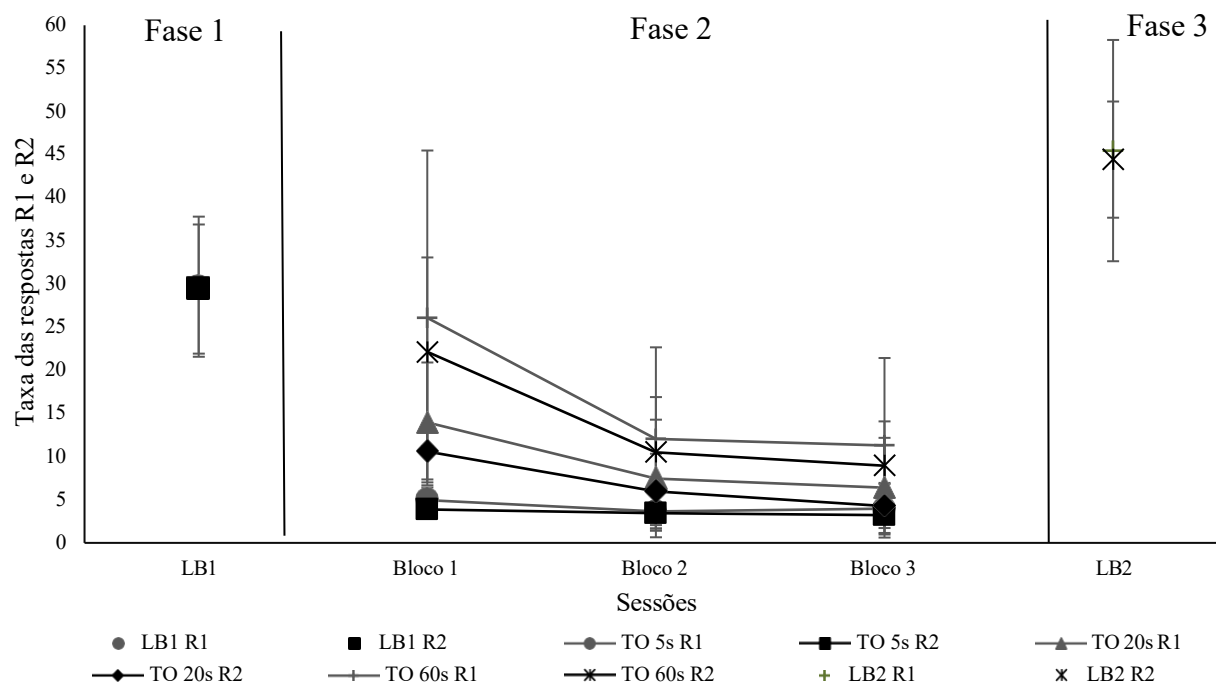
**FASE 3
(Linha de base)**

	Ses	19	18	15	17	15	17	18	18	18
LB2	R1	49,69 (6,06)	45,69 (7,73)	29,41 (2,61)	25,52 (4,55)	61,63 (7,44)	58,97 (3,51)	56,10 (2,27)	36,23 (3,85)	46,03 (2,93)
	R2	44,60 (5,99)	42,52 (3,36)	46,92 (7,58)	43,38 (5,83)	56,08 (4,78)	31,17 (3,20)	46,45 (5,31)	40,11 (4,54)	48,63 (2,41)
	Prop	1,12 (0,02)	1,07 (0,05)	0,64 (0,11)	0,58 (0,05)	1,09 (0,07)	1,9 (0,23)	1,21 (0,12)	0,91 (0,14)	0,94 (0,08)

Nota: A tabela 5 apresenta os dados do Experimento 1. Dentre eles estão descritos: total de sessões (Ses) realizadas durante a Fase 1 e 3, a média da taxa de resposta das últimas seis sessões para a opção 1 (R1) e opção 2 (R2) e a proporção da resposta (Prop) da Fase 1 e 3; e a média das taxas de respostas e proporção dos blocos de sessões da Fase 2. Os números entre parênteses representam o desvio padrão de cada média.

Apêndice B

Média Geral da Taxa de Resposta das Opções R1 e R2 Durante o TO na Fase 2 do Experimento 1



Nota: A figura 13 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e Fase 3 e dos blocos de sessões da Fase 2 durante o período de TO do Experimento 1. O TO foi apresentado contingente apenas a opção R2 da Fase 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Apêndice C

Taxa e Proporção da Resposta Durante as Fases Experimentais do Experimento 2

		Sujeitos			
		Y1	Y7	Y8	Y9
FASE 1 (Linha de base)					
	Ses	19	18	18	18
I R1	R1	49,69 (6,06)	56,10 (2,27)	36,23 (3,85)	46,03 (2,93)
	R2	44,60 (5,99)	46,45 (5,31)	40,11 (4,54)	48,63 (2,41)
	Pron	1,12 (0,02)	1,21 (0,12)	0,91 (0,14)	0,94 (0,08)
FASE 2 (TO)					
TO 5s					
Bloco 1	R1	58,62 (11,14)	70,32 (8,70)	43,16 (4,86)	63,00 (5,79)
	R2	40,15 (6,85)	37,10 (4,30)	59,57 (4,27)	39,40 (7,46)
	Pron	1,47 (0,29)	1,91 (0,30)	0,72 (0,04)	1,62 (0,16)
Bloco 2	R1	60,27 (3,11)	61,20 (19,53)	40,87 (4,00)	67,05 (0,96)
	R2	48,89 (4,33)	28,47 (8,85)	57,57 (5,85)	26,98 (4,13)
	Pron	1,24 (0,12)	2,14 (0,14)	0,71 (0,08)	2,53 (0,42)
Bloco 3	R1	65,83 (7,53)	71,37 (5,90)	45,74 (3,37)	68,89 (5,12)
	R2	47,17 (4,54)	38,69 (2,98)	60,12 (9,15)	33,86 (1,24)
	Pron	1,41 (0,30)	1,85 (0,17)	0,77 (0,16)	2,03 (0,14)
TO 20s					
Bloco 1	R1	52,48 (11,16)	66,27 (4,00)	41,14 (3,71)	40,13 (11,08)

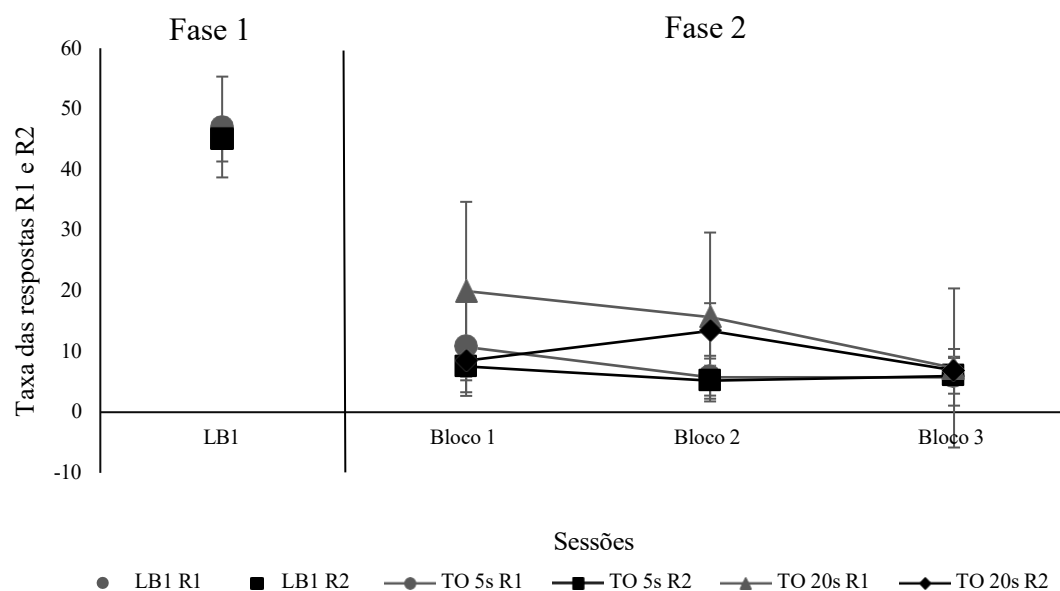
	R2	33,87 (5,07)	25,97 (1,03)	59,13 (3,21)	31,87 (11,97)
	Prop	1,61 (0,55)	2,55 (0,25)	0,69 (0,08)	1,32 (0,24)
Bloco 2	R1	54,60 (6,85)	68,16 (7,42)	42,10 (4,31)	58,51 (7,60)
	R2	35,72 (1,66)	30,13 (3,24)	57,15 (1,08)	27,63 (2,76)
	Prop	1,52 (0,17)	2,28 (0,37)	0,73 (0,06)	2,14 (0,46)
Bloco 3	R1	55,49 (7,15)	74,90 (4,55)	40,23 (2,37)	60,12 (1,05)
	R2	39,92 (4,43)	29,87 (6,24)	55,68 (3,99)	26,90 (7,37)
	Prop	1,39 (0,16)	2,60 (0,66)	0,72 (0,05)	2,38 (0,75)

Nota: A tabela 6 apresenta os dados do Experimento 2. Dentre eles estão descritos:

total de sessões (Ses) realizadas durante a Fase 1, a média da taxa de resposta das últimas seis sessões para a opção 1 (R1) e opção 2 (R2) e a proporção da resposta (Prop) da Fase 1; e a média das taxas de respostas e proporção dos blocos de sessões da Fase 2. Os números entre parênteses representam o desvio padrão de cada média.

Apêndice D

Média Geral da Taxa de Resposta das Opções R1 e R2 Durante o TO na Fase 2 do Experimento 2



Nota: A figura 14 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e dos blocos de sessões da Fase 2 durante o período de TO do Experimento 2. O TO foi apresentado contingente apenas a opção R2 da Fase 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.

Apêndice E

Taxa e Proporção da Resposta Durante as Fases Experimentais do Experimento 3

		Sujeitos								
		B1	B2	B3	B4	B5	B6	B7	B8	
		FASE 1 (Linha de base)								
LB1	R1	4,94 (0,69)	4,21 (1,43)	7,09 (1,87)	6,97 (2,53)	8,10 (2,62)	6,95 (0,90)	7,96 (0,82)	8,39 (1,48)	
	R2	5,02 (1,72)	5,25 (1,07)	8,25 (0,57)	10,82 (1,87)	14,35 (4,32)	7,28 (1,40)	11,34 (1,90)	12,28 (1,84)	
	Pron	0,63 (0,12)	0,83 (0,33)	0,85 (0,23)	0,65 (0,27)	0,62 (0,31)	0,97 (0,17)	0,71 (0,14)	0,69 (0,16)	
		FASE 2 (TO)								
		TO	20s	5s	20s	5s	5s	20s	5s	20s
Bloco 1	R1	7,96 (1,91)	8,96 (1,72)	8,70 (2,19)	12,25 (3,77)	12,17 (2,56)	6,93 (1,50)	12,08 (2,23)	11,08 (2,13)	
	R2	7,91 (1,53)	5,41 (2,19)	7,54 (3,08)	10,60 (2,13)	12,95 (0,73)	10,10 (1,89)	13,07 (1,57)	8,38 (1,02)	
	Pron	1,02 (0,29)	1,87 (0,73)	1,23 (0,36)	1,14 (0,18)	0,94 (0,23)	0,70 (0,17)	0,94 (0,23)	1,32 (0,26)	
Bloco 2	R1	10,11 (1,26)	7,89 (1,70)	7,86 (1,04)	6,01 (4,18)	14,25 (3,37)	7,57 (1,56)	9,80 (3,27)	6,10 (2,86)	
	R2	7,01 (0,37)	5,94 (1,45)	5,70 (0,47)	5,77 (3,62)	8,89 (1,32)	8,76 (0,64)	10,31 (1,35)	6,42 (1,91)	
	Pron	1,45 (0,26)	1,35 (0,23)	1,38 (0,22)	1,04 (0,18)	1,62 (0,41)	0,86 (0,16)	0,93 (0,25)	0,90 (0,28)	
Bloco 3	R1	10,98 (2,64)	7,10 (0,46)	6,66 (1,59)	9,04 (3,77)	11,64 (3,34)	10,52 (2,70)	13,74 (3,44)	8,24 (1,65)	
	R2	8,55 (1,25)	7,36 (1,60)	5,66 (0,51)	8,43 (3,56)	11,43 (2,37)	8,18 (2,49)	13,60 (3,60)	6,77 (1,20)	
	Pron	1,33 (0,36)	1,01 (0,28)	1,17 (0,26)	1,11 (0,24)	1,04 (0,29)	1,35 (0,45)	1,02 (0,15)	1,26 (0,43)	
		FASE 3 (Linha de base)								
LB2	R1	10,03 (1,91)	6,67 (1,61)	8,42 (2,17)	9,88 (1,01)	10,90 (1,51)	8,44 (1,73)	9,30 (1,42)	9,81 (1,20)	
	R2	10,67 (1,21)	9,93 (1,30)	11,47 (2,03)	12,90 (2,31)	15,89 (2,78)	15,86 (3,96)	12,96 (1,96)	8,91 (1,44)	

		Prop	0,93 (0,11)	0,66 (0,13)	0,76 (0,29)	0,77 (0,09)	0,69 (0,03)	0,55 (0,15)	0,72 (0,14)	1,13 (0,27)
FASE 4 (TO)										
		TO	5s	20s	5s	20s	20s	5s	20s	5s
Bloco 1	R1	11,26 (3,25)	7,04 (2,32)	6,87 (1,41)	10,22 (0,82)	13,62 (3,39)	10,29 (2,00)	7,24 (1,08)	8,21 (1,58)	
	R2	9,22 (2,03)	7,47 (0,71)	10,24 (0,39)	11,29 (1,61)	14,41 (2,30)	11,33 (1,73)	11,25 (2,56)	10,11 (1,44)	
	Pron	1,30 (0,60)	0,98 (0,39)	0,67 (0,14)	0,92 (0,20)	0,93 (0,14)	0,94 (0,30)	0,67 (0,17)	0,80 (0,06)	
Bloco 2	R1	12,54 (2,77)	6,69 (0,93)	7,75 (1,90)	8,74 (1,60)	13,61 (1,45)	13,54 (2,24)	9,67 (0,40)	11,78 (1,87)	
	R2	9,04 (1,89)	5,87 (1,12)	10,94 (0,81)	11,98 (1,04)	13,48 (1,42)	11,14 (2,96)	9,93 (1,51)	8,11 (1,32)	
	Pron	1,30 (0,21)	1,15 (0,16)	0,71 (0,19)	0,72 (0,09)	1,01 (0,08)	1,34 (0,62)	0,99 (0,15)	1,46 (0,41)	
Bloco 3	R1	11,16 (2,00)	7,15 (1,01)	5,58 (0,78)	8,96 (1,84)	13,61 (1,33)	11,73 (1,49)	11,48 (2,14)	14,13 (1,47)	
	R2	8,21 (0,90)	4,57 (0,60)	8,23 (0,80)	8,95 (0,90)	12,29 (1,66)	10,98 (1,92)	9,28 (0,88)	8,60 (0,81)	
	Pron	1,32 (0,29)	1,58 (0,34)	0,67 (0,04)	1,00 (0,23)	1,11 (0,11)	1,07 (0,08)	1,25 (0,33)	1,65 (0,18)	

Nota: A tabela apresenta os dados do Experimento 3. Dentre eles estão descritos: a

média da taxa de resposta das últimas seis sessões para a opção 1 (R1) e opção 2 (R2) e a proporção da resposta (Prop) dos elos iniciais das Fase 1 e Fase 3; e a média das taxas de respostas, e proporção dos elos iniciais dos blocos de sessões das Fase 2 e Fase 4. Os números entre parênteses representam o desvio padrão de cada média.

Apêndice F

Taxa e Proporção da Resposta Durante as Fases Experimentais do Experimento 4

		Sujeitos							
		A1	A2	A3	A5	A6	A7	A8	A9
		FASE 1 (Linha de base)							
	Ses	23	19	19	19	21	23	21	19
T D1	D1	32,83 (4,08)	32,98 (5,18)	18,02 (3,62)	21,74 (3,03)	33,83 (1,51)	23,00 (4,38)	29,69 (2,74)	32,42 (6,10)
	D2	29,93 (5,14)	22,33 (2,95)	23,83 (3,64)	25,26 (2,60)	32,06 (2,51)	36,19 (2,49)	44,31 (5,17)	23,67 (4,67)
	Pron	1,11 (0,14)	1,48 (0,26)	0,76 (0,15)	0,86 (0,10)	1,06 (0,11)	0,63 (0,10)	0,67 (0,07)	1,37 (0,10)
		FASE 2 (TO)							
		TO VI 3s							
Bloco 1	D1	55,21 (12,07)	34,12 (8,00)	28,15 (7,25)	35,93 (8,31)	45,70 (10,24)	29,43 (10,48)	40,52 (11,44)	59,19 (19,44)
	D2	39,77 (12,53)	27,85 (4,59)	35,48 (10,10)	19,81 (3,47)	23,29 (4,04)	26,06 (5,41)	35,65 (6,54)	25,50 (4,42)
	Pron	1,70 (1,31)	1,22 (0,14)	0,81 (0,12)	1,80 (0,24)	2,07 (0,87)	1,99 (0,57)	1,19 (0,54)	2,29 (0,62)
Bloco 2	D1	78,41 (3,05)	54,38 (6,88)	54,84 (4,51)	51,10 (6,43)	44,09 (6,17)	61,71 (10,10)	50,28 (4,06)	75,78 (8,02)
	D2	19,62 (5,21)	28,89 (2,78)	30,29 (2,89)	28,50 (3,44)	21,05 (4,69)	24,30 (3,73)	25,35 (3,25)	36,96 (1,90)
	Pron	4,34 (1,64)	1,90 (0,35)	1,82 (0,21)	1,83 (0,45)	2,20 (0,73)	2,57 (0,52)	1,99 (0,20)	2,06 (0,32)
Bloco 3	D1	77,64 (6,29)	69,30 (3,60)	61,49 (8,16)	48,23 (2,99)	53,44 (5,90)	67,51 (4,06)	45,00 (5,56)	67,65 (6,47)
	D2	26,82 (8,55)	28,44 (2,13)	31,71 (5,02)	35,50 (3,41)	20,62 (2,72)	29,16 (2,81)	31,45 (8,81)	36,48 (1,91)
	Pron	3,14 (1,06)	2,44 (0,21)	2,00 (0,55)	1,36 (0,14)	2,64 (0,57)	2,34 (0,34)	1,55 (0,59)	1,85 (0,09)
		TO VI 5s							
Bloco 1	D1	46,72 (12,26)	32,99 (6,29)	25,86 (4,21)	34,76 (12,38)	40,66 (5,97)	30,78 (11,82)	35,58 (10,70)	57,62 (14,93)

	R2	43,43 (3,79)	28,97 (4,41)	39,28 (10,53)	21,08 (2,57)	30,54 (6,46)	28,28 (6,48)	35,82 (6,68)	27,18 (1,61)
	Pron	1,06 (0,22)	1,15 (0,26)	0,68 (0,16)	1,63 (0,44)	1,36 (0,23)	1,19 (0,62)	1,05 (0,53)	2,12 (0,59)
Bloco 2	R1	73,84 (1,67)	53,62 (7,34)	46,73 (6,39)	51,08 (5,81)	46,72 (3,45)	59,73 (6,46)	46,34 (8,15)	74,10 (12,85)
	R2	21,27 (10,97)	29,66 (2,80)	38,26 (5,66)	28,86 (3,31)	21,97 (4,31)	22,98 (7,59)	28,16 (3,69)	37,45 (3,21)
	Prop	3,01 (1,01)	1,82 (0,29)	1,25 (0,29)	1,80 (0,40)	2,20 (0,53)	2,76 (0,76)	1,66 (0,34)	2,00 (0,47)
Bloco 3	R1	80,14 (6,19)	64,75 (4,43)	53,18 (5,80)	45,95 (4,09)	50,05 (6,59)	65,70 (4,38)	41,22 (6,18)	62,98 (4,55)
	R2	20,07 (6,50)	32,70 (3,51)	33,78 (5,99)	38,34 (2,99)	25,14 (3,70)	30,20 (2,83)	37,39 (10,90)	38,78 (2,74)
	Pron	4,44 (1,78)	2,00 (0,28)	1,61 (0,33)	1,20 (0,14)	2,03 (0,47)	2,19 (0,28)	1,27 (0,73)	1,63 (0,22)

TO VI 10s

Bloco 1	R1	45,89 (13,01)	35,08 (4,03)	28,21 (7,62)	35,70 (6,18)	39,89 (5,72)	28,34 (12,04)	32,07 (5,89)	57,33 (15,97)
	R2	46,02 (10,10)	31,40 (4,09)	42,22 (14,15)	20,25 (0,63)	33,81 (4,62)	32,90 (4,96)	47,68 (4,81)	32,08 (5,32)
	Pron	1,01 (0,27)	1,74 (0,23)	0,70 (0,16)	1,76 (0,32)	1,19 (0,22)	0,89 (0,43)	0,67 (0,11)	1,77 (0,33)
Bloco 2	R1	75,58 (5,29)	47,00 (4,19)	43,29 (4,42)	49,02 (11,06)	35,90 (6,04)	59,22 (9,36)	45,10 (2,24)	75,26 (11,37)
	R2	20,28 (8,35)	33,00 (3,44)	46,65 (3,31)	30,41 (7,33)	27,43 (6,84)	25,97 (6,26)	30,79 (5,22)	40,56 (3,35)
	Pron	3,11 (1,06)	1,44 (0,23)	0,93 (0,15)	1,70 (0,64)	1,40 (0,49)	2,39 (0,68)	1,51 (0,36)	1,86 (0,34)
Bloco 3	R1	79,03 (7,38)	62,38 (5,49)	48,21 (7,38)	45,93 (4,76)	46,14 (3,01)	63,53 (7,96)	38,83 (7,15)	60,40 (2,47)
	R2	23,70 (7,70)	36,53 (5,73)	38,81 (7,77)	36,25 (3,14)	27,48 (5,93)	32,41 (6,39)	37,83 (8,61)	42,23 (1,92)
	Pron	3,73 (1,76)	1,74 (0,39)	1,29 (0,34)	1,27 (0,17)	1,75 (0,43)	2,06 (0,64)	1,08 (0,34)	1,43 (0,05)

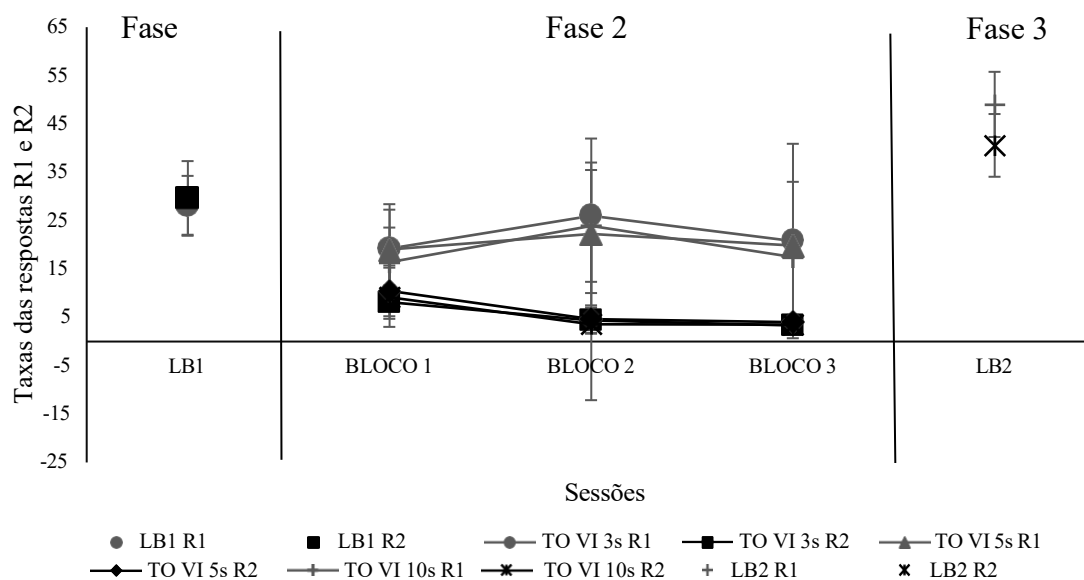
**FASE 3
(Linha de base)**

	Ses	10	10	10	10	10	10	10	10
IR2	R1	54,03 (4,95)	51,83 (7,43)	55,11 (5,74)	43,62 (2,37)	34,84 (3,49)	50,51 (4,63)	49,44 (4,24)	52,89 (5,51)
	R2	54,54 (6,08)	34,84 (5,14)	37,73 (8,82)	41,69 (1,67)	43,19 (5,25)	36,29 (4,28)	41,20 (5,55)	34,97 (6,68)
	Pron	0,99 (0,10)	1,49 (0,17)	1,54 (0,43)	1,04 (0,08)	0,82 (0,14)	1,41 (0,21)	1,22 (0,21)	1,59 (0,24)

Nota: A tabela apresenta os dados do Experimento 4. Dentre eles estão descritos: total de sessões (Ses) realizadas, a média da taxa de resposta das últimas seis sessões para a opção 1 (R1) e opção 2 (R2) e a proporção da resposta (Prop) da Fase 1 e Fase 3; e a média das taxas de respostas e proporção dos blocos de sessões da Fase 2. Os números entre parênteses representam o desvio padrão de cada média.

Apêndice G

Média Geral da Taxa de Resposta das Opções R1 e R2 Durante o TO na Fase 2 do Experimento 4



Nota: A figura 15 apresenta a média geral da taxa de resposta das opções R1 e R2 das últimas seis sessões da Fase 1 e Fase 3 e dos blocos de sessões da Fase 2 durante o período de TO do Experimento 4. O TO foi apresentado contingente apenas a opção R2 da Fase 2. As barras sobrepostas a cada ponto de dado representam os respectivos desvios padrão.