



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO  
PUC-SP

Fábio Nakabashi

O dilema do prisioneiro em ratos:  
níveis de restrição alimentar e seu efeito no comportamento de cooperação

MESTRADO EM PSICOLOGIA EXPERIMENTAL:  
ANÁLISE DO COMPORTAMENTO

São Paulo

2019



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO  
PUC-SP

Fábio Nakabashi

O dilema do prisioneiro em ratos:  
níveis de restrição alimentar e seu efeito no comportamento de cooperação

MESTRADO EM PSICOLOGIA EXPERIMENTAL:  
ANÁLISE DO COMPORTAMENTO

Dissertação apresentada à Banca  
Examinadora da Pontifícia Universidade  
Católica de São Paulo, como exigência  
parcial para obtenção do título de MESTRE  
em Psicologia Experimental: Análise do  
Comportamento, sob orientação da Prof.<sup>a</sup>  
Dra. Nilza Micheletto.

São Paulo

2019

Banca Examinadora:

---

---

---

Autorizo, exclusivamente para fins acadêmicos ou científicos, a reprodução total ou parcial desta dissertação por fotocópias ou processos eletrônicos.

São Paulo, \_\_\_\_ de \_\_\_\_\_ de 2019.

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

O presente trabalho foi realizado com apoio da Fundação São Paulo (FUNDASP), na  
forma de bolsa-dissídio integral

Agradeço à FUNDASP pelo apoio recebido.

## **Agradecimentos**

À minha esposa e companheira na vida, Akico, pela compreensão e apoio nesse desafio, e ao meu filho, Kevin, pela torcida e companhia nas noites de segunda.

À minha professora e orientadora, Prof.<sup>a</sup> Dra. Nilza Micheletto, pela visão, experiência e conhecimento e também pelas palavras de incentivo.

Aos professores da graduação (Ziza, Thomas, Emerson e Amilcar) e mestrado (Mare, Paola, Paula, Mônica, Fani e Nicolau), do curso de Psicologia Experimental da PUC-SP, pela dedicação e paixão que exercem suas funções. Levarei para sempre como exemplo esses docentes empenhados num mundo melhor.

Ao Marcos Alexandre de Medeiros pelo programa de computador, sem o que a pesquisa não seria realizada. Agradeço também pela primeira aula que tive de análise de comportamento ministrada por ele na graduação.

Aos alunos de mestrado e doutorado que conheci, pela paciência com minhas dúvidas e amizade.

Aos amigos e funcionários do Biotério (Neusa, André e Maurício), pela ajuda prestada na pesquisa.

A João Marinho, pela revisão e formatação com normas da APA desta dissertação.

Aos meus animais de pesquisa.

À Pontifícia Universidade Católica de São Paulo, pela bolsa concedida e incentivo à capacitação intelectual dos funcionários.

Nakabashi, F. (2019). *O dilema do prisioneiro em ratos: níveis de restrição alimentar e seu efeito no comportamento de cooperação* (Dissertação de mestrado). Pontifícia Universidade Católica de São Paulo, São Paulo, Brasil.

**Orientadora:** Prof.<sup>a</sup> Dra. Nilza Micheletto.

**Linha de Pesquisa:** Processos Básicos na Análise do Comportamento.

### Resumo

O Dilema do Prisioneiro é um conceito da teoria dos jogos muito utilizado para estudar a cooperação. Neste experimento, foram utilizadas caixas operantes separadas por uma tela vazada e transparente, nas quais o comportamento de cooperar foi simbolizado pelo acionar da barra, e a matriz de pagamento (reforçamento) foi baseada no Dilema do Prisioneiro, replicando o estudo de Wood et al. (2006), mas adicionando níveis de restrição alimentar: *ad libitum*; restrição alimentar de 70% do peso *ad libitum*; e restrição de 90% do peso *ad libitum*, em esquema ABACA. Resultados indicaram que a cooperação unilateral, ou seja, de apenas um dos sujeitos de três duplas de ratos Wistar, aumenta proporcionalmente quanto mais severa a restrição alimentar.

*Palavras-chave:* cooperação, ratos, dilema do prisioneiro, restrição alimentar, análise do comportamento

Nakabashi, F. (2019). *The prisoner's dilemma in rats: levels of food restriction and its effect on cooperative behavior* (Master's thesis). Pontifical Catholic University of São Paulo, São Paulo, Brazil.

**Thesis Advisor:** Prof. Nilza Micheletto, PhD.

**Line of Research:** Basic Processes in Behavior Analysis.

### **Abstract**

The Prisoner's Dilemma is a concept from game theory widely used to study cooperation. In this experiment, operant boxes separated by a perforated and transparent screen were used, in which the cooperating behavior was symbolized by the activation of a bar trigger, and the payment matrix (reinforcement) was based on the Prisoner's Dilemma, replicating the study of Wood et al. (2006), with the addition of food restriction levels: *ad libitum*; food restriction of 70% *ad libitum* weight; and food restriction of 90% *ad libitum* weight, in an ABACA scheme. Results indicated that the unilateral cooperation, that is, of only one of the subjects of three pairs of Wistar rats, increases proportionally to more severe food restriction.

*Keywords:* cooperation, rats, prisoner's dilemma, food restriction, behavior analysis



## Lista de Figuras

Figura 1 – Desenho esquemático da caixa operante.....	47
Figura 2 – Representação esquemática das condições em que as respostas produziam pelotas.....	50
Figura 3 – Média do número de tentativas classificadas como delação mútua, cooperação mútua e cooperação unilateral das quatro últimas sessões das fases de testes do Dilema do Prisioneiro dos sujeitos da dupla S2 e S7 em que vigoraram diferentes níveis de restrição alimentar. ....	54
Figura 4 – Número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S2 e S7 nas sessões das fases de teste do Dilema do Prisioneiro nas diferentes condições de restrição alimentar.....	56
Figura 5 – Média do número de tentativas classificadas como delação mútua, cooperação mútua e cooperação unilateral, das quatro últimas sessões das fases de testes do Dilema do Prisioneiro dos sujeitos da dupla S5 e S10 em que vigoraram diferentes níveis de restrição alimentar. ....	57
Figura 6 – Número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S5 e S10 nas sessões das fases de teste do Dilema do Prisioneiro, nas diferentes condições de restrição alimentar.....	58
Figura 7 – Média do número de tentativas classificadas como delação mútua, cooperação mútua e cooperação unilateral, das quatro últimas sessões das fases de testes do Dilema do Prisioneiro dos sujeitos da dupla S8 e S9 em que vigoraram diferentes níveis de restrição alimentar. ....	60
Figura 8 – Número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S8 e S9 nas sessões das fases de teste do Dilema do Prisioneiro, nas diferentes condições de restrição alimentar.....	62
Figura 9 – Porcentagem de respostas no escuro durante as sessões de treino discriminativo dos sujeitos de cada dupla, nas diferentes etapas do treino.	64
Figura 10 – Frequência acumulada de respostas por tentativa das últimas quatro sessões de cada condição alimentar da dupla S2 e S7.....	66
Figura 11 – Frequência acumulada de respostas por tentativa das últimas quatro sessões de cada condição alimentar da dupla S5 e S10.....	69
Figura 12 – Frequência acumulada de respostas por tentativa das últimas quatro sessões de cada condição alimentar da dupla S8 e S9.....	72
Figura 13 – Número de pelotas recebidas por dupla S2 e S7, S5 e S10 e S8 e S9 nas últimas quatro sessões de cada condição de restrição da fase do Dilema do Prisioneiro.....	75

**Lista de Tabelas**

Tabela 1 – Matriz de Pagamento no Jogo do Dilema do Prisioneiro (adaptada de Axelrod, 1984, p. 8).....	14
Tabela 2 – Fases do Experimento e Condições de Restrição Alimentar .....	53

## Sumário

Introdução.....	1
Definições de Cooperação na Análise do Comportamento .....	2
Procedimentos de desempenho.....	8
Procedimentos de escolha.....	8
O Dilema do Prisioneiro .....	9
Teoria dos jogos e o Dilema do Prisioneiro. ....	9
Origem e esquema do Dilema do Prisioneiro.....	12
<i>Tit for tat</i> (olho por olho).....	14
Jogo do Dilema do Prisioneiro sob a óptica da Análise do Comportamento.....	15
Dilema do Prisioneiro em sujeitos hão humanos (pesquisas em análise de comportamento).....	20
Dilema do Prisioneiro em sujeitos não humanos (pesquisas em biologia evolutiva).....	32
Privação: uma variável motivacional. ....	41
Objetivo .....	45
Método.....	46
Sujeitos.....	46
Aparato.....	46
Procedimento .....	47
Restrição alimentar. ....	47
Fase de condicionamento operante.....	48
Fase de treino discriminativo.....	49
Fase de testes do Dilema do Prisioneiro.....	51
Resultados e Discussão.....	54
Análise de Pelotas Recebidas nas Quatro Sessões Analisadas em Cada Condição de Restrição.....	74
Conclusão .....	81
Referências .....	82

Como o biólogo evolucionista britânico e geneticista John Maynard Smith (1982) observou, se uma espécie parasita aumenta sua população utilizando recursos que podem ser fatais ao hospedeiro, os outros organismos no mesmo hospedeiro limitam sua própria reprodução e reduzem seu consumo de recursos, aumentando a sobrevivência da fonte de alimento. Esse tipo de dilema em que indivíduos egoístas se beneficiam de atividades que comprometem benefícios compartilhados é denominado *tragédia dos comuns*, e é uma das principais razões pelas quais a *cooperação* é fundamental na história da evolução (Hardin, 1968).

A importância da cooperação destaca-se em exemplos que encontramos na natureza, como pássaros que voam em formato de “V” e mudam constantemente de posição entre si, a fim de vencer a resistência do ar em viagens de longa distância e evitar a fadiga do que está à frente recebendo a maior carga dessa resistência (Portugal et al., 2014). Predadores unem-se para caçar uma presa maior, atacando-a de forma estratégica e obtendo maior êxito. Uma presa busca segurança ao andar em bandos e defender-se em grupos. Nos humanos, encontramos exemplos cotidianos de como, ao fazer um protesto ou uma passeata, demonstramos que juntos temos mais força do que isolados. A cooperação é importante para a manutenção das espécies e para a evolução do que chamamos de cultura e sociedade.

Influenciado pela literatura de Darwin, Kropotkin (1902/2009), naturalista e revolucionário russo, interpreta que há duas lutas pela sobrevivência: entre os indivíduos da mesma espécie por recursos finitos, o que leva à competição; e contra outras espécies predadoras e condições adversas do meio ambiente, levando à cooperação e apoio mútuo entre os indivíduos. A partir dessa premissa, Kropotkin (1902/2009) desenvolve o argumento de que a cooperação é um fator evolutivo essencial, observando grupos de animais e suas relações inter e intraespecíficas.

Ademais, segundo Skinner (1953/2003), “as consequências reforçadoras geradas pelo grupo excedem facilmente os totais das consequências que poderiam ser conseguidas pelos membros se agissem separadamente. O efeito reforçador total é enormemente acrescido” (p. 341).

Conclui-se, portanto, que a cultura somente pode existir mediante dois requisitos básicos: a capacidade dos membros do grupo de aprenderem uns com os outros; e a existência da sociedade – e uma verdadeira sociedade inclui a cooperação (Baum, 2005/2008, p. 281).

### **Definições de Cooperação na Análise do Comportamento**

Do ponto de vista da análise do comportamento, a cooperação tem sido definida de diferentes formas por diferentes autores (Echagüe, 2006). De forma geral, a cooperação, sendo um comportamento mediado pelo comportamento de outro organismo, é caracterizada como comportamento social, sujeita tanto ao caráter dinâmico e flexível das contingências sociais que são constituídas pelo ambiente social (Guerin, 1994), quanto às leis comportamentais, como: seleção por reforçamento e operações estabelecedoras/motivacionais. No entanto, para que a relação seja denominada como cooperativa, deve haver reforçamento mútuo, de modo que os indivíduos envolvidos recebam reforçadores se o *desempenho do grupo* atingir o critério específico, seja de forma equitativa, seja de forma não equitativa entre os membros, mas proporcional ao trabalho efetuado. Dessa forma, difere-se da relação competitiva, na qual a distribuição de reforçadores é desigual e excludente: a liberação de reforçadores para um indivíduo limita ou mesmo anula a obtenção de reforçadores pelos demais (Abreu-Rodrigues & Ribeiro, 2005).

Skinner (1953/2003) define *comportamento social* como o comportamento de duas ou mais pessoas uma em relação a uma outra ou em conjunto em relação ao

ambiente comum e complementa teorizando que o comportamento de grupos é formado por indivíduos cujo reforçamento é mediado por outro organismo, além de que, diferentemente do ambiente mecânico, o reforçamento social depende da condição do agente reforçador. Assim, diferentes respostas podem ter a mesma consequência e diferentes consequências surgem para uma mesma resposta, proporcionando um “comportamento mais extenso” (p. 327) do que o comportamento reforçado em um ambiente não social, ainda que obedecendo às mesmas leis. O comportamento reforçado socialmente, portanto, é flexível e comumente reforçado de maneira intermitente.

Ao analisar um episódio social, Skinner (1953/2003) descreve como o comportamento de um organismo pode estar sob controle da resposta do outro, e, para isso, exemplifica os comportamentos de conduzir e ser conduzido em uma dança, em que o líder está sob controle das variáveis externas, e o liderado está sob controle do líder, com consequências reforçadoras de múltiplas contingências: quem conduz estabelece o padrão da dança respeitando o espaço disponível; o conduzido segue os movimentos do líder; e ambos têm o comportamento regulado a partir dos comportamentos um do outro.

No estabelecimento de situações cooperativas, ou seja, cooperação entre dois ou mais organismos, as respostas operantes dos envolvidos devem ser coordenadas ao seu respectivo produto. Na cooperação, há produção de reforçadores que retroagem sobre os comportamentos envolvidos e que somente podem ser produzidos de uma maneira específica, uma vez que a cooperação envolve respostas inter-relacionadas ou inter cruzadas ao seu respectivo produto (produzir um reforçador positivo, remover um negativo). Como exemplo, Skinner (1953/2003) descreve um experimento feito com pombos colocados numa caixa e separados por uma tela transparente. Em cada um dos lados, botões de diversas cores deveriam ser acionados. A resposta apenas era reforçada quando ambos apertassem ao mesmo tempo os botões de mesma cor, que era alternada

continuamente ao acaso. Nesse caso, o pombo-líder explorava os botões, enquanto o liderado ficava sob controle do líder, repetindo sua sequência.

Além de compartilharem a essência da definição de Skinner (1953/2003), Keller e Schoenfeld (1950/1974) afirmam que o comportamento social não deve se contrapor à psicologia do indivíduo e que os estímulos sociais não diferem, nas suas dimensões, dos outros estímulos, mas apenas em sua origem, já que provêm do comportamento do outro indivíduo ou de seus produtos. Além disso, os estímulos sociais também não diferem quanto à função. Os autores descrevem comportamentos sociais em animais primitivos, como organelas unicelulares, insetos, pássaros e mamíferos não humanos. Os estímulos sociais que controlam o comportamento animal são numerosos, e um exemplo seria o ciclo estral mediando o comportamento do organismo e controlando o de seu companheiro. Outro seria o cuidado maternal de filhotes ainda dependentes, que favorece o estabelecimento de estímulos discriminativos: aves papa-figos criadas longe dos pais, por exemplo, não têm o padrão de canto típico da espécie. Keller e Schoenfeld (1950/1974) exemplificam os vários tipos de relação social dos animais descrevendo três deles: a *dominância*, a *imitação* e a *cooperação*.

Na cooperação, segundo os autores, o comportamento combinado de dois ou mais organismos é necessário para obter reforçador positivo, ou remover negativo, ou ambos para um ou outro. Em animais não humanos e humanos, a cooperação pode ser conseguida por meio de um ensino inicial e observação de seu desenvolvimento. A eficácia da instrução verbal em seres humanos adultos depende, também, de uma história de ensino, mas o esforço cooperativo em algumas atividades somente é conseguido com dificuldade e, às vezes, não acontece.

Em qualquer caso, quer com seres humanos ou animais, em qualquer tarefa, a cooperação supõe duas coisas: (1) cada ação do organismo deve ser

discriminativa para o desempenho do outro; e (2) cada organismo deve ser reforçado pela parte que desempenha no esquema cooperativo. (Keller & Schoenfeld, 1950/1974, p. 374)

Keller e Schoenfeld (1950/1974) citam, ainda, um experimento clássico de cooperação, “Cooperative Problem Solving in Rats”, de Daniel (1942). No experimento, dois ratos foram colocados para cooperar em uma gaiola na qual o alimento era distribuído sobre uma placa eletrificada que somente podia ser desligada acionando uma plataforma situada em um piso superior, impossibilitando que ambos se alimentassem, ao mesmo tempo, sem receber o choque. Após o treino, esses ratos se alternavam de forma consistente, e o rato que ficava na plataforma podia tocar levemente o outro rato que se alimentava, o que possivelmente controlava a resposta de cessar a alimentação e alternar. O toque, possivelmente, constituiu-se em estímulo pré-aversivo que antecedia o estímulo aversivo (choque), provocando certa ansiedade, pois caso o rato que estava no piso superior acionando a plataforma e suspendendo o choque saísse, o choque podia retornar para o rato que se alimentava.

Segundo Keller e Schoenfeld (1950/1974), Daniel (1943) realizou um outro estudo em que faz alterações nesse procedimento. O experimentador retirou gradualmente o estímulo aversivo (choque), e o aparato consistiu em um comedouro com tampa automática. Como no experimento anterior, um rato devia permanecer na plataforma para a tampa ser levantada e liberar o alimento para outro rato. Não houve comportamento cooperativo na ausência de punição. Nesse caso, concluiu-se que o único motivo para trabalhar foi a fome, e o único reforçador para ir à plataforma, longe do comedouro, foi a chance de comer mais tarde se o parceiro cooperasse da mesma maneira. Keller e Schoenfeld (1950/1974) afirmam que a resposta de ir à plataforma envolveu um reforçador muito atrasado, e os toques dificilmente foram associados à



retirada da comida. Concluem, ainda, que esse tipo de cooperação, quando um estímulo inicialmente neutro para determinada resposta (comportamento de cooperação) é emparelhado com o estímulo condicionado (companheiro na plataforma/tampa do comedouro descoberto) é muito difícil de obter com animais não humanos, apesar de verificado rotineiramente no comportamento social entre pessoas.

Gardner, Corbin e Beltramo (1984) apontam que a alternância de resposta entre os sujeitos para obtenção de reforçadores nos remete ao trabalho de Mowrer (1940), ilustrado por um filme (*An Experimentally Produced Social Problem in Rats*, disponível na internet), cujos resultados demonstraram que apenas um rato pode trabalhar para outros dois na mesma caixa, saciando ambos, antes de poder alimentar-se. Pela descrição do filme, três ratos são treinados individualmente para acionar uma barra para ganhar alimento, e, em seguida, são alocados os três ratos, ao mesmo tempo, numa caixa operante com a barra e o comedouro colocados em lados opostos. Ao acionar a barra, o sujeito tem de atravessar todo o diâmetro da caixa para adquirir o alimento, dando oportunidade para que outro se aproveite do reforçador. No primeiro dia, todos os animais trabalharam, ou seja, pressionaram a barra de forma alternada, apesar de nenhum receber reforçador por sua respectiva resposta ao pressionar a barra, e o rato mais próximo do comedouro “furtou” o alimento liberado pela resposta à barra. Após 90 min no aparato, mesmo estando todos ainda famintos, a taxa de resposta diminuiu sensivelmente, havendo a quebra da relação resposta na barra – reforçador. No segundo dia, os ratos competiram pelo lugar à frente do comedouro. Independentemente de o reforçador ser produzido, a resposta de pressão à barra diminuiu mais ainda. No terceiro dia, com a resposta de pressão à barra quase extinta, iniciou-se reação de agressividade contra a barra, até que um dos ratos, e somente um deles, aumentou sua taxa de resposta, mesmo sem que ele mesmo recebesse o reforçador ou, pelo menos, reforçador

proporcional à taxa de resposta, restando-lhe apenas alguns grãos que os demais não conseguiram comer. No quarto dia, o rato que acionou a barra mais vezes no dia anterior aumentou ainda mais sua frequência de resposta, acionando a barra 1.156 vezes, contra três de seu companheiro e zero do outro rato. Após 90 minutos, o rato “cooperativo” saciou os companheiros e pôde alimentar-se, “resolvendo” o problema. A “solução” permaneceu inalterada indefinidamente: um rato trabalhador e os outros parasitando seu trabalho. Mowrer (1940) discute suas observações baseando-se na hierarquia por submissão – ascendente, que define a dominância entre indivíduos do mesmo grupo – e argumenta que, quando a competição é imposta, os ratos não aprendem a dividir o alimento, como havia ocorrido em experimentos anteriores do mesmo autor. Similarmente, Gardner et al. (1984), em seu estudo, citam outros trabalhos em que houve desenvolvimento de trabalhador – parasita (Baron & Littman, 1961; Mowrer, 1960; Oldfield-Box, 1967).

Hake e Vukelich (1972) revisaram e classificaram procedimentos de cooperação sob a óptica da análise do comportamento. Os aspectos essenciais de qualquer procedimento de cooperação, segundo seu trabalho, são: (a) que os reforçadores de ambos os indivíduos sejam, ao menos em parte, dependentes das respostas de outro indivíduo; e (b) que essas respostas, denominadas *respostas cooperativas*, resultem em uma divisão equitativa de respostas e reforçadores. Os processos básicos de cooperação mostrados na literatura sofrem, porém, considerável variação. Dessa forma, a unidade básica de procedimento, que dá o requisito mínimo de resposta para o reforçador, foi denominada *episódio cooperativo* (Hake & Vukelich, 1972).

O episódio cooperativo, então, refere-se a todas as respostas ou, pelo menos, uma resposta cooperativa necessária para a entrega de um reforçador. Essas respostas podem ser de apenas um ou de ambos os indivíduos e podem incluir reforçador para um ou para

ambos, podem incluir distribuição equitativa ou não de trabalho e/ou reforçadores, mas devem incluir pelo menos uma resposta cooperativa e um reforçador. Entretanto, para que a definição de cooperação seja atingida, ou seja, a distribuição equitativa de trabalho e/ou reforçadores, vários episódios devem ocorrer ao longo do tempo, pois dificilmente a cooperação é alcançada em um único episódio cooperativo. O tempo decorrido para que episódios sucessivos se transformem em cooperação foi denominado *período de cooperação* (Hake & Vukelich, 1972).

Hake e Vukelich (1972) dividiram, ainda, as pesquisas de cooperação em duas categorias:

**Procedimentos de desempenho.** Medem o comportamento que ocorre durante a realização de uma solução cooperativa, ou seja, o próprio comportamento cooperativo. Segundo Faleiros (2009), os procedimentos de desempenho são procedimentos de “cooperação forçada”, pois não há reforçamento sem cooperação.

**Procedimentos de escolha.** Medem apenas a seleção de uma solução. Segundo Hake e Vukelich (1972), essas pesquisas não medem o comportamento necessário para realizar uma solução. Nos procedimentos de escolha, como no Dilema do Prisioneiro, que será analisado neste trabalho, existem várias soluções alternativas para a escolha, incluindo uma em que os reforçadores são igualmente distribuídos (cooperação) e uma em que os reforçadores são desigualmente distribuídos (competição).

## O Dilema do Prisioneiro

O Jogo do Dilema do Prisioneiro é um jogo de soma não zero<sup>1</sup> que constitui um problema clássico da teoria dos jogos.

**Teoria dos jogos e o Dilema do Prisioneiro.** A teoria dos jogos, por definição de von Neumann e Morgenstern (1943/2007), matemático húngaro naturalizado americano e economista austríaco respectivamente, sistematiza matematicamente as decisões de estratégia, o resultado de interações competitivas ou cooperativas e as sequências dos comportamentos de duas ou mais pessoas envolvidas em modelos de jogos, ou seja, sujeitas a um conjunto específico de regras.

---

<sup>1</sup> Em jogos de soma zero, o benefício total para todos os jogadores, para cada combinação de estratégias, sempre soma zero (ou, descrevendo mais informalmente, um jogador somente lucra com base no prejuízo de outro). Em jogos de soma diferente de zero, o ganho de um dos jogadores não necessariamente corresponde à perda dos outros.

Esses autores consolidaram as bases de uma moderna teoria dos jogos, na qual o conceito de *utilidade*<sup>2</sup> é fundamental, detalhando a prova do teorema minimax (na qual há sempre uma solução racional para um conflito entre dois indivíduos, cujos interesses são completamente opostos) e a formalização dos jogos de soma zero (jogos em que o ganho de um jogador representa necessariamente a perda para o outro jogador).

A teoria dos jogos, segundo von Neumann e Morgenstern (1943/2007), é aplicada em jogos não cooperativos e pode ser comparada à teoria darwinista da seleção natural, na qual, em uma disputa de soma zero e reduzido o jogo de vários sujeitos para duplas, o mais adaptado sobrevive.

Um dos principais personagens da teoria dos jogos foi John Nash (1951), matemático norte-americano que procurou generalizar o teorema de minimax. Em Nash (1951), a rivalidade, que era individual, passou a ser estudada em grupo. A solução

---

<sup>2</sup> “Uma função de utilidade é a quantificação numérica das preferências de uma pessoa com relação a certos objetos, ou seja, é a tradução de ganhos qualitativos em números” (Feliciano, 2007).

“Na teoria dos jogos, admite-se que os jogadores tentem maximizar seus ganhos ou minimizar as perdas. A seleção natural, por sua vez, privilegia os ‘bons jogadores’. Os ganhos que cada jogador recebe em um determinado jogo são afetados pela utilidade que esses ganhos efetivamente têm para cada jogador. Daí se extrai um conceito importante da teoria dos jogos: a teoria da utilidade. Os bons jogadores, do ponto de vista da seleção natural, equivalem aos indivíduos mais capazes de sobreviver e reproduzir em um determinado ambiente; já na teoria da utilidade os bons jogadores ganham mais a despeito do ganho do outro, logo alcançam uma pontuação maior” (Alencar & Yamamoto, 2008, p. 524).

encontrada foi batizada de *ponto de equilíbrio*, uma situação em que cada jogador não tem incentivo em mudar sua estratégia se os demais jogadores não o fizerem. Em qualquer tipo de jogo, existe pelo menos um ponto de equilíbrio que pode ser encontrado utilizando estratégias comuns a todos (mistas) como uma variável na aplicação de estratégias individuais (egoístas). Em outras palavras, é possível maximizar ganhos individuais cooperando com o adversário, ou seja, cada jogador tem duas contingências para formular sua estratégia: a individual e a coletiva, e a melhor estratégia consiste em equilibrar a magnitude do reforçador para evitar punições e maximizar o ganho individual. O teorema de Nash é aplicável a jogos não cooperativos, grupos e de soma zero ou não zero, mas, para que ocorra o equilíbrio, é necessário que os jogadores se comportem racionalmente e não se comuniquem antes do jogo, a fim de evitar acordos.

Fora do contexto de economia e matemática, a pergunta sobre como a cooperação surge e se mantém entre espécies era uma grande questão em biologia evolutiva. Smith (1982) introduz a ideia das *estratégias evolutivamente estáveis* (EEE; em inglês, *evolutionarily stable strategy* – ESS), mostrando como a teoria dos jogos pode ser aplicada para o comportamento dos animais, utilizando-se os conceitos do teorema de Nash e analisando-se a dinâmica global de uma população de jogadores interagindo entre si. Se, na teoria dos jogos, o conceito de estratégia se baseia na escolha feita pelos participantes, na biologia evolutiva, trata de fenótipos herdados por um ou outro gene. Na ESS, o equilíbrio de Nash de uma população dá-se quando um grupo dessa população não é invadida por uma população mutante (Almeida & Chacha, 2014).

O livro *A Evolução da Cooperação*, publicado por Axelrod (1984), descreveu como a cooperação pode emergir e persistir com base na aplicação da teoria dos jogos. Ele popularizou torneios com simulações de estratégias computacionais para o Dilema do Prisioneiro Iterado, que veremos adiante, com a finalidade de analisar como a

cooperação poderia emergir até mesmo em participantes utilizando estratégias individuais, e o resultado observado foi que uma estratégia simples, no caso a *tit for tat*, explicada adiante, teve melhores resultados contra outros algoritmos mais complexos.

Na biologia evolutiva, há duas teorias para explicar o surgimento de cooperação a partir de um processo evolutivo genético: *teoria do parentesco* e *teoria da reciprocidade*. Ambas utilizam o Dilema do Prisioneiro em seus estudos, sendo que a teoria do parentesco parte de uma maior cooperação quanto maior o material genético compartilhado, enquanto a teoria da reciprocidade, quando há: (a) igual probabilidade de que os dois indivíduos interajam novamente em um contexto biológico; e (b) a viabilidade de uma estratégia cooperativa em um ambiente dominado por indivíduos não cooperativos (Faleiros & Hubner, 2008).

**Origem e esquema do Dilema do Prisioneiro.** Em janeiro de 1950, os matemáticos norte-americanos Melvin Dresher e Merrill Flood, da Rand Corporation, prepararam uma experiência com a finalidade de verificar se o comportamento humano encontraria estratégias de equilíbrio previstas pela teoria de John Nash. Realizaram os testes com a participação do matemático John Williams, chefe de departamento da Rand, e do economista Armen A. Alchian, da Universidade da Califórnia, que interagiam entre si. Entretanto, em vez de buscarem suas estratégias individuais dominantes (estratégia de um jogador que produz o melhor resultado, independentemente da estratégia escolhida pelo outro) e pararem no *ponto de equilíbrio*, os jogadores optaram por cooperar mais vezes do que competir um com outro. Posteriormente, esse famoso experimento ficou conhecido como Dilema do Prisioneiro, graças à história contada pelo canadense Albert William Tucker em um seminário para psicólogos na Universidade de Stanford, que simplificava o entendimento do experimento.

Uma descrição simples do Dilema é a seguinte: dois presos foram interrogados separadamente. Se o suspeito denunciasse o amigo e este ficasse calado, ou seja, resposta de delação frente à resposta de cooperação (D-C), denominada *tentação* na teoria dos jogos, ou *delação unilateral* neste trabalho, o suspeito ficaria livre e o amigo arcaria sozinho com a pena máxima, e vice-versa; na resposta de cooperação frente à resposta de delação do oponente (C-D), chamada de *sucker/simplório* na teoria de jogos, ou *cooperação unilateral* neste trabalho, o suspeito seria condenado à pena máxima e o amigo sairia livre. Se ambos ficassem calados – resposta de *cooperação mútua* neste trabalho (C-C) e de *recompensa* na teoria dos jogos –, ambos pegariam uma sentença mais branda e leve. Finalmente, se ambos se denunciassem mutuamente – resposta de *delação mútua* neste trabalho (D-D) e de *punishment/penalidade* na teoria de jogos –, ambos pegariam uma pena menor que a máxima, mas maior que a pena leve.

O Dilema do Prisioneiro, portanto, segue uma matriz na qual são distribuídos os pontos para: tentação (T), que explora o outro; *sucker/simplório* (S), que se deixa explorar pelo outro; recompensa/cooperação (R), por ambos não se denunciarem; e *punishment/penalidade* (P), por delação mútua.

A matriz de pagamento deve ser obtida sob três condições para que exista o Dilema. A primeira determina que a tentação (T) seja superior à recompensa (R), a recompensa maior que a penalidade (P) e que todas sejam maiores que o ganho do simplório (S):  $T > R > P > S$ , em que o traidor ganha mais que o cooperador ao jogar com ele (Rapoport & Chammah, 1970).

Na segunda condição, a recompensa (R) deve ser maior ou igual à média entre a tentação (T) e o simplório (S):  $R > [(T + S) / 2]$ . No jogo iterado, dois jogadores que alternam entre cooperação e traição têm um ganho menor do que se cooperassem sucessivamente (Axelrod, 1984).



Finalmente, em terceiro lugar, a escolha por R não produz o maior reforçador individual da matriz, mas o resultado de ambos (R/R) é uma soma que resulta no maior reforçador geral:  $R/R > S/T$ , ou  $T/S$  ou  $P/P$ .

Tabela 1

*Matriz de Pagamento no Jogo do Dilema do Prisioneiro (adaptada de Axelrod, 1984, p. 8)*

		Cooperar	Delatar
Jogador A	Cooperar	Ganho médio/Ganho médio C-C (ex.: $R = 3/R = 3$ )	Maior perda/Maior ganho C-D (Ex.: $S = 0/T = 5$ )
	Delatar	Maior ganho/Maior perda D-C (Ex.: $T = 5/S = 0$ )	Média perda/Média perda D-D (Ex.: $P = 1/P = 1$ )
		Jogador B	

***Tit for tat (olho por olho).*** No estudo de uma teoria da evolução, Axelrod (1984) promoveu um torneio envolvendo pesquisadores de várias áreas, no qual diferentes estratégias seriam confrontadas por meio de programas de computador. Esses confrontos são chamados de *simulações computacionais* ou apenas *simulações*, que Pegden, Sadowski e Shannon (1995) definem como um processo de programar um modelo computacional imitando um sistema de comportamentos reais e conduzir experimentos com esse modelo, com o propósito de entender e/ou avaliar estratégias de todo o processo.

Numerosos pesquisadores e especialistas da teoria dos jogos de diferentes disciplinas submeteram suas propostas de estratégias a serem assumidas pelos participantes. As estratégias foram, então, confrontadas entre si e comparadas por modelos matemáticos, e a que forneceu o melhor resultado (menor prejuízo e maior

estabilidade) foi a *tit for tat* (TFT), na qual o primeiro lance era sempre cooperativo e depois repetia sempre o comportamento do adversário. Axelrod (1984) justifica o sucesso da TFT nos torneios pelas suas características: (a) simpatia: não ser o primeiro a delatar; (b) perdão: não guardar ressentimento e cooperar assim que seu parceiro o fizer; (c) retaliação: penalizar imediatamente um comportamento de traição; e (d) robustez: clareza no comportamento, demonstrando uma estratégia facilmente identificável.

### **Jogo do Dilema do Prisioneiro sob a óptica da Análise do Comportamento**

Em uma análise comportamental, Komorita, Hilty e Parks (1991) argumentaram que a estratégia TFT promove cooperação não por haver simpatia, perdão, vingança ou bondade, mas porque a cooperação é reforçada pelo outro jogador na próxima tentativa. Na estratégia TFT, escolhas de delação não são reforçadas e encorajam o jogador a mudar para a cooperação. Os autores argumentam que, enquanto a biologia evolutiva busca prever as classes de organismos que têm maior probabilidade de adaptação (sobrevivência) de uma geração para a outra, a análise do comportamento busca prever o comportamento dos organismos sob determinadas condições.

Rachlin, Brown e Baker (2000) estudaram o Dilema do Prisioneiro, especialmente o iterado, sob a óptica da análise do comportamento. Segundo os autores, a cooperação no Jogo do Dilema do Prisioneiro depende de fatores extrínsecos ao momento do próprio jogo, como a experiência em jogos anteriores e expectativas sobre jogos futuros, e, por essa razão, os estudos em Jogos do Prisioneiro são iterados. No estudo, a abordagem dos autores ocorreu em termos de contingências de reforçamento e punição, sendo que o contexto relevante da escolha não foi um “evento interno”, como simpatia, perdão, vingança ou bondade, e a “experiência anterior” foi uma história de reforçamento ou punição. “Acordos e garantias”, que são termos econômicos e

matemáticos, por seu turno, foram postos como instruções laboratoriais (matriz de pagamento/magnitude de reforço) e tratados como estímulos discriminativos (SD).

As contingências no Jogo do Dilema do Prisioneiro Iterado são as seguintes: na estratégia *tit for tat* (TFT), o Jogador B reforça a cooperação do Jogador A na tentativa seguinte à cooperação ( $n + 1$ ) e pune a delação de A na mesma situação, ambos aprendendo a cooperar consistentemente. A estratégia TFT essencialmente converte esse dilema social em um dilema individual assemelhando a um problema individual de autocontrole. O termo “escolha do Jogador B” pode ser substituído por “escolha anterior do Jogador A”. O Jogador A ganha um maior número de pontos na delação unilateral (tentação), mas paga esse ganho na próxima tentativa, por ter de escolher cooperação unilateral (simplório) ou delação mútua (penalidade), em vez de delação unilateral (tentação) ou cooperação mútua (recompensa), pois o Jogador B espelhará a jogada anterior. Nesse caso, o Jogador B, na tentativa corrente, coopera e reforça a cooperação do Jogador A na tentativa prévia, ou delata na tentativa corrente, punindo a delação do Jogador A na tentativa prévia. Essa estratégia TFT, em jogos repetitivos, geralmente ensina A e B a cooperarem entre si, pois A delatar na tentativa prévia e esperar que B coopere na tentativa corrente são situações incompatíveis (o Jogador A devia ter cooperado na tentativa prévia).

Killingback e Doebeli (1999) estudaram o Dilema do Prisioneiro Iterado padrão e contínuo, sendo que *contínuo* os autores definem como uma extensão natural do Dilema do Prisioneiro aplicado a populações definidas geograficamente, a fim de compreender como a cooperação surge e evolui nas diversas populações. Os autores introduzem o termo *investimento variável* (Killingback & Doebeli, 2002), que explica como pequenos comportamentos surgem em pequenas populações espacialmente definidas. Mucholin, Faleiros e Ferreira (2018) descrevem as contingências de reforçamento do Dilema do

Prisioneiro na visão de Killingback e Doebeli (2002) da seguinte forma: cada jogador tem duas alternativas, “cooperar ou competir”, considerando que o jogo é simétrico e as condições são idênticas para ambos os jogadores; então, a cooperação mútua é mais vantajosa que a competição mútua, pois cada uma das escolhas combinadas dos participantes está relacionada a diferentes consequências: reforçamento ou punição. No caso de uma cooperação unilateral, o Jogador A será punido e o Jogador B será reforçado; na delação unilateral, será o contrário; na cooperação mútua, ambos são reforçados e, a longo prazo, terão acesso a maior ganho; e, na delação mútua, ambos também ganham, mas uma quantia menor do que na cooperação mútua.

Fidelis e Faleiros (2017) fizeram uma revisão de literatura sistemática sobre o Jogo do Dilema do Prisioneiro na análise do comportamento. Os principais temas encontrados nos resultados da pesquisa foram: comportamento social, autocontrole e metacontingências. Até a década de 1990, os experimentos eram realizados somente para investigar variáveis relacionadas ao comportamento social. A partir dos anos 1990, os experimentos que utilizam o Dilema do Prisioneiro passaram a investigar os efeitos do comportamento de variáveis relacionadas a contingências denominadas de autocontrole em conjunto com o comportamento social. Metacontingências no Dilema do Prisioneiro foram objetos de estudo a partir de 2012.

Segundo Fidelis e Faleiros (2017), o período de maior produção na área ocorreu entre 2000 e 2010, com 15 artigos, correspondendo a 70% de toda a produção selecionada. Com relação às variáveis manipuladas, os resultados encontrados, separados por temas, foram:

- (a) Comportamento social envolvendo interação verbal (possibilidade dos participantes se comunicarem em determinados momentos durante o jogo): tipo de *feedback* (manipulados tipos de informação ao participante

sobre a possibilidade de cooperação ou competição por parte do outro jogador); falso *feedback* (disponibilizado ao participante um comentário discrepante de qual escolha o outro jogador faria); presença do outro participante (investigava o efeito da presença ou não do outro participante no momento em que as escolhas eram realizadas); modelo de aprendizagem (investigado, em uma simulação computacional, o efeito da aprendizagem prévia na emissão de respostas cooperativas); desconto social (avaliação dos efeitos da proximidade e familiaridade entre os participantes sobre as respostas cooperativas); acesso à pontuação com valor da consequência (variáveis independentes eram o acesso à pontuação após cada tentativa no jogo e os valores dos reforçadores, respectivamente); presença ou ausência do outro jogador e valor da consequência; tamanho do grupo; e interação verbal com pausa entre tentativas (esta última variável indicava uma “trégua” durante o jogo com cinco participantes);

- (b) Autocontrole: atraso do reforçador (intervalo de tempo entre a resposta e o reforçador); atraso e valor do reforçador (ou magnitude, que é o valor do reforçador no momento em que o mesmo é disponibilizado); e resposta de comprometimento (comportamento do participante que dá acesso somente à recompensa maior e atrasada, impossibilitando a escolha da recompensa menor e imediata);
- (c) Comportamento social e autocontrole: valor do reforçador (ganho da pontuação durante o jogo, alterada a depender da escolha feita pelo outro participante); probabilidade de reciprocidade e presença de outro participante (referiam-se, respectivamente, à probabilidade que o

participante tinha de receber uma resposta recíproca do outro no jogo e a como a presença de outro participante podia alterar a probabilidade de respostas cooperativas); e desconto social (envolvia a proximidade e familiaridade que os jogadores tinham entre si) e temporal (referia-se a um procedimento em que o participante do experimento devia escolher entre duas opções que envolviam reforçadores imediatos ou a longo prazo);

- (d) Metacontingências: interação verbal (possibilidade de os participantes se comunicarem em determinados momentos durante o jogo) e valor da consequência cultural (alteração e manutenção das contingências entrelaçadas); acesso e valor da consequência cultural (coordenação de comportamentos de múltiplos indivíduos); e possibilidade de tirar pontos do parceiro (oportunidade de o jogador diminuir a quantidade de pontos obtidos pelo outro participante, resultando no aumento dos pontos do primeiro).

Com base nos dados obtidos, os autores afirmaram que a análise experimental do comportamento tem ampliado a investigação em relação ao Dilema do Prisioneiro, inserindo os conceitos de contingências de reforçamento na aquisição e manutenção de respostas de cooperação. Além disso, é pioneira ao usar o conceito de autocontrole como variável nos estudos de cooperação com o Dilema do Prisioneiro, ampliando, em consequência disso, o estudo de fenômenos sociais sob a óptica desse conceito.

Segundo Rachlin (2000), estudos típicos de autocontrole em não humanos envolvem, basicamente, uma escolha entre um reforçador menor e mais rápido contra um reforçador maior e mais demorado numa determinada tentativa, enquanto, num jogo do Dilema do Prisioneiro com um parceiro usando a estratégia TFT, as escolhas são (dentro

de uma tentativa) entre um reforçador menor e um reforçador maior, ambos imediatos, sendo que a escolha de um reforçador menor leva a uma recompensa global maior e a escolha maior leva a uma recompensa global menor. Rachlin (2000), classifica esses estudos, respectivamente, como ambivalência simples e ambivalência complexa.

Segundo Faleiros e Hubner (2008), os estudos envolvendo o Jogo do Dilema do Prisioneiro por analistas do comportamento podem investigar variáveis do autocontrole, como magnitude e atraso do reforçador e controle de estímulos para sinalizar a probabilidade de reciprocidade em estudos sobre cooperação. Ainda segundo os autores, quando os estudos envolvem autocontrole e comportamento social, há a possibilidade de se estudar a reciprocidade não apenas em relação ao espaço temporal dos reforçadores e sua magnitude individual, mas também o espaço social, havendo necessidade de reciprocidade de indivíduos no contexto social, ou seja, a cooperação de um indivíduo depende fortemente do grau em que os outros cooperam com ele.

**Dilema do Prisioneiro em sujeitos não humanos (pesquisas em análise do comportamento).** Poucos são os estudos sobre o Dilema do Prisioneiro envolvendo sujeitos não humanos. São descritas, neste estudo, pesquisas na análise do comportamento e outros artigos ligados diretamente ao problema proposto, sendo alguns de autores de fora da área da análise do comportamento.

O experimento de Green, Price e Hamburger (1995) testou pombos e o controle por consequência imediata no Dilema do Prisioneiro Iterado (IPD – *Iterated Prisoner's Dilemma*, na sigla em inglês) em pombos. Nessa pesquisa, o objetivo foi testar pombos com restrição alimentar a 80% do peso *ad libitum*, jogando o IPD em estratégias *tit for tat* (TFT) e respostas randomizadas (RND), sendo que, por randomizadas, entendia-se que o computador cooperava em metade das tentativas randomicamente e delatava na outra metade, independentemente da escolha do pombo; e, por TFT, entendia-se que a

primeira resposta era sempre cooperativa, mas as subsequentes eram espelhadas nas respostas do oponente, ou seja, o computador respondia para o pombo com a mesma escolha do pombo na tentativa precedente. A matriz de pagamento do TFT e RND seguiu o esquema: (a) cooperação do pombo (CP) + cooperação do computador (CC) = três pelotas; (b) delação pombo + cooperação do computador = cinco pelotas; (c) cooperação do pombo (CP) + delação do computador (DC) = nenhuma pelota; e (d) delação do pombo (DP) + delação do computador (DC) = uma pelota.

Green et al. (1995) fizeram três experimentos com diversas fases em cada experimento, alternando significados das luzes (em algumas fases, a chave vermelha significava cooperar e a chave verde significava delatar para o pombo; em outras fases, o significado era o oposto. O mesmo acontecia para as respostas do computador, que eram azul e amarelo) e estratégias implementadas: no Experimento 1, foi adicionado o TFT-dom, cuja matriz de pagamento favorecia a cooperação, ficando a matriz de pagamento da seguinte forma: cooperação do pombo (CP) + cooperação do computador (CC) = três pelotas; delação do pombo (DP) + cooperação do computador (CC) = nenhuma pelota, em vez de cinco; cooperação do Pombo (CP) + delação do Computador (DC) = duas pelotas, em vez de nenhuma; delação do pombo (DP) + delação do computador (DC) = uma pelota.

No Experimento 2, foi adicionado o TFT+1, no qual todos os resultados de pagamento foram acrescidos de uma pelota, *Chicken*, com a delação mútua (DP e DC) tendo o pior resultado para ambos, no seguinte esquema: CP e CC = 3; CP e DC = 1; DP e CC = 5; DP e DC = 0. Posteriormente, introduziu-se TFT x 2, que dobrava a matriz de pagamento de TFT, a fim de testar se a delação continuava mesmo com pagamento maior: CP e CC = 6; CP e DC = 0; DP e CC = 10; DP e DC = 2. Esta última condição foi proposta considerando dados da literatura que indicavam que, em humanos, a cooperação



aumenta com pagamento maior. Finalmente, o Experimento 3 aumentou o número de sujeitos de dois para sete e testou apenas as estratégias TFT e TFT-dom.

Foram analisadas as respostas das últimas cinco sessões em cada condição. O resultado do Experimento 1 foi que os pássaros delataram na maioria das tentativas, sendo que, em algumas condições, os pássaros ficaram sob controle das cores quando elas foram alternadas em termos de significado. Ambos os pássaros cooperaram em quase todas as tentativas em TFT-dom. A troca de cooperação na TFT-dom para delação na TFT produziu um decréscimo acentuado na comida recebida, de 150, que era o máximo, para 50 pelotas, valor perto do mínimo.

A conclusão no Experimento 1 foi de que, sob estratégia TFT e RND, a escolha foi delatar, descartando-se os resultados em que o pássaro se encontrava sob controle da cor da luz (quando houve alternância). Na estratégia RND, na qual predominou delação, os pássaros chegaram perto da escolha ótima (apenas delação), e, na estratégia TFT, delataram mesmo quando havia interesse em cooperar, acumulando perto da quantidade mínima de pelotas que podiam ganhar (Green et al., 1995).

A resposta do computador teve claro efeito no comportamento dos pombos, sugerindo que os pássaros foram sensíveis aos valores de reforçamento. Resultados em TFT-dom também demonstraram que ambos os pássaros foram sensíveis às diferenças no reforçamento com resultados de cooperação pura nessa condição; não se determinou, porém, se os pombos apenas evitavam um pagamento zero de pelotas. Também não ficou claro se suas escolhas estavam somente sob controle dos reforçadores imediatos (Green et al., 1995).

Os resultados do Experimento 2 demonstraram que ambos os pássaros delataram em quase todas as tentativas em TFT, continuaram a delatar com o acréscimo de uma pelota, na inversão das luzes e no retorno das luzes ao padrão original, concluindo-se que

a esquivar do pagamento zero não explicava os resultados dos experimentos. Na condição *Chicken*, a cooperação ocorreu em metade das tentativas. Na condição TFT-dom, os pássaros cooperaram em quase todas as tentativas, chegando perto do máximo de pelotas que podiam ganhar. Na volta à TFT original, voltaram a delatar. Na condição TFT x 2, os pássaros continuaram a delatar – e também com a inversão das luzes (Green et al., 1995).

Finalmente, o Experimento 3 da pesquisa testou sete pássaros, com apenas duas estratégias, TFT e TFT-dom. O objetivo foi eliminar o pequeno número de sujeitos dos experimentos anteriores. O resultado foi conclusivo para cooperação em TFT-dom e delação em TFT (Green et al., 1995).

A conclusão geral dos três experimentos foi que pombos respondem à maior magnitude do reforçador na impulsividade, mesmo que, na tentativa seguinte, se produza um valor mais baixo de pagamento. Falha a expectativa de que um sujeito jogando com um oponente usando estratégia TFT coopere ao longo das tentativas. “Os pombos cooperaram somente quando tal comportamento produziu a recompensa maior imediata (e.g., contra TFT-dom)” (Green et al. 1995, p. 15). Os autores indicam que a generalidade dos resultados deve ser, ainda, avaliada, considerando que eles resultaram de uma matriz de valores de pagamento limitada e do uso apenas de consequências positivas. Green et al. (1995) relacionam os resultados aos obtidos em estudos de autocontrole. Pombos não apresentam autocontrole, são impulsivos. Os autores definem autocontrole, genericamente, como responder nas contingências de recompensa molar (de maior magnitude e maior atraso); e impulsividade como responder nas contingências de recompensa mais imediata (molecular). Entretanto, pombos podem apresentar resposta de compromisso (preferência pelo reforçador de maior magnitude com maior atraso).

Baker e Rachlin (2002) analisaram o estudo de Green et al. (1995) e indicaram aspectos do método que poderiam ter produzido os resultados obtidos. Como as tentativas duravam 25 s cada,

a vantagem da cooperação só poderia ser obtida se o comportamento e reforçamento compusesse pelo menos um período de 50 segundos – muito além do que normalmente é eficaz com pombos. Quando estímulos salientes são emparelhados confiantemente com reforço, no entanto, não humanos conectam e podem abstrair relações entre comportamento e reforço até mesmo com longos atrasos. (Baker & Rachlin, 2002, p. 483)

Diante disso, Baker e Rachlin (2002) propuseram como objetivos de seu estudo: (a) examinar se o autocontrole (de pombos jogando TFT) seria obtido fornecendo estímulos mais salientes e confiáveis (luzes piscantes sinalizando reforçadores de maior e menor magnitude) com atraso constante (entre a resposta e o reforçador) de 3 s; (b) testar se os espaços temporais de escolha afetariam o nível de cooperação, ou seja, se o Dilema do Prisioneiro com oponente jogando estratégia TFT fosse um procedimento de autocontrole (cooperação gera recompensa maior a longo prazo e delação recompensa menor e imediata), os pombos teriam maiores resultados de cooperação quando o ITI (intervalo entre tentativas, ou *inter-trial interval*) fosse menor; por isso, seriam testados ITIs de 0 a 18 s; e (c) verificar se a apresentação de um grupo de escolhas juntas, seguido de longo atraso para depois apresentar um novo grupo de escolhas, aumentaria o nível de cooperação dos pombos, tal como identificado em estudos com humanos.

Foram utilizados seis pombos com treinamento prévio em bicar um painel para ganhar uma mistura de grãos, mantidos a 80% do peso *ad libitum*. O aparato consistia em uma caixa operante com duas chaves de resposta iluminadas com luzes coloridas. Grãos eram liberados num comedouro no meio do painel. Durante o reforçamento, o

comedouro era aceso, e a luz de teto e das chaves eram apagadas. As tentativas iniciavam-se com as duas chaves iluminadas (uma verde e uma vermelha). O verde era sempre a cor de cooperação, e o vermelho era a escolha de delação. Os lados (posição) das chaves verdes e vermelhas variavam conforme a fase do experimento (Baker & Rachlin, 2002).

Uma primeira bicada em uma das chaves apagava a luz da outra chave e não dava consequência à resposta na outra chave e, após nove bicadas na chave de escolha, apagava-se a luz da mesma sem entrega de reforçador para resposta de bicar a chave. A seguir, ocorria um piscar da luz do teto: rápido (100 ms), indicando que o pombo cooperou na tentativa anterior e receberia um reforçador de alta magnitude na tentativa presente; ou lento (500 ms), indicando que o pombo delatou na tentativa anterior, e o reforçador seria de pequena magnitude na tentativa presente, por 3 s. Segundo os autores, esses seriam os estímulos mais salientes e confiáveis. Reforçadores de grande magnitude consistiram em 5 a 6 s de acesso à comida, enquanto reforçadores pequenos, em 1 a 2 s de acesso. Após piscar por 3 s, a luz de teto foi extinta, e uma das quatro possibilidades de reforçamento da matriz de pagamento foi apresentada (Baker & Rachlin, 2002).

A matriz de pagamento correspondia a: 5 s de acesso à mistura de grãos se ambos cooperassem (C-C); 6 s, se o Jogador A delatasse e B cooperasse (D-C); e 1 s, se vice-versa (C-D); e, finalmente, 2 s em caso de delação mútua (D-D). Se um pombo delatasse na tentativa anterior e depois cooperasse na tentativa corrente (D→C), ganhava 1 s de acesso à comida e 5 s de *blackout*. Se o pombo delatasse na tentativa anterior e na corrente (D→D), o reforçador eram 2 s de acesso e 4 s de *blackout*. Se o pombo cooperasse nas duas tentativas, anterior e corrente (C→C), ganhava 5 s de acesso à comida e 1 s de *blackout*. Por fim, se o pombo cooperasse na tentativa anterior e delatasse na atual (C→D), o ganho era de 6 s de acesso ao alimento. Depois de uma das alternativas ser

reforçada, era iniciado um ITI sem sinalização com um *blackout* da caixa, cujo tempo foi manipulado. Cada sessão foi finalizada com 30 tentativas.

A Fase 1 do experimento avaliou a preferência lateral. Nessa fase, o ITI foi de 9 segundos fixos. A chave de cooperação variou em três condições (em 12 sessões era apresentada randomicamente em ambos os lados, em sete sessões apenas na esquerda e em sete sessões apenas na direita).

Na Fase 2, a chave de cooperação foi apresentada de forma aleatória em ambos os lados, sendo que o ITI variou em três condições: 0 s, 18 s e padrão (em que, repetidamente, ocorriam três tentativas rápidas seguidas por 18 s). Para cada condição, desta e da próxima fase, foram realizadas um mínimo de 24 sessões.

Na Fase 3, a chave de cooperação foi apresentada apenas em um lado durante a sessão corrente e, depois, foi alternada entre as sessões seguintes. Esse procedimento objetivou confirmar a preferência pela chave de cooperação em vez de um controle pelo lado da Fase 2. Nessa fase, os pombos teriam de bicar nas duas chaves para cooperar, ao contrário da Fase 2. O ITI variou como na Fase 2, e, numa preferência lateral, o resultado sofreria uma mudança drástica na taxa de reforçamento de uma sessão para outra, pois os sujeitos não estariam expostos à contingência do TFT.

Na última condição da Fase 3 (delação), a duração do ITI foi de 0 s, foi retirado o esquema TFT, o computador somente delatava (a luz de teto sempre respondeu lentamente depois da escolha), e o pombo somente recebia um reforçador da matriz de pagamento. Durante essa fase, a delação maximizava o reforçamento tanto no tempo curto como no longo, já que o computador que respondia delatava em todas as vezes. Pombos foram expostos a essa fase para verificar se eram sensíveis às contingências de pagamento (Baker & Rachlin, 2002).

Os resultados na Fase 1 indicaram forte controle lateral nos pombos 4, 5 e 6. Os outros pombos, 1 e 2, preferiram cooperar, e o pombo 3 escolheu delatar. Os resultados na Fase 2 não demonstraram diferenças de porcentagem média entre escolhas de cooperação durante a primeira e segunda exposição de 0 e 18 s de ITI, o mesmo acontecendo na Fase 3. Na comparação entre as Fases 2 e 3, revelou-se um nível de cooperação significativamente alto, mas não houve diferenças significativas entre as condições de ITI 18 s e padrão; porém, numa análise de medida repetitiva de variância, mostrou-se uma diferença significativa no nível de cooperação, sendo que o ITI de 0 s teve maiores índices de cooperação do que o ITI de 18 s e padrão. Ademais, o nível de cooperação diminuiu durante a condição de delação (Baker & Rachlin, 2002).

Conclusão: o estudo demonstrou que, quando as escolhas são de reciprocidade, os pombos podem aprender a cooperar (apresentar autocontrole) no Dilema do Prisioneiro Iterado (IPD). Além disso, o nível de cooperação é afetado pelo atraso temporal das escolhas, ou seja, nível de cooperação diminui com o aumento do desconto temporal,<sup>3</sup> o que sugere que a solução de dilemas e problemas de autocontrole pode depender de evidenciar eventos temporariamente estendidos.

Segundo Stephens, McLinn e Stevens (2002), Axelrod e Hamilton (1981) argumentam que, no Dilema do Prisioneiro, a repetição e a reciprocidade se combinam para tornar a cooperação mútua uma estratégia viável, porque, embora um delator receba um reforço imediato, ele sofre por essa escolha a longo prazo na reciprocidade. Uma possível explicação para a fragilidade da cooperação no Dilema do Prisioneiro e a falta de sucesso empírico em animais não humanos aplicando a estratégia TFT em laboratório

---

<sup>3</sup> Desconto temporal: depreciação do valor de um reforçador relacionado ao tempo que leva para ser liberado.

seria o desconto temporal muito forte. No Dilema do Prisioneiro Iterado, os pagamentos a longo prazo são maiores na cooperação; porém, estudos de autocontrole mostram preferências por reforçadores imediatos (impulsividade). Outra explicação para a falha da cooperação seria a falta de aplicação de uma estratégia apropriada, uma estratégia em que há aumento de reciprocidade visando a uma maior probabilidade de cooperação futura, sem o que não há razão para cooperar mesmo sem o desconto temporal. Por isso, em seu estudo, Stephens et al. (2002) manipularam o efeito do atraso do reforço e a estratégia para avaliar a contribuição do desconto e da estratégia de reciprocidade no equilíbrio cooperativo.

Foram utilizados 16 pássaros Gaio azul (*Cyanocitta cristata*) de sexo indefinido, divididos em pares. Não houve privação nesse estudo. O aparato era uma câmara dividida em dois compartimentos em “V” lado a lado com três poleiros cada, e os poleiros eram ligados a um computador. No início de cada tentativa, uma luz era ligada, e o pássaro voava para um dos dois poleiros à sua frente, indicando sua escolha. Para manipular o efeito do atraso do reforço, os pássaros deviam completar uma sequência de quatro tentativas para obter o reforço acumulado dessa sequência em um comedouro transparente (podiam ver a comida ganha, mas não consumir), e, nas tentativas não acumulativas, a comida era dispensada após cada tentativa. Para manipular o efeito das estratégias, um dos pássaros era designado sujeito, e o outro, *stooge*. O sujeito era livre para escolher entre cooperar e delatar; o *stooge*, no entanto, era forçado a jogar delatando sempre ou aplicando estratégia TFT, de acordo com uma estratégia pré-programada (Stephens et al., 2002).

Os resultados mostraram que, jogando com um oponente delator, a frequência de cooperação diminuía quase a zero, não importando se havia acúmulo de comida. Quando o jogo foi contra um oponente que respondia em reciprocidade, os resultados foram altas

taxas de cooperação; entretanto, no jogo acumulado, as taxas de cooperação foram elevadas e estáveis, e, no jogo não cumulativo, as respostas começaram na cooperação, mas declinaram até o nível dos resultados do oponente que sempre delatava. Os resultados desse estudo demonstraram que o desconto (atraso no reforço) e a estratégia (reciprocidade ou não) afetam a estabilidade da cooperação; porém, destacam que a estratégia complementa o desconto temporal, uma vez que o momento do reforçamento pode ser o que determina a cooperação estável (TFT com reforçamento acumulado), ou a cooperação inicial que se transforma em delação mútua (TFT com reforçamento não acumulado (Stephens et al., 2002).

Estudos citados anteriormente utilizaram pássaros como sujeitos de experimentos. A seguir, descreveremos estudos que utilizaram ratos como sujeitos experimentais.

Gardner, Corbin, Beltramo (1984) investigaram cooperação e competição por meio de um Jogo do Dilema do Prisioneiro em pares de ratos. Os sujeitos foram 12 pares de ratos fêmeas com 80% de peso *ad libitum* por todo o experimento. O aparato consistia em dois labirintos em “T” de madeira separados por divisória em plexiglás e portas deslizantes em plexiglás divididas em box de partida (*start box*), box de segurança (*holding box*) e box de pontuação (*goal box*). O reforço deu-se com pelotas de 45 mg de ração. Matriz de pagamento no estudo: C-C (cooperação mútua) = 3-3 (três pelotas para cada um); C-D<sup>4</sup> (cooperação unilateral) = 0-5 (zero pelota para quem coopera e cinco pelotas para o delator); D-C (delação unilateral) = 5-0; e D-D (delação mútua) = 1-1.

---

<sup>4</sup> No experimento, os autores simbolizam a resposta não cooperativa (delação) como “U” (*uncooperative response*).



Os ratos inicialmente passaram por duas sessões de 24 tentativas, com o objetivo de adaptação ao labirinto. Eram liberados na *start box* na base do “T” e podiam escolher qualquer corredor. Ao chegar a um dos lados (esquerdo ou direito na *holding box*), teriam uma recompensa randomizada de 50%, ou seja, seriam recompensados em metade das tentativas ao escolherem qualquer um dos lados.

Na fase seguinte (fase de treino com a matriz de pagamento), foi determinado que o lado direito seria resposta de cooperação, e o lado esquerdo seria resposta de delação para metade dos ratos; para a outra metade, o significado seria o inverso. Somando-se a isso, metade dos pares foram condicionados a respostas contíguas (do mesmo lado do labirinto), significando C-C e D-D; e outra metade, resultado das respostas em C-D e D-C (para a mesma resposta). Esse procedimento visou a eliminar qualquer preferência pela posição e eliminar a tendência de um rato seguir a resposta do outro por companhia ou imitação. Os animais foram liberados simultaneamente na *start box*. Apenas uma das portas era aberta de um dos lados, forçando a resposta nessa fase, sendo os ratos expostos a todas as matrizes de pagamento (C-C, D-D, D-C e C-D). Assim que o animal entrava na *holding box*, o pagamento era depositado na *goal box* e liberado logo em seguida. Cada par de ratos teve três dias de sessão com 24 tentativas cada (Gardner et al., 1984).

No teste de cooperação, as portas de ambos os lados foram abertas, e os animais foram liberados simultaneamente. O primeiro a responder ficava na *holding box* até o oponente fazer sua escolha, e ambos ficavam na *holding box* até dispensar a recompensa. Pares foram testados 24 tentativas/dia por 10 dias. O plexiglás foi, então, substituído por um cartão não transparente até o *goal box*, e houve testes por mais 10 dias de 24 tentativas cada. Esse procedimento visou a observar o comportamento dos ratos sem a comunicação visual entre ambos. O cartão somente era liberado quando os ratos entravam em suas

respectivas *holding boxes* e, assim, podiam verificar o resultado de ambos pela posição de seu companheiro.

Dados foram analisados por pares. Respostas de mútua delação (D-D) foram predominantes com diferença significativa nos primeiros 10 dias em relação à cooperação mútua (C-C). As respostas de delação tornaram-se acentuadas com a barreira de cartão. Respostas C-D, D-C e C-C não apresentaram diferenças significativas. Ao comparar D-D e D-C quando o primeiro a responder não cooperava, a resposta de delação na tentativa seguinte era preponderante. Quando a resposta de colaboração vinha em primeiro, as respostas foram iguais no começo, mas a delação aumentava sensivelmente após o oitavo dia (Gardner et al., 1984).

Gardner et al. (1984) demonstram que ratos são sensíveis à resposta de um oponente, respondendo de acordo com o pagamento envolvido e regras estabelecidas. Em estudos como os de Scodel, Minas, Ratoosh e Lipetz (1959) e Minas, Scodel, Marlowe e Rawson (1960) com humanos jogando Dilema do Prisioneiro Iterado, as taxas de cooperação também decresceram com o passar do tempo, algo semelhante ao encontrado no estudo de Gardner et al. (1984), ao retirar a interação; porém, a barreira visual causou maior declínio das respostas cooperativas no início e aumento leve (não significativo) no decorrer da sessão. Durante a interação (plexiglás), animais faziam uma espécie de dança um para o outro, como que tentando influenciar a resposta do outro, o que durava de 15 a 20 s. O mesmo acontecia quando o rato que escolhia primeiro e escolhia cooperar. Caso o segundo rato fosse em direção à resposta de delação, o rato na *holding box* tentava abrir a porta para escapar, o que não acontecia quando a resposta era de cooperação mútua. Os animais nem checavam se havia comida nas escolhas D-C/C-D. Quando a tampa era removida muitas vezes, o rato pulava para o outro labirinto para

agredir o outro rato. Esse comportamento não era observado nas respostas de cooperação e delação mútuas.

**Dilema do Prisioneiro em sujeitos não humanos (pesquisas em biologia evolutiva).** O estudo de Viana, Gordo, Sucena e Moita (2010) propôs estudar as interações entre ratos no Dilema do Prisioneiro Iterado, em que a perda de reforçadores por cooperar (custo) e a escolha no reforçador imediato (tentação) estavam presentes, manipulando estratégias TFT e pseudorandomizada.

Nas fases do experimento que envolviam privação de comida, ratos eram mantidos com 85% do peso. O aparato consistia em um duplo labirinto em “T” disposto frente a frente; cada labirinto consistia em uma pequena caixa de partida que dava acesso à uma porta deslizante que conduzia a dois compartimentos. Cada compartimento era separado por uma divisória móvel, sendo que a resposta era considerada quando o rato colocava todas as quatro patas no compartimento escolhido por ele. As paredes opostas ao labirinto em “T” e a divisória eram transparentes e perfuradas (assim, os ratos podiam ver-se e cheirarem um ao outro em todo o labirinto) e as demais paredes eram pintadas de preto. Como reforçadores, foram utilizadas pelotas de ração.

A matriz de pagamento utilizada nesse experimento foi de: tentação (delação unilateral) = seis pelotas; recompensa (cooperação mútua) = quatro pelotas; *sucker/simplório* (cooperação unilateral) = três pinçadas na ponta do rabo; *punishment*/penalidade (delação unilateral) = uma pinçada na ponta do rabo. A razão de se utilizar um método com estímulo aversivo foi baseada nas conclusões de Stephens e Clements (1998), de que matrizes de pagamento heterogêneas, ou seja, em que haja reforçador (respostas em cooperação mútua e delação unilateral) e delação mútua e cooperação unilateral gerando punição, levariam a taxas de cooperação maiores comparadas a pagamentos homogêneos, nos quais há somente reforçamento.

Na fase de treinamento, Viana et al. (2010) utilizaram apenas um lado do labirinto para um simples teste de preferência. Inicialmente, foi feito um treino para determinar se os ratos podiam discriminar os valores de reforço e magnitude de punições. Primeiramente, entre quatro e seis pelotas e entre um e três pinçamentos de cauda. Para reforçamento positivo, o rato escolheu as seis pelotas em vez de quatro em 80% no último dia dos cinco, com uma sessão de 20 tentativas. Para a punição, a escolha por um pinçamento foi de 59%, em vez da de três pinçamentos.

Na fase experimental com o Jogo do Dilema do Prisioneiro, os compartimentos do labirinto foram arbitrariamente definidos como C (cooperação) e D (delação), alternando-se aleatoriamente entre os ratos: se, para um rato, em C, ele cooperava e, em D, ele delatava, para o rato seguinte, os compartimentos eram invertidos. Isso garantiu que pistas de odores do rato anterior pudessem eliciar respostas opostas para o rato seguinte (cheiro de urina ou fezes podiam indicar sinal de medo ou perigo, fazendo o rato seguinte evitar o compartimento).

Foram utilizados cinco a seis ratos denominados *alvos*, e, para cada jogo, um dos ratos, denominado *stooge*, era designado a jogar uma estratégia fixa: TFT (cooperava na primeira e imitava respostas subsequentes), ou pseudorandomizada (estratégia randomizada seria como jogar “cara ou coroa”; na pseudorandomizada, não poderia haver mais de três respostas iguais seguidas), ou seja, em cada tentativa, o rato *stooge* era colocado nos compartimentos C ou D de acordo com a estratégia definida pelo próprio experimentador. O outro rato, o rato-alvo, era colocado no box de partida e podia escolher livremente entre os compartimentos C e D. Antes da primeira sessão, houve uma fase de habituação de três dias sem o rato *stooge* e cinco dias com o rato *stooge* (5 min/dia), apenas para acostumar o rato às mudanças. Uma vez que as divisórias eram transparentes e perfuradas, o rato-alvo podia usar o *stooge* como sugestão para guiar sua

escolha, enquanto o rato *stooge* era, simplesmente, colocado num dos compartimentos à frente dos compartimentos do alvo, de acordo com a estratégia que se queria empregar: por exemplo, para fixar a estratégia TFT no rato *stooge*, o rato era colocado no compartimento C na primeira tentativa e, nas demais tentativas, em C ou D, de acordo com as regras do TFT – repetir a resposta anterior do oponente. Feita a escolha, a divisória era fechada, e o reforçador era entregue de acordo com a matriz de pagamento. Caso o rato não escolhesse em 30 s do início da tentativa, o experimentador fechava vagarosamente a divisória, forçando o rato-alvo a escolher um dos compartimentos e finalizando a tentativa, que, depois, era iniciada com um rato diferente. A média de intervalo entre tentativas era de 4 a 5 min (Viana et al., 2010).

Não houve, como se esperava, cooperação significativa na primeira sessão como resultado da estratégia de TFT imposta; contudo, a partir da segunda sessão, uma cooperação significativa emergiu e foi mantida até o final do jogo, em um resultado similar ao encontrado em humanos. Houve mais respostas de cooperação mútua (C-C) do que de delação mútua (D-D) e considerável incidência de respostas delação e cooperação unilateral D-C (tentação) e C-D (simplório), sendo que o resultado mais baixo foi de delação mútua, D-D.

Esses resultados são explicados pelo fato de que, após um movimento de delação, a probabilidade de cooperação era alta, e, depois de um movimento de cooperação, a probabilidade de cooperação era a mesma da delação, conforme dados obtidos nesse experimento (Viana et al., 2010). Em outras palavras, a probabilidade de repetir a delação era muito baixa, levando a resultados de delação mútua infrequentes. Em contraste, depois de haver cooperação, era mais provável manter a cooperação mútua ou delatar para obter resultado em delação unilateral (tentação) na outra tentativa. Em resumo: ratos cooperam mais do que delatam e, quando delatam, rapidamente reverterem para

cooperação, evitando um ciclo de delação mútua. Após um treino inicial, as estratégias utilizadas pelos diferentes ratos eram notavelmente similares. Quando o oponente retribuía, o comportamento dos ratos era de cooperação mútua ou de reciprocidade alternada (alternância entre tentação e *sucker*/simplório) (Viana et al., 2010).

Para testar uma estratégia de não reciprocidade, uma vez que as taxas de cooperação são contingentes a um comportamento de reciprocidade, os ratos foram testados em uma estratégia pseudorandomizada, na qual as escolhas do rato-alvo não influenciavam os movimentos subsequentes. Na estratégia pseudorandomizada, o *stooge* cooperava em 50% das tentativas, sendo que a sequência de cooperação e delação foi randomizada, sem repetir, porém, a mesma resposta mais de três vezes consecutivas. Num Dilema do Prisioneiro randomizado e único (não iterado), a melhor estratégia é delatar – e, de fato, na estratégia pseudorandomizada, o rato delatou na maioria das tentativas. Isso corrobora as conclusões da estratégia TFT, que determina que os ratos cooperem dependendo da estratégia do oponente e que a reciprocidade é necessária para a emergência e sustentabilidade da cooperação (Viana et al., 2010).

Os resultados da estratégia pseudorandomizada demonstraram, ainda, que os comportamentos dos ratos são sensíveis à estratégia de seu oponente, ainda que abaixo do esperado (comportamento adotado não produz o melhor ganho possível). Na TFT, haveria mais cooperação mútua (C-C) e nenhum resultado em delação mútua (D-D), e, na estratégia pseudorandomizada, os ratos deveriam sempre delatar, maximizando a recompensa e minimizando os resultados em delação mútua (D-D) (Viana et al., 2010).

O estudo de Viana et al. (2010) analisou, ainda, se os ratos adaptaram seu comportamento aos termos econômicos da matriz de pagamento. Num Jogo do Dilema do Prisioneiro Iterado, o pagamento maior e imediato resulta de uma tentativa de delação. No entanto, o maior ganho ao longo das sessões é alcançado quando ambos os jogadores

cooperam. Sendo assim, na estratégia TFT e com pleno conhecimento dos movimentos do adversário, o jogador-alvo tende a cooperar sempre, porém a tendência à delação unilateral pode ser alta após uma tentativa anterior de cooperação mútua. A hipótese é de que os ratos adotam uma estratégia mista (cooperação mútua e reciprocidade alternada), na qual a cooperação mútua implica um pagamento maior do que a reciprocidade alternada pura, mas a tentação para delatar é significativa por sua recompensa imediata. Para testar isso, diminuiu-se o resultado da cooperação mútua (de quatro para duas pelotas), enquanto se manteve o nível de delação unilateral (seis pelotas). Os resultados confirmaram a previsão, com achados de delação unilateral similares à primeira fase do experimento. Os níveis de cooperação mútua, por seu turno, foram significativamente menores do que os de delação mútua, mostrando que os termos econômicos são percebidos pelos ratos. Segundo os autores, isso significa que os ratos são sensíveis à discriminação numérica. Além disso, dados sugerem que os ratos ficam sob controle do reforçamento do jogo, pelo menos até a tentativa anterior, uma vez que a cooperação somente surgiu com as seguintes condições: (a) o oponente jogou TFT (que dependia dos movimentos da tentativa anterior); e (b) quando o pagamento para cooperação mútua era maior do que para reciprocidade alternada, e isso somente foi percebido após duas ou mais tentativas acumuladas. Esses resultados reforçam os achados anteriores, de que ratos são capazes de reciprocidade direta, que, segundo os autores, requerem memória de interações anteriores, mostrando ainda que eles têm capacidade cognitiva para se engajarem em uma estratégia de reciprocidade (Viana et al., 2010).

O estudo de Viana et al. (2010) também testou se o comportamento dos ratos era contingente ao estado motivacional. Nas fases anteriores, os animais tinham livre acesso à alimentação, o que pode explicar o sucesso na observação de cooperação nesse estudo, que contrasta com estudos anteriores, que haviam demonstrado que a impulsividade e o

desconto temporal desempenhavam um papel importante na taxa de cooperação. Além disso, importa dizer que os estudos anteriores utilizaram animais moderadamente privados. Para testar essa possibilidade, foi utilizada uma privação moderada de alimento 85% do peso *ad libitum*. Os resultados confirmaram a previsão, e ratos moderadamente privados de alimento, mesmo jogando contra TFT, diminuíram a taxa de cooperação comparado aos níveis das fases anteriores (Viana et al., 2010).

O estudo a seguir, diferentemente do anterior, concebeu um aparato em que as respostas produzidas pelos ratos foram mais condizentes com o teorema de Nash, segundo o qual os sujeitos experimentais não devem visualizar as escolhas de seus parceiros. Com respostas simultâneas e rápidas, os animais somente vislumbravam a escolha do parceiro ao receberem os pagamentos.

Em Wood, Kim e Li (2016), foi utilizado o Jogo Dilema do Prisioneiro Iterado (IPD) e teste de reciprocidade em ratos, no qual foram testados: sexo, *status* dominante, familiaridade e saciedade com estratégias TFT e Pavlov.<sup>5</sup> A hipótese dos autores era que respostas de cooperação seriam maiores em ratas fêmeas, ratos subordinados, ratos bem-alimentados e ratos com maior grau de parentesco.

Nesse estudo, foram utilizadas caixas de condicionamento operante divididas ao meio com telas transparentes e removíveis, de modo que cada um dos ratos podia observar o que o outro fazia. Foram utilizados ratos com seis semanas de vida, mantidos em baixo nível de crescimento (3 a 4 g de ganho de peso diário), sendo 32 machos e 16

---

<sup>5</sup> Também chamado de *win-stay-lose-shift*. O sujeito responde por reforçamento positivo e punição. Dada duas escolhas, ele inicialmente responde randomicamente; depois, repete a última escolha se o resultado for bom, dentro de um cálculo realizado (Kraines & Kraines, 1989).



fêmeas. Após modelar a resposta de pressão à barra, os sujeitos foram treinados individualmente no seguinte procedimento: cada sessão iniciava-se com a caixa escura e sem barra. A caixa era iluminada por 2 s, e a barra retrátil era inserida; o rato tinha, então, 10 s para acionar a barra e ganhar a pelota de sacarose. Após esse tempo, a barra era retirada e acendia-se a luz do teto por 30 s (intervalo entre tentativas – ITI). Caso o animal não acionasse a barra, a mesma retraía-se, e seguia-se um intervalo entre tentativas. Nesse caso, a tentativa era contada como omissão. O tempo para responder foi gradualmente diminuído para 5 e 2 s. O critério de encerramento desse treino foi de 25 respostas por 20 min de sessão (35 tentativas/dia, resultando em aproximadamente 70% de respostas), por dois dias consecutivos. Logo em seguida, os ratos foram testados aos pares em quatro condições para o Jogo do Dilema do Prisioneiro Iterado. Os testes eram conduzidos em sessões diárias de 25 tentativas, que se iniciavam com estímulo luminoso por 2 s. A seguir, a barra era inserida, e os ratos tinham 2 s para a resposta de pressionar a barra. Após 2 s, a barra era retirada, e a luz da caixa era acesa por 30 s. Em cada tentativa, o sujeito podia cooperar ou delatar (Wood et al., 2016).

Na cooperação unilateral (apenas um rato pressionava a barra, entregando cinco pelotas para o parceiro), na cooperação mútua e na delação mútua, a resposta era inversa dependendo do grupo de sujeitos. Ratos RM- e RF- foram treinados a dar resposta de cooperação *não* pressionando a barra, enquanto ratos RM+ foram treinados a cooperar pressionando a barra para eliminar qualquer viés do treino discriminativo, no qual os sujeitos são treinados a acionar da barra para ganhar reforço (o controle devia ficar sob a matriz de pagamento, e não da resposta). Pelotas eram entregues a cada 0,5 s com um som de *clicker*, de forma que ambos os ratos podiam reconhecer a liberação das pelotas. Os ratos recebiam três pelotas por cooperação mútua, cinco pelotas por delação unilateral e nenhuma pelota para cooperação unilateral e delação mútua.

Foram realizados vários testes. Entre eles, foram destacados os de restrição alimentar, dada a relação com o presente estudo. Outros testes incluíram: sexo, reciprocidade direta, dominância e o fato de o par ser companheiro de gaiola ou rato não familiar. Os ratos foram testados durante 10 dias, sendo que, nos últimos quatro dias, os sujeitos estavam em regime de restrição alimentar.

Wood et al. (2016) analisaram as respostas consideradas de cooperação mútua (resultado de recompensa) para os três grupos de ratos (RM+, RM- e RF-). Na análise geral dos dados, ratos machos e fêmeas que tinham sido treinados a cooperar não pressionando a barra (RM- e RF-) escolheram menos tentativas de cooperação mútua nos 10 dias de teste em comparação ao grupo cuja resposta de cooperação era pressionar a barra (RM+), pois precisavam *não* responder para cooperar e tinham sido treinados para responder à barra a fim de receber pelotas. Nos últimos quatro dias, houve aumento significativo para machos (RM-) nas respostas de cooperação mútua (quando ambos não pressionavam a barra). Fêmeas (RF-) também aumentaram o número de respostas de cooperação mútua, mas o aumento entre as fêmeas não foi significativo estatisticamente ( $p = 0,07$ ). Para os ratos expostos ao treino em condição invertida (RM+), ou seja, que pressionavam a barra para cooperar, nos quatro últimos dias, houve escolha em cooperar em metade das tentativas, número que excedia os dos dois grupos anteriores.

A análise de respostas em contraste com pelotas recebidas por dia em ratos dominantes e subordinados em condição *ad libitum* e com restrição alimentar concluiu que não houve diferença entre as condições para os grupos dominante e subordinado. Entretanto, a restrição alimentar influenciou diferentemente a frequência de resposta e recebimento de pelotas nos três grupos. Os ratos machos sob restrição alimentar responderam mais e receberam menos pelotas, significando mais respostas na delação. A cooperação foi reduzida sob restrição alimentar (Wood et al., 2016).

Já a análise de resultados de tentativas mostra que RM- e RM+ mostraram diferenças significativas sob alimentação *ad libitum* e restrição alimentar. Na condição *ad libitum*, no par RM-, os resultados de pagamentos em tentação (delação unilateral) e *sucker/simplório* (cooperação unilateral) foram significativamente maiores do que os de *punishment*/penalidade (delação mútua). Com restrição alimentar, a frequência de cooperação/delação unilateral em contraste com a delação mútua de ambos foi similar, mas os resultados de recompensa (cooperação mútua) foram infrequentes. Pares RM+ na condição *ad libitum* mostraram distribuição similar de resultados de tentativas: a frequência de resultados de delação mútua foi menor do que a de cooperação mútua ou cooperação/delação unilateral. Todavia, em RM+ com restrição alimentar, a frequência de resultados em delação mútua e cooperação/delação unilateral não diferiu.

A mudança de resultados na restrição alimentar em RM+ e RM- é consistente com o aumento na taxa da resposta, assim como há aumento na delação entre ratos para obter um pagamento alto com delação unilateral. Em *ad libitum*, cinco dos 32 machos receberam mais pelotas do que podiam receber em cooperação mútua. Com restrição, apenas três machos receberam mais. Em contraste, para as fêmeas RF-, não houve diferença no efeito da condição alimentar, apresentando resultados similares ao *ad libitum* (Wood et al., 2016).

Na análise da probabilidade de cooperação em resultados seguidos de R, T, S e P na tentativa prévia, RM+, RM- e RF- não seguiram regra TFT nem Pavlov. RM- foram os mais propensos a cooperar quando sem restrição de alimento, comparados com restrição, e também mais propensos a cooperar após uma tentativa prévia de cooperação do que de uma tentativa com delação. O resultado foi igual para RM+. Para RF-, a restrição não teve efeito na cooperação; entretanto, a probabilidade de cooperação

seguida de uma tentativa de cooperação foi significativamente maior do que na tentativa anterior com delação mútua.

O estudo de Wood et al. (2016) discutiu a utilização de reforçador primário e generalizado. Em humanos, utiliza-se dinheiro como reforçador, e essa condição interfere na impulsividade, uma vez que o dinheiro pode acumular-se em sucessivas tentativas. Por outro lado, o reforçador primário varia conforme a motivação (privação). A restrição alimentar utilizada como variável afeta a taxa de resposta de cooperação. Como essa variável também vai ser manipulada neste experimento, é pertinente um aprofundamento nos conceitos relacionados a esse tema.

**Privação: uma variável motivacional.** Entre as variáveis das quais o comportamento é função, incluem-se as variáveis motivacionais. A privação de alimento é condição primária para que o mesmo possa, como consequência de uma resposta, adquirir função reforçadora, assim como evocar comportamentos que, no passado, o produziram. A privação perturba a homeostase do organismo, e o comportamento reforçado tende a restaurá-la, o que permite prever em que “direção” o organismo vai alterar seu comportamento como resultado de uma variável independente.

Skinner (1953/2003) destaca a importância das variáveis motivacionais na análise do comportamento, sugerindo que estas são tão importantes quanto o esquema de reforçamento no controle do comportamento – e apresenta o conceito de *impulso (drive)*, que estabelece os limites e alcances conferidos à motivação na ciência do comportamento.

Para Keller e Schoenfeld (1950/1974), o *drive* seria mais adequadamente definido como um conjunto de relações entre o organismo e o ambiente. Demonstraram os autores que podemos tratar a variável motivacional como uma variável independente que pode ser manipulada de maneira experimental – e introduziram o termo *operação*

*estabelecedora*. Os autores ainda incluíram a estimulação aversiva como fator motivacional.

Millenson (1967/1975) destaca a privação de um reforço como uma operação fundamental para afetar o valor reforçador de reforçadores primários, entre outras operações (exercício forçado, estímulo cerebral, aplicação de medicamentos, etc.). Os reforçadores primários sofrem periodicidade cíclica no valor reforçador quando apresentados em estoque ilimitado. O autor afirma que a medida do valor reforçador de um estímulo é feita pela coleta de dados do efeito comportamental do mesmo, contingente à resposta operante, e cita diversas técnicas comportamentais para medir o valor reforçador primário dos estímulos: resistência à saciação, taxa de respostas reforçadas intermitentemente, resistência à extinção, aquisição, superar obstáculos, técnica do trabalho máximo, adulteração do reforçador e preferência, definindo-as todas como operações de estabelecimento de reforçador.

Michael (1993) distingue operações estabelecedoras de estímulos discriminativos. Define o autor operação estabelecedora como operação ou condição de estímulo que afeta um organismo alterando momentaneamente: (a) a eficácia como reforçador de outros eventos; e (b) aumentando a frequência de respostas do organismo relacionadas a esses eventos como consequências. O primeiro efeito é classificado como efeito estabelecedor, e o segundo, evocativo. A privação aumenta o valor reforçador e aumenta a frequência de respostas anteriormente reforçadas, evocando qualquer comportamento que tenha sido seguido pelo reforçador. Esse efeito evocativo é o resultado direto da operação estabelecedora sobre o comportamento, um aumento na eficácia de todos os estímulos discriminativos e um aumento na frequência de respostas condicionadas cuja eficácia depende da privação. Michael (1993) classifica as operações estabelecedoras em duas categorias: (a) operação estabelecedora incondicionada (OEI):

os efeitos de alteração da eficiência de reforçador que esse tipo de operação estabelecadora produz abrangem eventos ou estímulos reforçadores incondicionados (dependem da história de evolução de cada espécie); (b) operação estabelecadora condicionada (OEC):<sup>6</sup> abrangem os efeitos estabelecadores de reforço de reforçadores de origem ontogenética, isto é, relacionados à história de reforçamento de cada indivíduo.

De interesse neste estudo, são as operações estabelecadoras incondicionadas (de origem filogenética). Michael (1993) identifica vários tipos de OEI: (a) privação e saciação – privações de água, comida, sono e atividade servem como OEI, resultando em efeitos de reforço, estabelecedor ou evocativo; a saciação tem o efeito oposto da privação; (b) mudanças de temperatura – essa operação, quando está acima do nível normal das condições de conforto, tem como efeito estabelecedor do reforço a diminuição do nível da temperatura como um evento reforçador eficaz. O efeito evocativo é verificado pelo aumento momentâneo da frequência de qualquer comportamento que tenha sido reforçado pela diminuição do nível da temperatura, e vice-versa, quando a temperatura está abaixo dos padrões de normalidade. Essa OEI pode ser tratada sob a categoria de estimulação aversiva; (c) variáveis relacionadas ao reforçamento do comportamento sexual – características ambientais ou hormonais podem servir como OEI para manter contato sexual ou eliciar comportamentos que servem de OEI para o sexo oposto; (d) estimulação dolorosa (esquiva) – essa operação tem como efeito estabelecedor do reforço o aumento da eficácia momentânea da cessação ou da remoção do estímulo doloroso, e o efeito evocativo é demonstrado pelo aumento momentâneo de qualquer tipo de comportamento que tenha sido reforçado pela cessação ou pela remoção da estimulação

---

<sup>6</sup> As operações estabelecadoras condicionadas são classificadas em três tipos: (a) OEC substituta; (b) OEC reflexiva; e (c) OEC transitiva.

dolorosa. Nesse caso, a OE difere do estímulo discriminativo, pois a ausência do estímulo aversivo se assemelha mais à ausência de privação do que à ausência do estímulo discriminativo (sem o choque presente, a eliminação do choque não é uma forma efetiva de reforçamento). Michael (1993) sugere a utilização do termo apetitivo como o oposto de aversivo para manter o termo reforçamento específico para sua função de fortalecimento e como termo geral para eventos positivos. (e) Estimulação dolorosa (agressão) – para reduzir a estimulação dolorosa, o comportamento que produziu a redução na presença de outro organismo evoca comportamento agressivo contra esse organismo tendo como base fatores ontogênicos e filogênicos envolvidos; (f) outras operações estabelecidas emocionais – estimulação dolorosa evocando agressividade pode ser estendida para outras operações que produzem outros comportamentos.

Em Laraway, Snyckerski, Michael e Poling (2003), Jack Michael e colaboradores introduzem o conceito de *operações motivadoras*, refinando ainda mais os termos utilizados nos efeitos dos eventos motivacionais. Segundo os autores, o termo operação estabelecida apresenta limitações: implica apenas aumento da eficácia de uma consequência como reforçador ou punidor; contudo, muitas variáveis motivacionais *diminuem* essa eficácia. Michael (1993) reconheceu o problema e introduziu o termo *operação abolidora*, sendo que ambas – operações estabelecidas e operações abolidoras – são operações motivadoras, utilizadas no artigo como um termo geral. Estudos têm identificado os efeitos operações motivacionais sobre padrões comportamentais. Em Carlton (1962), restrição alimentar afetou o desempenho em respostas encadeadas. Em Senkowski, Vogel e Pozulp (1978), a restrição interferiu no estabelecimento de discriminação temporal de resposta em experimento que buscou estabelecer respostas de diferentes durações, e Souza (2007) verificou efeito de diferentes níveis de restrição alimentar sobre diferentes níveis de variabilidade .

## Objetivo

O objetivo deste estudo foi testar o efeito de diferentes níveis de restrição de alimento sobre o comportamento de cooperação de ratos em um Jogo do Dilema do Prisioneiro Iterado (IPD). Duplas de ratos machos receberam diferentes magnitudes do reforçador por cooperar ou delatar em tentativas repetidas, replicando sistematicamente o estudo de Wood et al. (2016), já apresentado anteriormente.

No estudo de Wood et al. (2016), foram manipuladas diferentes variáveis em um IPD: sexo, dominância, familiaridade e restrição alimentar. Cooperação foi identificada em altas taxas em ratos em condição *ad libitum*, diminuindo quando houve restrição alimentar.

O presente estudo buscou avaliar dois níveis de restrição e seus efeitos no comportamento de cooperação, considerando que outros estudos identificaram que níveis diferentes de restrição podem afetar padrões comportamentais. Buscou-se avaliar se níveis de restrição afetam diferentemente a cooperação, ou seja, se a restrição de alimento, variável independente manipulada, altera o valor do alimento como estímulo reforçador e, assim, modifica as respostas de cooperação em um Jogo do Dilema do Prisioneiro Iterado.



## Método

### Sujeitos

Seis ratos Wistar machos com idade aproximada de seis meses no início da experimentação, acondicionados isoladamente em condições controladas de temperatura ( $22^{\circ}\text{C} \pm 2^{\circ}\text{C}$ ), umidade ( $55 \pm 15\%$ ) e luminosidade (ciclo claro-escuro de 12 horas).

Na caixa-viveiro, os sujeitos, em condição *ad libitum* e sob restrição, foram alimentados com ração da marca Nuvital Quimtia CR-1 autoclavável. Na caixa experimental, os sujeitos receberam pelotas de 50 mg de sacarose (glóbulos homeopáticos n.º 5 inertes). Este estudo foi aprovado pelo Comitê de Ética no Uso de Animais (CEUA) da Pontifícia Universidade Católica de São Paulo (PUC-SP) sob o n.º 0022/2018.

### Aparato

Caixa de condicionamento operante de marca MedPlus com as seguintes dimensões: 40 cm de extensão x 17 cm de largura x 20 cm de altura, conectada à interface para registro das sessões, dividida ao meio com divisória em plexiglás (divisória transparente perfurada em toda sua extensão para que os ratos percebam a presença, comuniquem-se e sintam seus odores). Nas faces opostas, barras de acionamento em ambas as divisões fornecerão pelotas de 50 mg de sacarose em comedouro localizado ao lado da barra. Seguiu-se esquema proposto na Figura 1: luzes de estímulo (verde) localizadas no topo da parede (acima da barra de reforço) da caixa operante e luz de teto (branca) localizada no teto da caixa de isolamento acústico.

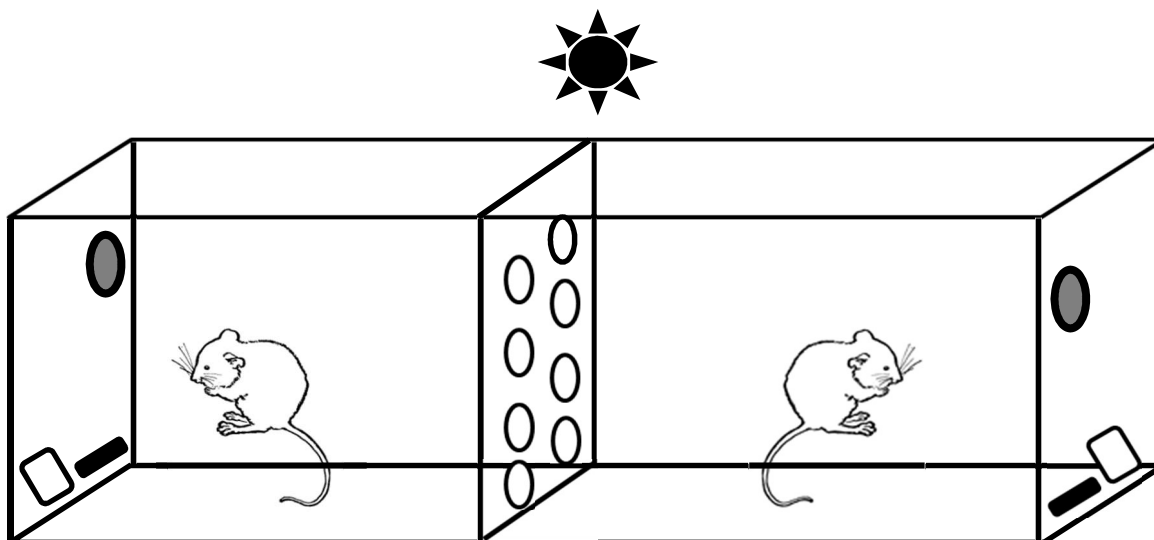


Figura 1. Desenho esquemático da caixa operante.

### Procedimento

**Restrição alimentar.** O procedimento de restrição alimentar seguiu o modelo de Souza (2007). Em uma primeira fase (do desmame ao 180.º dia de vida), os sujeitos ficaram com disponibilidade de comida e água livres para consumo e ficaram em gaiolas comunitárias com uma média de cinco sujeitos por gaiola.

A partir dos seis meses de vida, os sujeitos ficaram acomodados em gaiolas individuais (gaiolas de camundongo 30 x 19 x 13cm) com tampa alta.<sup>7</sup> Identificamos os sujeitos como Rato 1 a 10. O peso foi medido diariamente (exceto em finais de semana), para controle de peso *ad libitum*. O peso *ad libitum* foi calculado elaborando a média de

---

<sup>7</sup> A tampa alta se justifica como um enriquecimento ambiental, pois ficaram um período longo em isolamento. “Privação social ou interações sociais negativas não podem ser utilizadas. Quando espécies sociais necessitam permanecer fisicamente isoladas, o contato visual, auditivo e olfatório deve ser mantido” (Resolução Normativa n.º 33, de 18 de novembro de 2016, do Conselho Nacional de Controle de Experimentação Animal – CONCEA)

peso dos três últimos dias em que o sujeito atingiu estabilidade, antes do início da fase de restrição alimentar.

Os animais passaram pela fase de restrição alimentar 30 dias antes de entrarem nas fases experimentais. Dois níveis de restrições foram utilizados, sendo que, no experimento, foi utilizado um delineamento experimental ABACA (A = *ad libitum*; B= 70% do peso; e C = 90% do peso), para cada sujeito.

- (a) Sujeitos passam a receber inicialmente uma quantidade de alimento por dia correspondente a 70% do consumo médio diário de ração na condição *ad libitum*.<sup>8</sup> O peso foi medido, e a ração ajustada até atingirem a estabilidade de 70% do peso *ad libitum*, tendo livre acesso a água. A estabilidade de peso foi calculada a partir da fórmula:  $(I.^{\circ}D / M) - 1 \times 100$ , sendo  $I.^{\circ}D$  o dia de pesagem e  $M$  a média de peso dos últimos três dias. Quanto mais próximo de zero, o resultado estava mais próximo da estabilidade. Nessa fase, houve óbito de três animais (Ratos 1, 3 e 4);
- (b) Seguindo o mesmo esquema, os sujeitos tiveram o peso reduzido a 90% do peso *ad libitum* na condição C do esquema ABACA.

**Fase de condicionamento operante.** Privados a 70% do peso *ad libitum*, ratos foram colocados da caixa experimental e respostas de pressionar a barra tiveram como consequência pelotas de sacarose (automodelagem). Foram realizadas cinco sessões de reforçamento contínuo (CRF), em que a luz verde esteve acesa, sendo as sessões diárias,

---

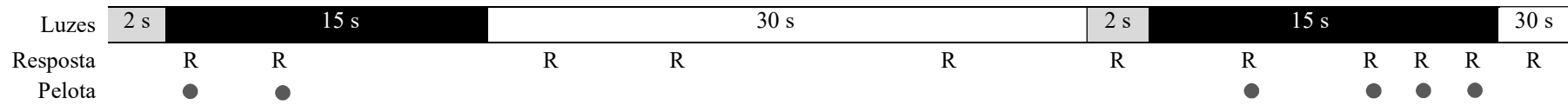
<sup>8</sup> Na condição *ad libitum*, eram administrados 50 g de ração diariamente e calculado no dia seguinte o consumo, subtraindo os gramas que restaram na grade do comedouro.

de 30 min ou com emissão de 35 respostas consecutivas. Nesta fase e nas subsequentes, os sujeitos foram colocados alternadamente em cada lado da caixa operante.

**Fase de treino discriminativo.** As fases seguintes foram baseadas no esquema de Wood et al. (2016). Ainda sob regime de restrição alimentar, os sujeitos foram colocados individualmente nas caixas operantes. O treino consistiu nas seguintes etapas:

- (a) Com a caixa totalmente escura, a tentativa se iniciava com a apresentação de uma luz verde que ficava acesa por 2 s. Com o apagar da luz verde, o sujeito tinha 15 s para responder. Caso animal pressionasse a barra, uma pelota era liberada no comedouro em CRF. Após os 15 s, a luz de teto acendia por 30 s e não havia reforçamento para as respostas de pressão à barra. Caso o sujeito pressionasse a barra no período de 10 s finais de luz de teto, a luz de teto se mantinha acesa por mais 10 s. Esse esquema foi introduzido para que respostas durante a luz branca não fossem reforçadas pelo aparecimento da luz verde, após a qual ocorria reforçamento. Após os 30 s de luz de teto, e caso não houvesse resposta com acréscimo de tempo, uma nova tentativa se iniciava. A sessão teve duração de 30 min. Esta etapa foi encerrada quando o sujeito atingiu índice de discriminação de, no mínimo, 75% em 03 sessões. O índice discriminativo foi calculado dividindo-se o número de respostas na luz apagada da sessão pela soma das respostas na luz apagada com as respostas durante a luz de teto e multiplicando-se por 100. O critério para encerramento das sessões foi por tempo: 30 min. O critério para encerrar essa etapa foi índice de discriminação em 70% em três sessões consecutivas

A Figura 2 ilustra as condições em que as respostas foram reforçadas.



*Figura 2.* Representação esquemática das condições em que as respostas produziam pelotas.

- (b) Na segunda etapa, o procedimento foi o mesmo. Entretanto, os 15 s iniciais foram diminuídos para 5 s no mesmo esquema CRF. Manteve-se também o critério para encerrar a etapa de índice de discriminação em 70% em três sessões consecutivas.
- (c) O término do treino discriminativo, na terceira etapa, deu-se com a diminuição dos 5 s para 2, reduzindo assim, gradativamente, o tempo em que a pelota podia ser produzida. O índice de discriminação foi de, no mínimo, 70% por três dias consecutivos. Caso um membro da dupla atingisse o critério antes do outro, sessões de treino nesta etapa foram mantidas até que ambos atingissem o critério.

**Fase de testes do Dilema do Prisioneiro.** Após o treino discriminativo individual, a restrição alimentar foi suspensa. Os sujeitos foram, então, testados em pares em sessões diárias de 25 tentativas cada. As sessões foram feitas três vezes por semana.

Os testes foram realizados em diferentes níveis de restrição, na seguinte sequência: (A) nenhuma restrição – *ad libitum*; (B) com restrição de 70% do peso *ad libitum*; (A) nenhuma restrição – *ad libitum*; (C) restrição de 90% do peso *ad libitum*; (A) nenhuma restrição – *ad libitum*.

Durante os testes, com a caixa escura, luz verde foi acesa por 2 s. Ratos tinham 2 segundos para pressionar a barra; após a pressão, a luz do teto se acendia por 30 s. Pelotas de sacarose foram entregues a cada 0,5 s para respostas no período escuro, segundo o esquema de pagamento.

Respostas de pressão à barra foram consideradas, neste estudo, como cooperação. A cooperação mútua (R) se deu quando ambos pressionaram a barra na mesma tentativa. Para isso, a consequência foi de três pelotas de sacarose para cada rato.

Quando ambos *não* pressionaram a barra na mesma tentativa, o resultado foi classificado como delação mútua (P). A consequência foi *nenhuma* pelota de sacarose.

Quando um dos ratos pressionava a barra e o outro não emitia resposta, o resultado foi de cooperação unilateral, denominado em outros estudos como *sucker* (S) ou tentação (T). A consequência foi de cinco pelotas para o que não emitiu resposta e nenhuma para o cooperador que pressionou a barra. Em cada tentativa, foram registradas respostas ou omissões de cada sujeito da dupla.

**Testes sem restrição.** Para cada dupla, nas primeiras 10 sessões, houve condições *ad libitum*, ou seja, sem restrição alimentar. Os dados destacados para análise são dos últimos quatro dias.

**Testes com restrição a 70% do peso.** A seguir, os sujeitos foram submetidos à restrição alimentar, com o mesmo procedimento de restrição anteriormente descrito, até atingirem 70% do peso *ad libitum*. O teste do Dilema do Prisioneiro foi mantido durante esse período, até que se atingisse a estabilidade de 70% do peso *ad libitum*, por quatro sessões consecutivas. Depois, novamente, os sujeitos experimentais foram alimentados livremente. Quatro testes do Dilema do Prisioneiro foram realizados nesse período.

**Testes com restrição a 90% do peso.** Na etapa seguinte, sujeitos experimentais sofreram restrição alimentar até atingirem 90% do peso *ad libitum* ao mesmo tempo que ocorriam as sessões do Dilema, até que os sujeitos da dupla atingissem a estabilidade de 90% do peso *ad libitum* por três dias de teste. Na última etapa, foram realizados quatro testes sem restrição.

Na Tabela 2, são apresentadas as fases do experimento e as condições de restrição.

Tabela 2

*Fases do Experimento e Condições de Restrição Alimentar*

		Restrição	Condição na caixa experimental
Condicionamento operante		70% do peso <i>ad lib.</i>	Individual
Treino discriminativo		70% do peso <i>ad lib.</i>	Individual
Testes do dilema do prisioneiro	1	Nenhuma	Dupla
	2	70% do peso <i>ad lib.</i>	Dupla
	3	Nenhuma	Dupla
	4	90% do peso <i>ad lib.</i>	Dupla
	5	Nenhuma	Dupla



## Resultados e Discussão

Os efeitos de diferentes níveis de restrição sobre o comportamento de cooperar no Jogo do Dilema do Prisioneiro Iterado foram avaliados em três duplas de sujeitos. Nas Figuras 3, 4 e 5, são representados os números de respostas de cooperação mútua, delação mútua e cooperação unilateral de cada membro da dupla, nas quatro últimas sessões das condições *ad libitum* e nos dois níveis de restrição avaliados das duplas S2 e S7, S5 e S10, S8 e S9, respectivamente. A média de respostas foi obtida somando-se o total de respostas de cada uma das quatro últimas sessões e calculando a média aritmética simples (divisão por quatro).

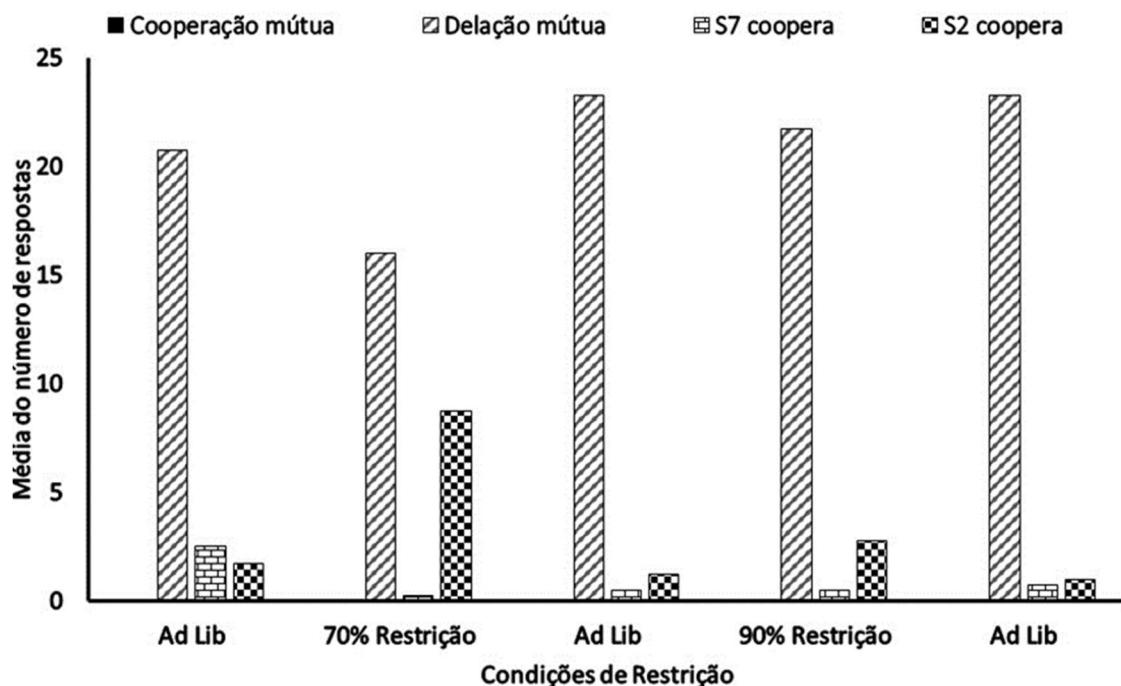


Figura 3. Média do número de tentativas classificadas como delação mútua, cooperação mútua e cooperação unilateral das quatro últimas sessões das fases de testes do Dilema do Prisioneiro dos sujeitos da dupla S2 e S7 em que vigoraram diferentes níveis de restrição alimentar.

Nos últimos quatro dias de teste com o Dilema do Prisioneiro, na primeira fase *ad libitum*, ambos os sujeitos da dupla S2 e S7 emitiram respostas classificadas como

cooperação unilateral, sendo que o S7 respondeu mais (média de 2,5 das tentativas das quatro sessões) do que o S2 (média de 1,75); porém, foram predominantes as tentativas classificadas como delação mútua (média de 20,75 respostas) e nenhuma resposta de cooperação mútua ocorreu.

Na fase de restrição de 70% do peso *ad libitum*, S2 aumentou consideravelmente a quantidade de respostas – cooperação unilateral, com média de 8,75 –, S7 praticamente não respondeu nesta fase (média de 0,25 respostas), e a delação mútua reduziu-se em relação à fase *ad libitum* (média de 16 respostas), mas ainda com predominância.

Na segunda fase *ad libitum*, a dupla S2 e S7 apresentou queda sensível de respostas de cooperação unilateral. O sujeito S2 respondeu poucas vezes (média de 1,25), assim como S7 (média de 0,5). O resultado foi predominantemente de delação mútua (média de 23,25).

Na fase de 90% de peso, houve aumento de resposta de cooperação unilateral de S2 (média de 2,75 das tentativas) em comparação com a fase *ad libitum* anterior, mas menor que a fase de restrição de 70% de peso *ad libitum*; S7 manteve níveis baixos de resposta (média de 0,5).

Na última fase *ad libitum*, as respostas ficaram perto de zero: (S2, média de 1; e S7, média de 0,75), sendo predominante a delação mútua das tentativas

A Figura 4 mostra o número de tentativas em que houve cooperação mútua ou unilateral e delação mútua em todas as sessões realizadas pela dupla S2 e S7 nas sessões das fases de teste do Dilema do Prisioneiro, nos diferentes níveis de restrição alimentar.

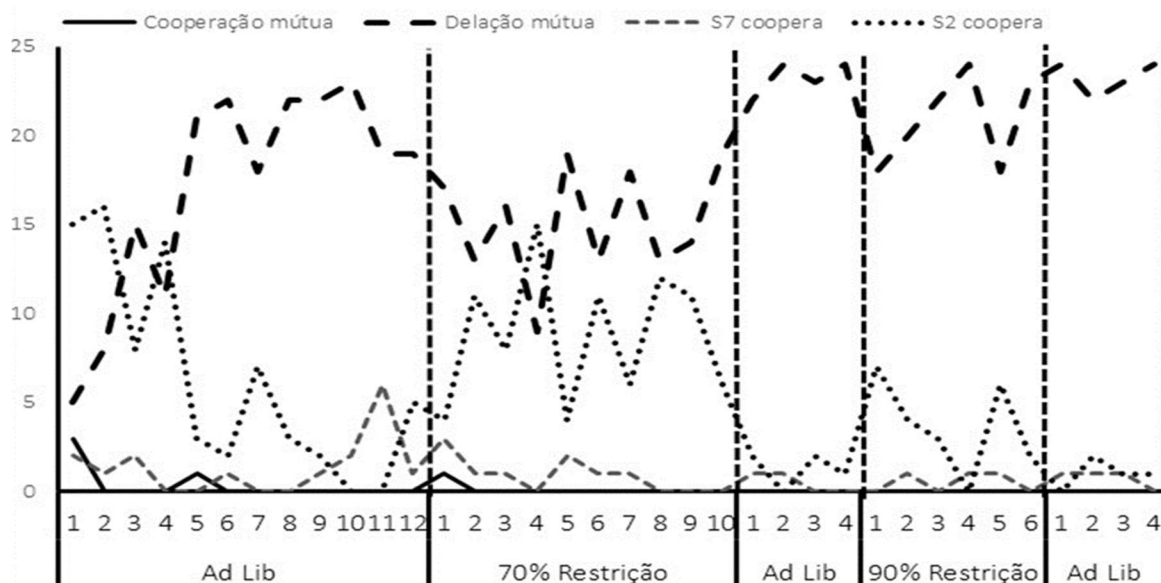


Figura 4. Número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S2 e S7 nas sessões das fases de teste do Dilema do Prisioneiro nas diferentes condições de restrição alimentar.

Fazendo uma análise geral de todas as sessões da dupla S2 e S7, houve um pequeno número de cooperação mútua no início do experimento, mas que logo desapareceu. A restrição alimentar afetou diretamente o número de respostas de apenas um sujeito da dupla – S2 –, aumentando o número de respostas quanto maior a restrição. Uma característica que se pode observar na Figura 4 é que o efeito da restrição alimentar se faz notar logo nas primeiras sessões (antes da estabilidade do peso proposto) e produz uma oscilação no número de respostas que não se mantém estável nas últimas quatro sessões analisadas.

A dupla S5 e S10 obteve um resultado semelhante à dupla S2 e S7, como pode ser observado na Figura 5. Na primeira fase *ad libitum*, também se apresentou preferência pela delação mútua (média de 21 respostas), nenhuma resposta de cooperação mútua e poucas respostas de cooperação unilateral, sendo que S10 respondeu (média de 2,5 respostas) mais do que S5 (média de 1,5 respostas).

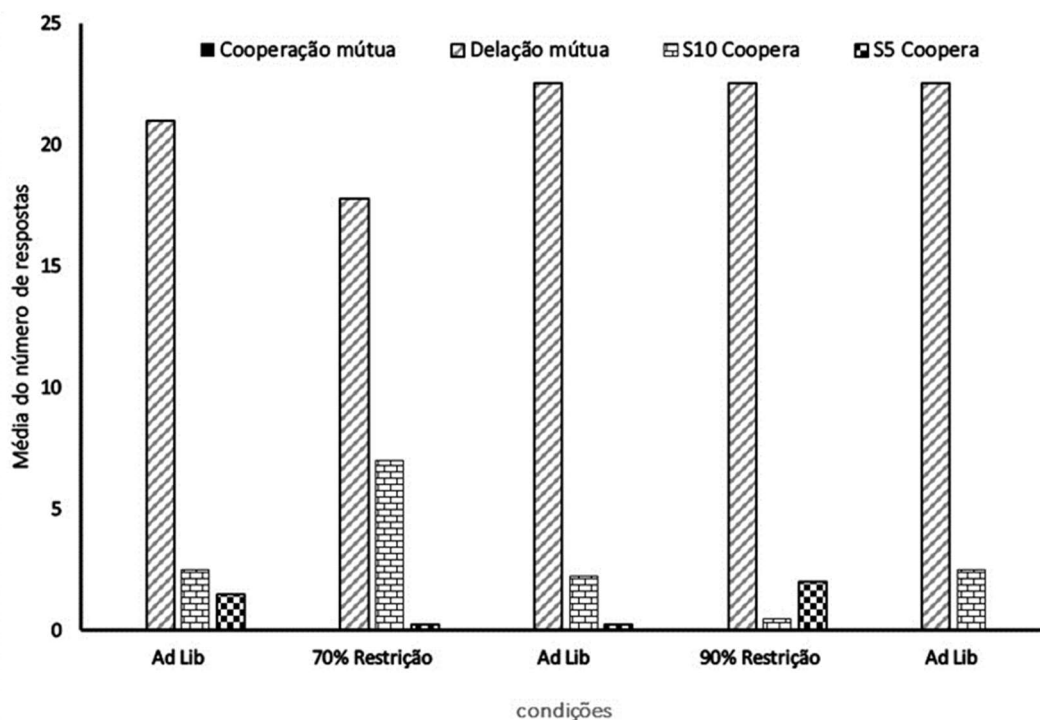


Figura 5. Média do número de tentativas classificadas como delação mútua, cooperação mútua e cooperação unilateral, das quatro últimas sessões das fases de testes do Dilema do Prisioneiro dos sujeitos da dupla S5 e S10 em que vigoraram diferentes níveis de restrição alimentar.

Na Fase de Restrição de 70% de peso *ad libitum*, houve aumento da resposta de cooperação de S10 (média de 7 respostas por sessão) e uma redução dessas respostas de S5 (média de 0,25). As respostas de delação mútua predominaram nessa fase (média de 17,75 respostas). Não houve respostas de cooperação mútua.

Na segunda fase *ad libitum*, ambos os sujeitos continuaram a não responder, aumentando as tentativas classificadas como delação mútua (média de 22,5); S10 diminuiu muito o número de respostas em comparação à fase anterior (de 7 para 2,25 em média, por sessão) e S5 respondeu, novamente, apenas uma vez nas quatro sessões analisadas (mantendo média de 0,25 respostas).

Na fase de 90% de peso *ad libitum*, predominaram tentativas de delação mútua (média de 22,5 das tentativas). Houve um leve aumento de respostas de S5 (média de duas respostas), enquanto S10 diminuiu suas respostas em relação às fases anteriores (média de 0,5 respostas).

Na última Fase *ad libitum*, S5 não respondeu, e S10 apresentou uma média de respostas maior (média de 2,5) comparado com a fase de 90% de peso *ad libitum*.

Na Figura 6, são apresentadas as tentativas em que houve cooperação mútua ou unilateral e delação mútua por S5 e S10, ao longo de todas as sessões de teste do Dilema do Prisioneiro.

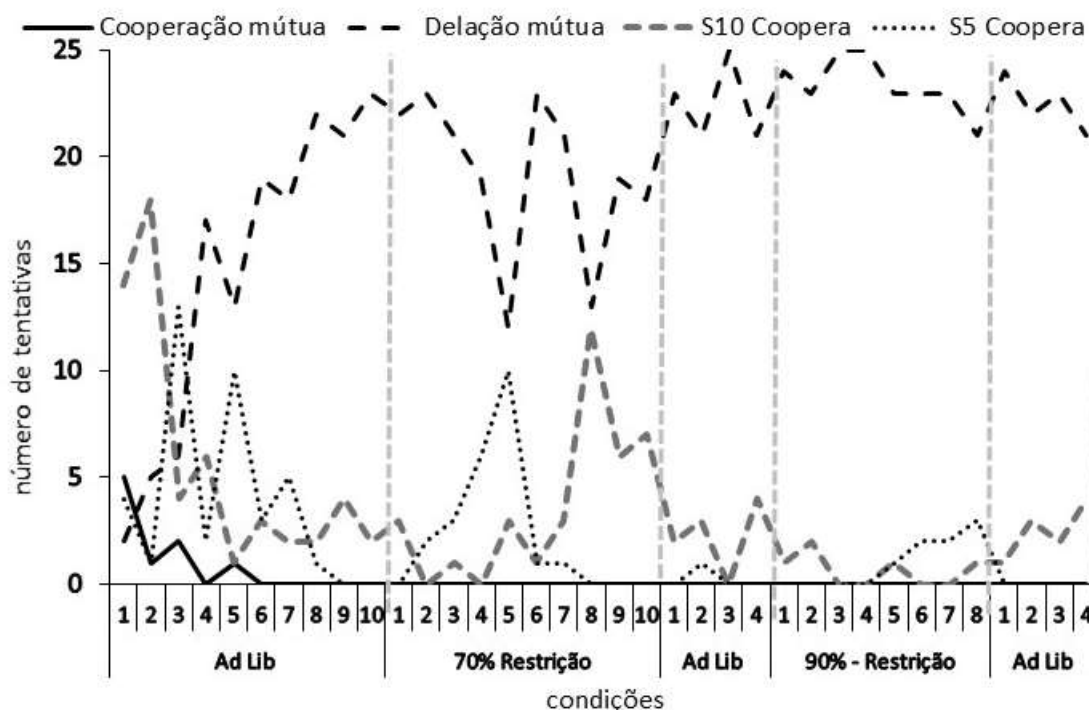


Figura 6. Número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S5 e S10 nas sessões das fases de teste do Dilema do Prisioneiro, nas diferentes condições de restrição alimentar.

Na análise geral de todas as sessões, S5 e S10 foi a dupla que mais respostas de cooperação mútua efetuou, mas também logo desapareceu (a partir da sexta sessão).

Na fase *ad libitum*, o número de respostas de cooperação mútua cai ao longo das seis primeiras sessões (de 5 para 0) e não mais ocorrem durante todo o experimento. Nessas sessões, respostas de cooperação unilateral se alternam entre os sujeitos, e, nas últimas quatro sessões desta condição, observa-se um aumento da delação mútua e uma diminuição de cooperação unilateral (S10 emitiu entre duas e quatro respostas, e S5 reduziu de cinco para zero tais resposta).

A fase de restrição de 70% de peso *ad libitum* inicia-se com uma alternância de respostas de cooperação unilateral, com um aumento no número de tentativas com respostas de S5 (chegando a 10 tentativas) na quinta sessão, seguida de uma queda abrupta. Ao compararmos as últimas quatro sessões da condição *ad libitum* com as da fase de restrição de 70%, em que o peso estava estável, observa-se um grande aumento no número de tentativas (12) em que S10 coopera unilateralmente, seguido por uma diminuição (7), mas ainda com valores acima da fase *ad libitum*.

Com o retorno à fase *ad libitum*, há uma diminuição de respostas de cooperação unilateral, retornando-se aos valores anteriores (entre dois e quatro), especialmente por parte de S10, sendo que a alternância se inverteu.

Nas sessões de restrição de 90% de peso *ad libitum*, há uma redução das respostas de S10 no início da restrição alimentar (sessões de estabilização do peso), mas, quando o peso se estabiliza, nas últimas quatro sessões, S5 aumenta um pouco o número de tentativas em que responde (chegando a três).

No final, apenas S10 voltou a responder na terceira fase *ad libitum*, como havia feito anteriormente nessa condição (entre uma e quatro respostas).

A média do número de tentativas em que dupla S8 e S9 cooperou mutuamente, unilateralmente e delatou nas quatro últimas sessões de cada condição dos testes do Dilema do Prisioneiro são apresentados na Figura 7.

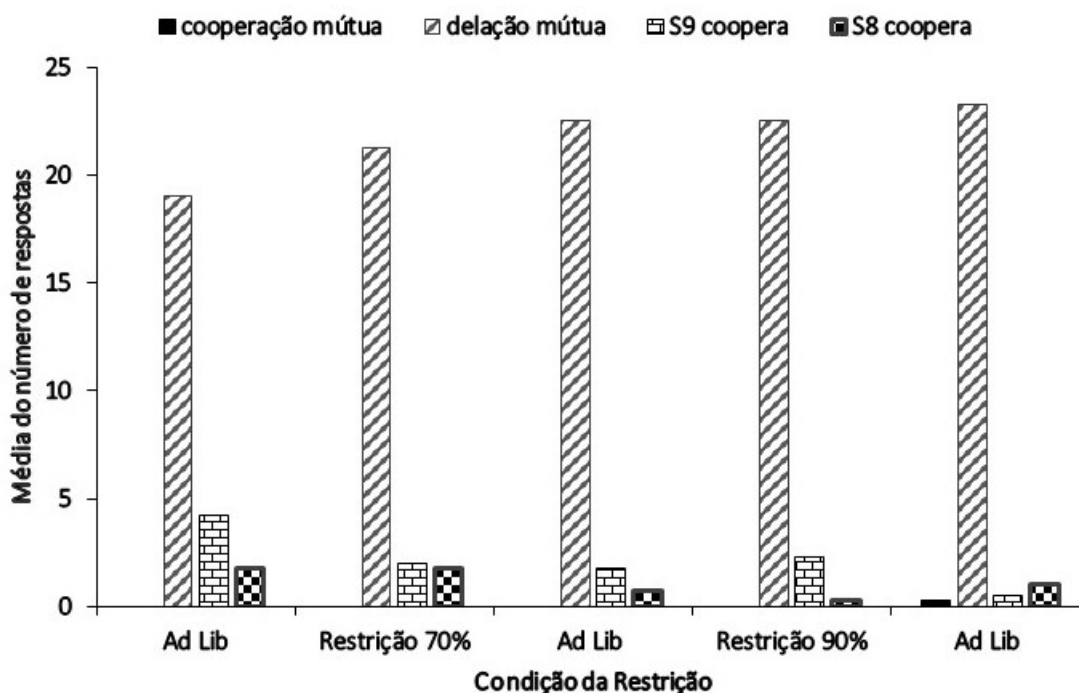


Figura 7. Média do número de tentativas classificadas como delação mútua, cooperação mútua e cooperação unilateral, das quatro últimas sessões das fases de testes do Dilema do Prisioneiro dos sujeitos da dupla S8 e S9 em que vigoraram diferentes níveis de restrição alimentar.

Na primeira fase *ad libitum*, houve predominância de respostas de delação mútua (média de 19 respostas), ausência de respostas de cooperação mútua e cooperação unilateral de ambos os sujeitos, sendo maior a de S9 (média de 4,25 respostas) em relação ao S8 (média de 1,75 respostas).

Na fase de restrição de 70% do peso *ad libitum*, a dupla S8 e S9 aumentou a média de tentativas de delação mútua (21,25) e cooperação unilateral em um número próximo de vezes (S8 com média de duas respostas, e S9 com média de 1,75). Não houve

tentativas com cooperação mútua. Entretanto, diferentemente das duas duplas anteriores, nas quatro últimas sessões analisadas da fase de 70% de restrição, não houve aumento das respostas de cooperação unilateral em relação à fase *ad libitum*.

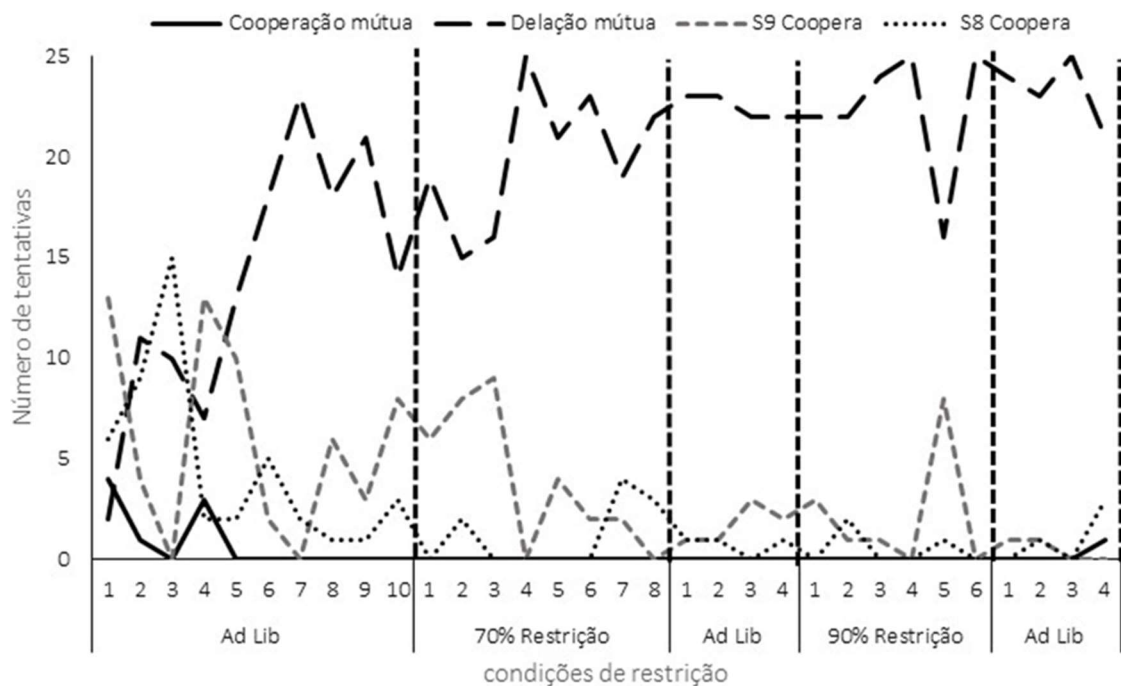
Na segunda fase *ad libitum*, a dupla S8 e S9 também diminuiu a média de respostas de cooperação unilateral, como nas duplas anteriores, sendo que S9 (1,75) respondeu pouco mais que S8 (0,75), que praticamente não respondeu, resultando num aumento de tentativas de delação mútua (média de 22,5).

Na fase de restrição de 90% de peso *ad libitum*, a dupla S8 e S9 teve pequeno aumento de respostas de cooperação unilateral (2,25) em comparação à fase anterior, e S8 (0,25) continuou respondendo muito pouco.

Na última fase *ad libitum*, S8 aumentou suas respostas de cooperação unilateral (uma resposta) comparada à fase anterior e S9 (0,5), ficou próxima a zero.

A Figura 8 permite identificar como foi o número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S8 e S9 ao longo das sessões.





*Figura 8.* Número de tentativas em que houve cooperação mútua ou unilateral e delação mútua pela dupla S8 e S9 nas sessões das fases de teste do Dilema do Prisioneiro, nas diferentes condições de restrição alimentar.

Na análise geral de todas as sessões da dupla S8 e S9, como aconteceu com as duplas anteriores, a cooperação mútua ocorreu em algumas sessões do início do experimento e logo desapareceu. Com relação às respostas de cooperação unilateral de S9, há uma variação sensível à restrição (antes de o peso estabilizar) entre seis e nove tentativas; porém, essa variação nem sempre se deu dentro do período das últimas quatro sessões em que os dados são analisados. Na Figura 8, que mostra todas as sessões da dupla S8 e S9, pode-se notar o aumento das respostas de S9 nas fases de 70% e em uma sessão da fase de 90%.

Comparativamente, as três duplas apresentaram reações semelhantes nas fases de restrição alimentar, e a emissão de respostas da dupla S8 e S9, na condição restrição 70%, foi bem menor após o peso se ter estabilizado. Houve alguma alternância de respostas entre os sujeitos das duplas. Entretanto, houve aumento maior de resposta de cooperação

unilateral, em geral, de apenas um dos sujeitos da dupla, e maior número de tentativas com resposta quanto maior a restrição.

Analisando o experimento em sua totalidade, o comportamento de cooperação mútua que ocorre nas primeiras sessões do teste do Dilema do Prisioneiro logo deixa de ocorrer (no máximo, até a quinta sessão) para todas as três duplas.

Na Figura 9, estão representadas todas as sessões de treino discriminativo dos sujeitos do experimento e as etapas a que foram submetidos.

O sujeito S2 foi o primeiro a atingir o critério, tendo de esperar S7 ter o mesmo resultado em três sessões consecutivas (mantendo-se em treino por mais três sessões); S5 foi o terceiro a atingir o critério que teve que esperar S10 conseguir 70% de discriminação por três sessões consecutivas; S8 foi o quinto a atingir critério e esperou S9 atingir o mesmo resultado.

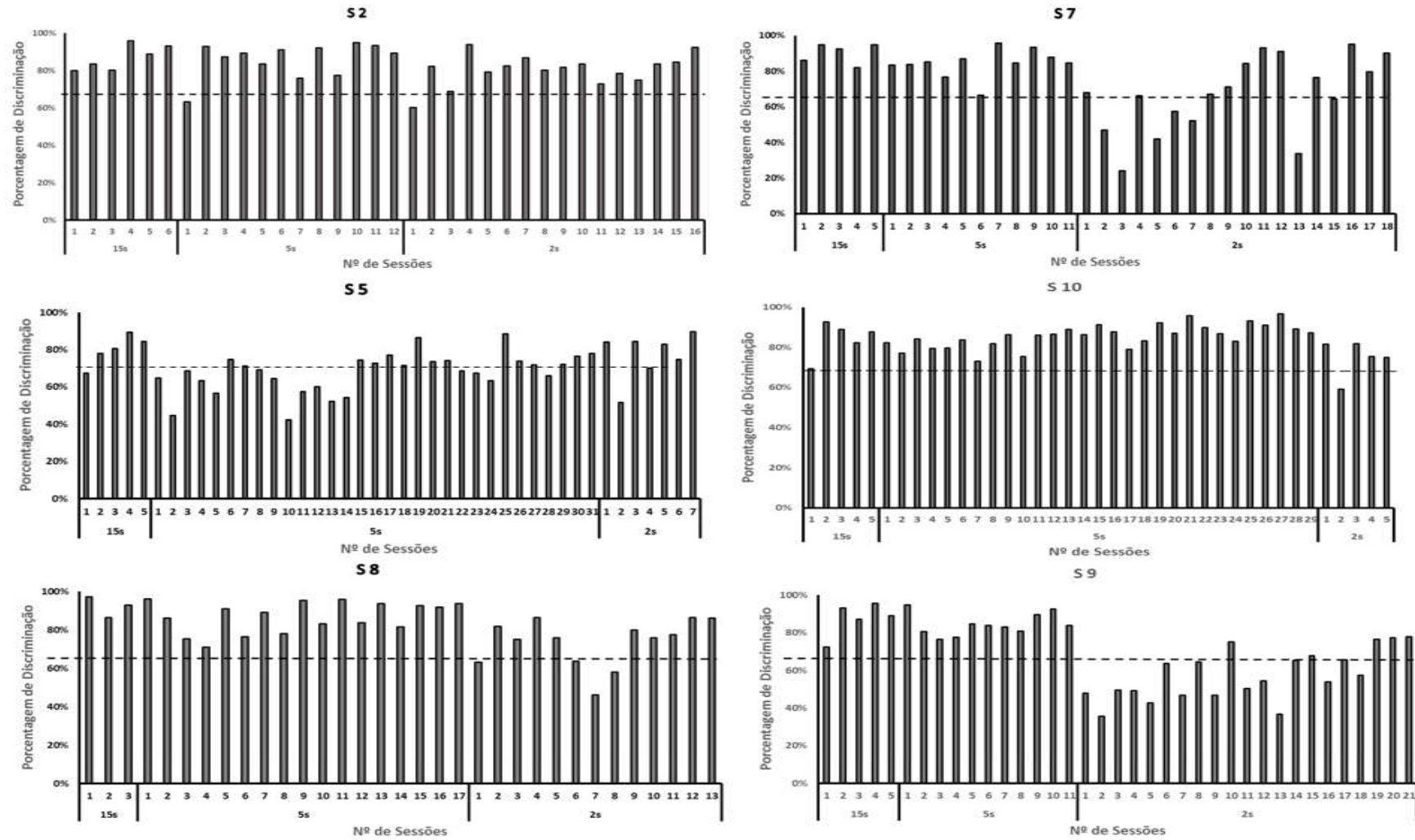


Figura 9. Porcentagem de respostas no escuro durante as sessões de treino discriminativo dos sujeitos de cada dupla, nas diferentes etapas do treino.

Ao analisar a Figura 9 e relacioná-la com os resultados do teste com o Dilema do Prisioneiro, o primeiro sujeito da dupla S2 e S7 que atingiu critério no treino discriminativo teve maior número de sessões em 2 s que S7, pois o parceiro, assim que o atingiu, já foi submetido ao teste do Dilema na última etapa, enquanto S2 treinava, esperando a evolução do parceiro. No cômputo geral, S2 foi o primeiro a atingir o critério do treino discriminativo e foi, visivelmente, o sujeito mais cooperativo da dupla.

Essa relação se inverte na dupla S5 e S10, pois S5 teve um treino mais longo em 2s, mas, no teste com o Dilema do Prisioneiro, foi *menos* cooperativo que S10, excetuando-se a condição de 90% de peso *ad libitum*, onde cooperou mais que S10. A dupla S8 e S9 teve resultado semelhante à dupla S5 e S10. S9 foi o mais cooperativo da dupla; porém, foi o segundo a atingir o critério. Portanto, não se identificou uma relação entre maior quantidade de treino em 2 s e respostas de cooperação dos sujeitos.

Uma análise das respostas durante as últimas quatro sessões de cada fase de teste do Dilema do Prisioneiro, emitidas ao longo das tentativas durante o escuro (em que elas poderiam produzir reforçamento) e durante a luz de teto (condição em que não havia reforçamento) é apresentada nas Figuras 10, 11 e 12, para as duplas S2 e S7; S5 e S10; e S8 e S9, respectivamente.

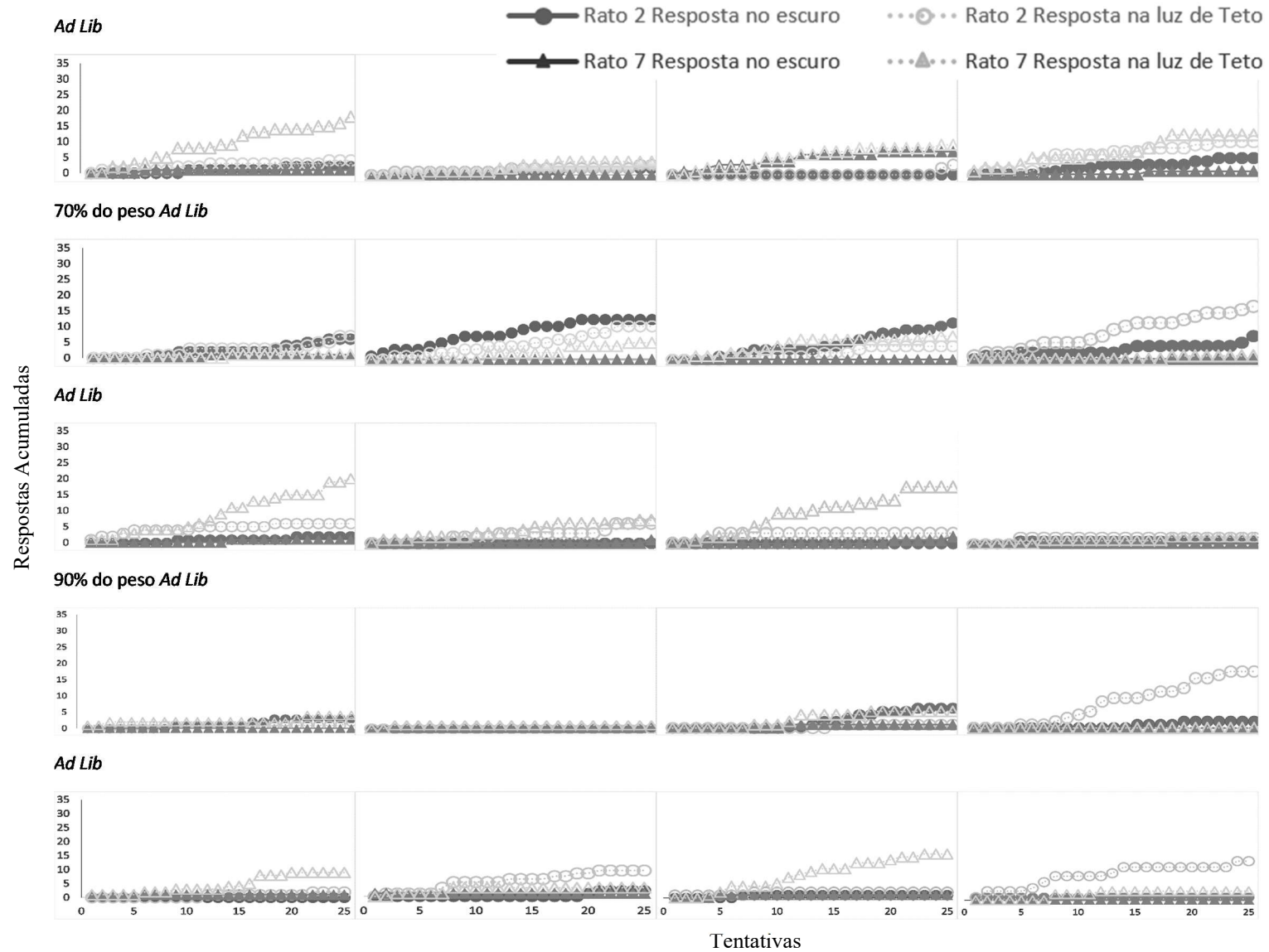


Figura 10. Frequência acumulada de respostas por tentativa das últimas quatro sessões de cada condição alimentar da dupla S2 e S7.

Ao analisar os resultados por tentativa das últimas quatro sessões da primeira fase *ad libitum* da dupla S2 e S7, observa-se que, na primeira sessão analisada, o S7 respondia mais frequentemente na luz de teto (18) que na luz apagada (1), assim como S2 em menor intensidade com quatro respostas na luz de teto contra duas na luz apagada ao longo das tentativas. Isso volta a ocorrer na última sessão dessa fase, com S7 respondendo 12 vezes na luz de teto contra uma vez na luz apagada, e S2 respondendo 10 vezes na luz de teto contra cinco vezes na luz apagada, como mostra a Figura 10.

O dado sugere que, na ausência de privação, o controle estabelecido durante o treino discriminativo (em privação) não ocorreu. Outra característica é que as respostas emitidas não foram consequências direta da resposta do companheiro, não obedecendo às regras de reciprocidade direta, que são: caso um sujeito coopere (responda na barra dando reforço ao seu parceiro), na tentativa a seguir, o parceiro retribui acionando a barra e reforçando reciprocamente o sujeito. Não havendo reciprocidade direta aqui, estratégias como TFT ou Pavlov não existem. Ademais, não se observam respostas alternadas consequenciadas dos sujeitos.

Na fase de 70% de peso *ad libitum*, S2 respondeu muito mais (entre seis a 12 respostas) que S7 (entre zero e uma resposta) em luz apagada, mas respostas em S-delta (luz de teto) também aconteciam (S2, entre sete e 16 respostas). Esse pequeno aumento do número de resposta já se tinha iniciado na última da fase *ad libitum*, especialmente nas tentativas finais da sessão. Isso se repete nas três primeiras sessões da fase 70% de restrição, diminuindo um pouco na última sessão dessa fase. Outra característica dessa fase é a quase ausência de resposta de S7 em luz apagada, que só as emitiu em S-delta (luz de teto, entre uma e sete respostas).

Na segunda fase *ad libitum*, houve várias respostas na luz de teto, quando o tempo é maior, principalmente de S7 (respostas em S-delta: 20 e 17 respostas) na primeira e

terceira sessão. Nessas sessões, a barra estava do mesmo lado, lembrando que os sujeitos eram trocados de lado da caixa experimental a cada sessão. No geral, nessa fase, respostas em SD não aconteceram para a dupla, e a quarta sessão quase não registrou respostas de nenhum dos sujeitos, fosse em SD ou S-delta.

Na fase de 90% de peso *ad libitum*, S2 apresenta mais respostas em SD que S7. No entanto, o total de respostas não excedeu o número de 10 para ambos os sujeitos, exceto na quarta sessão, quando S2 aumentou a resposta em S-delta (luz de teto), com 17 respostas nessa sessão.

A última fase, terceira *ad libitum*, há mais respostas em S-delta (luz de teto) de ambos os sujeitos (S2, duas a 13 respostas; e S7, de três a 15 respostas), e, para ambos os sujeitos, elas ocorreram no mesmo lado da caixa. Respostas em SD (luz apagada) não excederam duas respostas por sessão. Em resumo, S2 respondeu mais, fosse na luz de teto, fosse no escuro, em restrição alimentar; e S7 respondeu mais em *ad libitum*.

Na Figura 11, estão representadas as frequências acumuladas das sessões de teste do Dilema do Prisioneiro por tentativas da dupla S5 e S10.

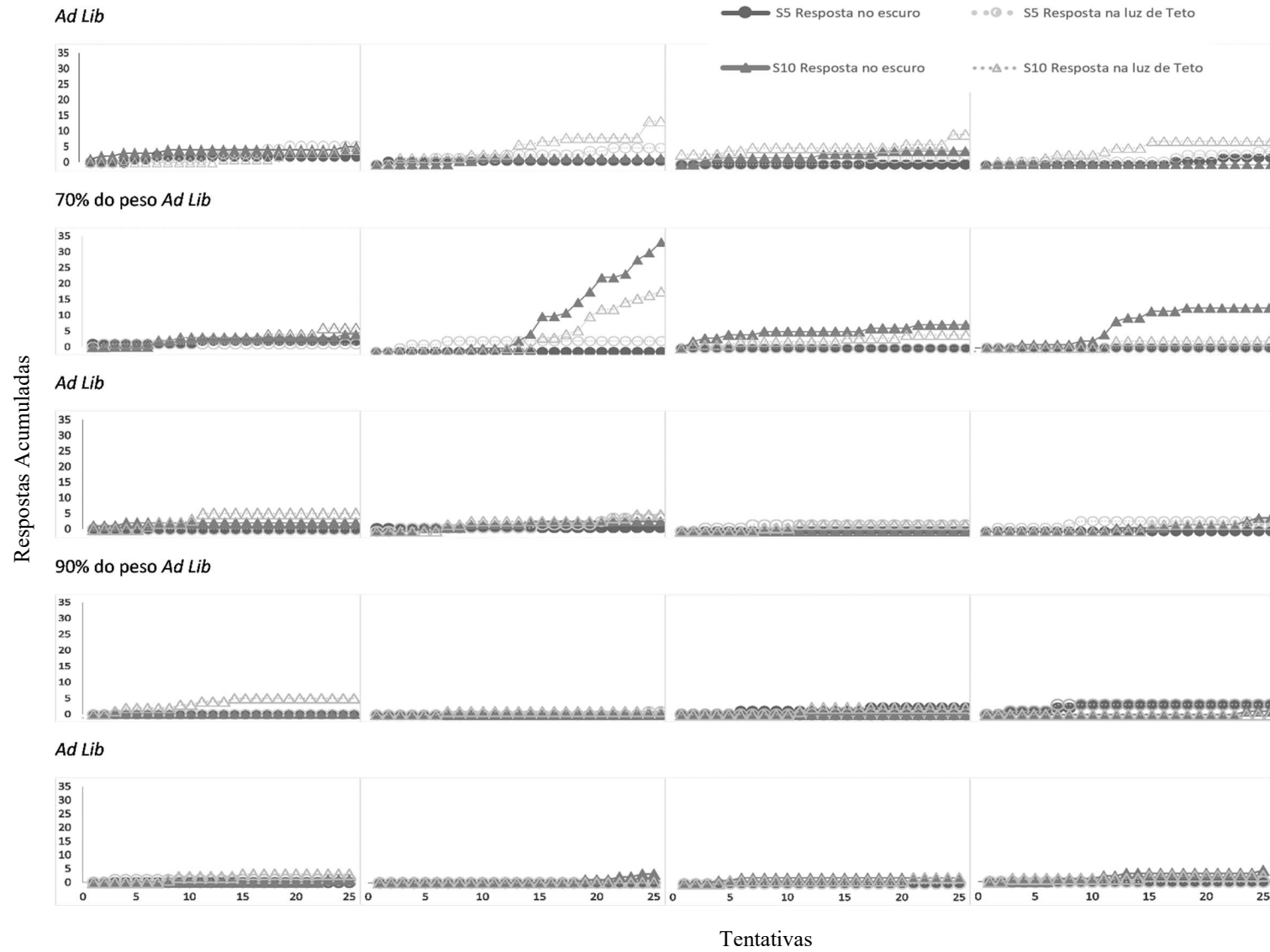


Figura 11. Frequência acumulada de respostas por tentativa das últimas quatro sessões de cada condição alimentar da dupla S5 e S10.



Na primeira fase *ad libitum*, S10 respondeu em luz apagada em todas as sessões, menos na quarta, e S5 deixou de responder apenas na terceira sessão; porém, ambos os sujeitos responderam entre zero e cinco por sessão, um resultado bem baixo comparado às outras duplas. S10 respondeu em S-delta em todas as sessões (respostas entre quatro e 13 vezes), e S5 também respondeu em todas as sessões em S-delta, mas não chegou a mais de cinco respostas por sessão.

Na fase de 70% de peso *ad libitum*, S10 teve aumento de todas as respostas, principalmente no escuro na segunda sessão, chegando a emitir 31 respostas. Respostas na luz de teto começaram com muita frequência e diminuíram nas sessões seguintes para S10, chegando a 17 respostas e diminuindo para duas na última sessão. Nas Sessões 2 e 4, observa-se um aumento no número de tentativas com respostas a partir da metade das sessões. Em nenhuma tentativa, houve respostas no escuro de S5 (que produziria reforços para S2). S5 quase não respondeu nessa fase, fosse em SD, fosse em S-delta, com respostas entre 3 a zero nas sessões.

A segunda fase *ad libitum* teve poucas respostas dos sujeitos, fossem no escuro, fossem durante a luz de teto, com apenas uma resposta na segunda sessão em SD e respostas entre zero e quatro em S-delta, para S5; e respostas em SD entre zero e quatro e S-delta, entre duas e cinco, para S10. Manteve-se, porém, a predominância de respostas em S-delta de ambos os sujeitos.

A fase de 90% *ad libitum* teve poucas respostas, fossem em SD, fossem em S-delta, não excedendo cinco respostas para ambos os estímulos. Na primeira sessão, S10 respondeu na luz de teto cinco vezes e apresentou uma resposta em SD na última sessão. S5 teve algumas respostas no escuro na terceira (2) e quarta sessões (3).

Na terceira fase *ad libitum*, S5 praticamente não respondeu nessa condição, com apenas uma resposta em S-delta. S10 teve algumas respostas no escuro (respostas entre

uma e quatro por sessão) e uma frequência similar na luz de teto (respostas entre zero e três).

Em resumo, S10 teve um aumento de respostas na condição de 70%, aproveitada por S5, que respondeu pouco, independentemente da condição.

Na Figura 12, está representada a frequência acumulada das sessões de teste do Dilema do Prisioneiro por tentativas da dupla S8 e S9.

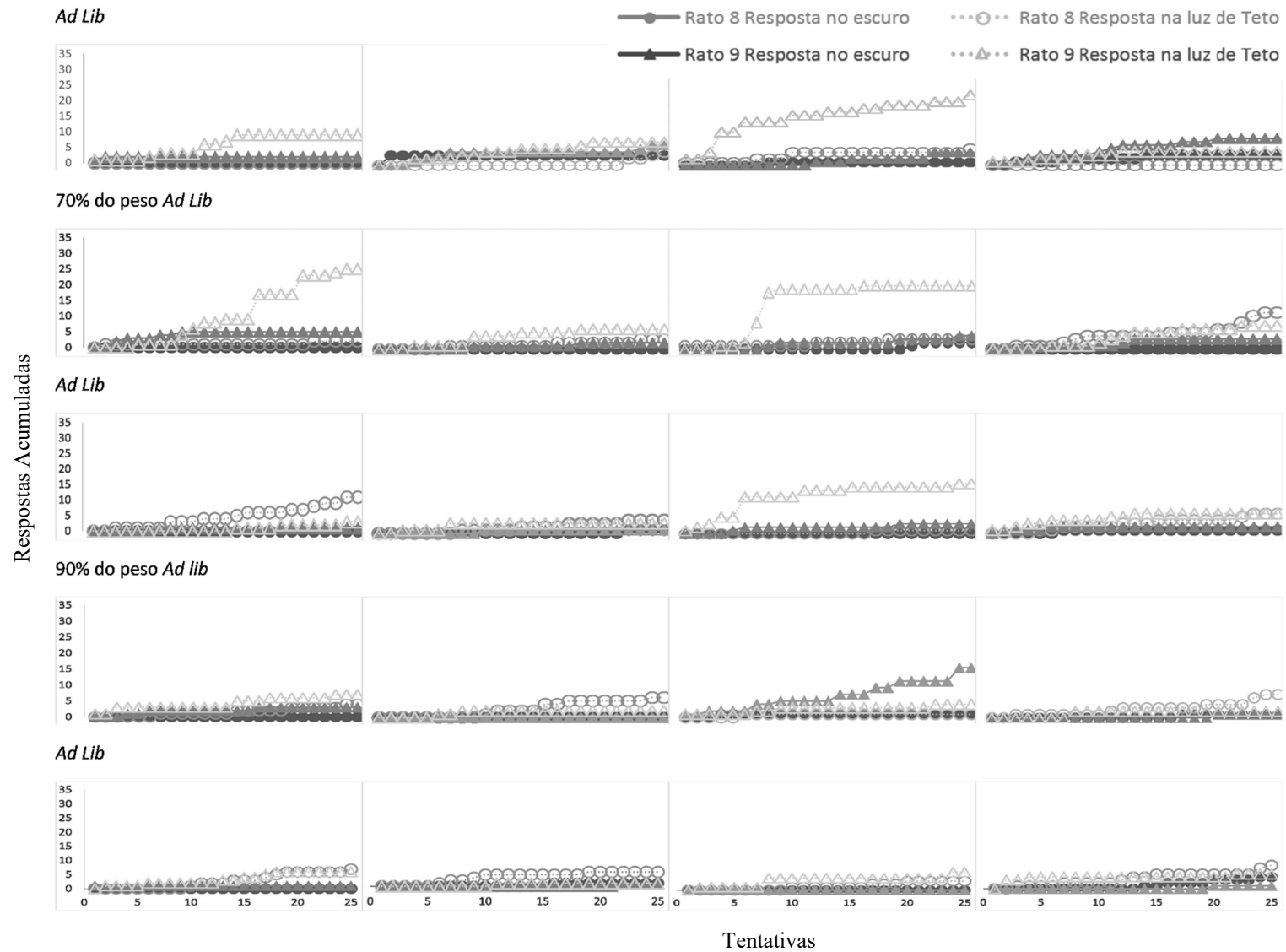


Figura 12. Frequência acumulada de respostas por tentativa das últimas quatro sessões de cada condição alimentar da dupla S8 e S9.

Na primeira sessão *ad libitum* da dupla S8 e S9, S9 responde com muita frequência na luz de teto (21 vezes na terceira sessão). Na segunda e terceira sessões, há um aumento de respostas no escuro de ambos os sujeitos (S8 entre zero e três respostas, e S9 entre duas e oito respostas); porém, não há aplicação de estratégias para explicar o fenômeno: como se vê na Figura 12, não há uma sequência de respostas alternadas dos sujeitos indicando reciprocidade.

Na fase de 70% do peso *ad libitum*, respostas no escuro ocorreram por parte de S9 (respostas entre duas a cinco por sessão), que, na primeira sessão dessa condição, responde em cinco tentativas, número semelhante ao da quarta sessão *ad libitum* (4) da fase anterior. Nas sessões seguintes, esse número diminui. S9 respondeu durante a luz de teto, principalmente na primeira (25) e terceira sessão (19), repetindo o padrão da condição *ad libitum*. S8 respondeu pouco no escuro, apenas duas vezes na terceira sessão. Na luz de teto, S8 emitiu poucas respostas, exceto na quarta sessão, quando registrou 11 respostas.

Novamente, a dupla responde pouco em luz apagada na segunda fase *ad libitum* (entre zero a três para ambos os sujeitos), e S8 apresenta respostas durante a luz de teto em todas as sessões, principalmente na primeira (11) e quarta sessões (6), não evidenciando controle pelo lado da caixa operante e talvez seguindo a tendência que apresentou na quarta sessão da condição antecedente. S9 também apresentou respostas na luz de teto, principalmente na terceira (15) e quarta sessões (6).

Em 90% do peso *ad libitum*, S9 teve um aumento considerável de respostas em SD na terceira sessão, chegando a 15 respostas. S8 teve poucas respostas em SD, apenas uma na terceira e uma na quarta sessão; e respondeu em luz de teto em todas as sessões (entre uma a sete respostas). S9, aliás, também respondeu em todas as sessões em S-delta (entre duas a sete respostas).

Na terceira fase *ad libitum*, S8 e S9 responderam durante a luz de teto em todas as sessões, mas com pouca frequência (S8 entre três e oito respostas, e S9 entre uma e seis respostas). Poucas respostas no escuro ocorreram, não excedendo quatro respostas por sessão.

### **Análise de Pelotas Recebidas nas Quatro Sessões Analisadas em Cada Condição de Restrição**

O Dilema do Prisioneiro como modelo de pesquisa de variáveis de autocontrole necessita de dois fatores: a estratégia de reciprocidade do oponente e a repetição do jogo. Assim, a cooperação pode ser reforçada, e a delação, punida nas tentativas subsequentes. O resultado final seria um ganho maior atrasado pela cooperação mútua do que ganhos esporádicos imediatos na delação unilateral ou ganhos globais menores com delação mútua. Neste estudo, o comportamento de cooperação mútua não se manteve, não havendo um ganho acumulado que possa ser calculado.

De modo geral, os sujeitos que mais responderam durante o escuro, ou seja, mais cooperativos, receberam menos pelotas – e, quanto mais pelotas receberam, menos o parceiro respondeu durante o escuro, em correspondência. A Figura 13 mostra as pelotas recebidas nas quatro últimas sessões de cada condição de restrição da fase do Dilema do Prisioneiro.

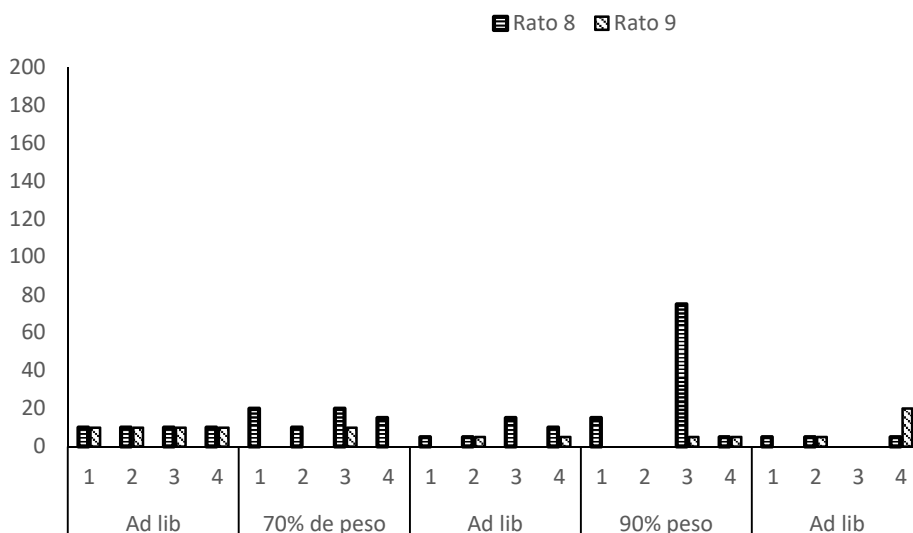
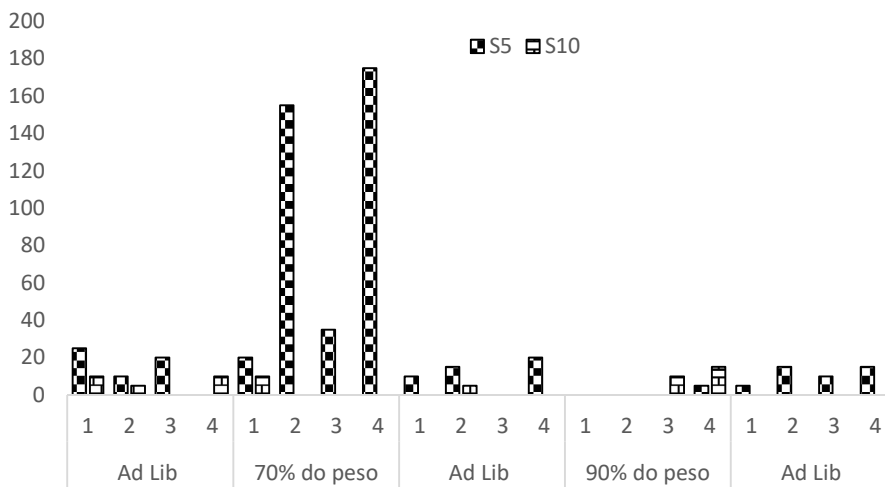
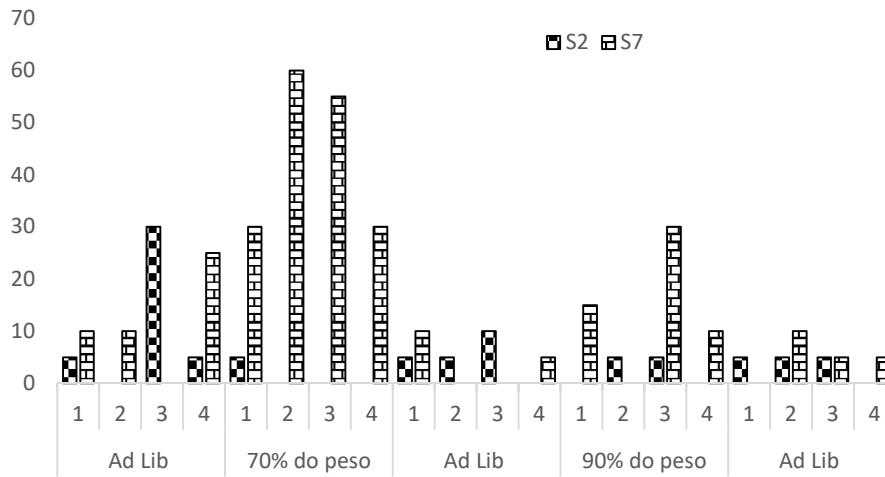


Figura 13. Número de pelotas recebidas por dupla S2 e S7, S5 e S10 e S8 e S9 nas últimas quatro sessões de cada condição de restrição da fase do Dilema do Prisioneiro.

Esse estudo replica o de Wood et al. (2016) com um diferencial: não foi utilizada barra retrátil que era introduzida na caixa experimental apenas quando a luz verde acendia por 2 s para, em seguida, ser retraída novamente e acender a luz de teto. No presente estudo, a barra permanecia na caixa durante todo o experimento, sendo que as respostas só foram reforçadas na luz apagada. Nas outras condições, luz verde e luz de teto, respostas de pressionar a barra não produziram consequências, utilizando-se, assim, um treino discriminativo.

Esse treino pode ter influenciado o número de respostas durante o teste com o Dilema do Prisioneiro, pois, na condição em que não havia possibilidade de respostas de pressão à barra no estudo de Wood et al. (2016), neste estudo, havia a possibilidade de os sujeitos apresentarem respostas em S-delta, como pode ser observado nas Figuras 10, 11 e 12, de respostas acumuladas.

No estudo de Wood et al. (2016), houve estabelecimento do comportamento de cooperação mútua que diminuiu com a restrição alimentar. Sujeitos aumentaram suas respostas de pressão à barra sob restrição alimentar; porém, na análise geral de recebimento de pelotas, houve diminuição das pelotas adquiridas. Em outras palavras, houve aumento de resultados de cooperação/delação unilateral na restrição alimentar, que, na contabilidade final, somam uma quantidade menor de pelotas para a dupla do que se cooperassem mutuamente. Recordar-se que os autores diferenciaram dois grupos com relação à resposta de cooperação: um grupo deveria acionar a barra como resposta de cooperação, e o outro grupo não responder na barra para cooperar.

No presente estudo, também houve aumento de pressão à barra durante a restrição alimentar, especialmente na maior restrição (70%) e para um dos sujeitos da dupla. No presente estudo, a resposta de cooperação não sofreu diferença com relação ao treino, sendo sempre de resposta na barra.

Os experimentos de Daniel (1942, 1943) indicaram que, sem o estímulo aversivo (choque), não houve cooperação e que, na cooperação sem estímulo aversivo, esse condicionamento de ordem superior em que o estímulo neutro para uma determinada resposta é emparelhado com outro estímulo já condicionado é difícil de acontecer em animais não humanos.

Stephens e Clements (1998, pp. 239-260) concluem que matrizes de pagamento heterogêneas de cooperação mútua e delação unilateral reforçadas positivamente e de delação mútua e cooperação unilateral punidas levam a índices maiores de cooperação mútua quando comparadas à matrizes de pagamento homogêneas, quando há apenas reforçamento positivo. No presente estudo, não houve estímulos aversivos. Wood et al. (2016), no entanto, citam que a ausência de pagamento na delação mútua e cooperação unilateral e a visão do parceiro recebendo comida e se alimentando é, em si, uma punição. Esta, porém, é uma informação difícil de confirmar. De qualquer forma, no presente estudo não houve, em geral, o acionamento simultâneo da barra dos sujeitos o que poderia levar à maior quantidade de pelotas para dupla ao longo do experimento.

Green et al. (1995) apresentaram resultados em experimentos com pombos, nos quais pombos só respondiam cooperando para o computador quando o resultado era uma recompensa maior e imediata. Baker e Rachlin (2002), trabalhando também com pombos e, corrigindo problemas encontrados pelos autores em Green et al. (1995), conseguiram comportamento de cooperação com um estímulo mais consistente do reforçador de maior magnitude e menor magnitude, além de diminuir o tempo entre tentativas.

O número de pelotas recebidos na cooperação mútua (3) e delação/cooperação unilateral (5) pode ter sido difícil para os sujeitos do presente estudo fazerem a distinção entre as magnitudes de reforço. Apesar de utilizado o mesmo método e matriz de pagamento de Wood et al (2016), a sinalização de reforços com relação à magnitude pode



ser mais consistente como foi em Baker e Rachlin (2002) utilizando pombos. Em vez de luzes, pode ser usado o *clicker*, para diferenciar a magnitude de reforço e uma discriminação maior entre cooperação mútua e delação unilateral.

O tempo de ITI de 30 s também pode ter sido excessivo. Baker e Rachlin (2002), em seu estudo com pombos, conclui que o ITI de 0 s teve um índice de cooperação significativamente maior do que o de 18 s. A média de ITI em Viana et al. (2010) foi de 4 a 5 minutos.

Stephens et al. (2002) demonstraram em seu experimento com pássaros sem restrição alimentar que, frente a um parceiro delator, não há cooperação; e, com um parceiro cooperativo, a cooperação mantém-se em reciprocidade. Com um parceiro jogando TFT (recíproco, portanto), a recompensa acumulada (maior magnitude do reforçador; porém, atrasado) mantém a cooperação indefinidamente, mas a recompensa não acumulada (menor magnitude imediata) iniciava-se com bons índices de cooperação e declinava para delação mútua. Esse estudo também aponta que a recompensa por cooperação mútua deve ser de magnitude considerável para que haja discriminação. No presente trabalho, houve um pequeno número de tentativas de cooperação mútua no começo do experimento, que logo desapareceu.

Em outro estudo que apresentou cooperação com ratos no Dilema do Prisioneiro, Viana et al. (2010) manipularam estratégias TFT de reciprocidade e pseudorandomizadas, que independem da resposta do sujeito (não considera a resposta dada na tentativa anterior), e utilizou estímulos aversivos (pinçamento da cauda) nos pagamentos de cooperação unilateral (*sucker*) e delação mútua (*punishment*). Resultados em TFT foram altos índices de cooperação mútua e cooperação/delação unilateral e um menor índice de delação mútua, resultado esse explicado pela maior probabilidade, na tentativa seguinte, de movimento de cooperação após resultado de delação mútua e, no

resultado de cooperação, pela probabilidade de ser a mesma que a resposta em cooperação ou delação, mesmo com uma matriz de pagamento para cooperação mútua e delação unilateral de quatro e seis pelotas, respectivamente. Já na estratégia pseudorandomizada o resultado de delação unilateral e mútua foi predominante. Todos os resultados anteriores foram em condição *ad libitum*, para verificar se a motivação (restrição alimentar) afeta o comportamento de cooperação, e ratos que foram colocados em restrição de 85% de peso *ad libitum* diminuíram as respostas de cooperação mútua jogando o TFT (Viana et al., 2010).

Vianna et al. (2010) utilizaram ratos *stooge* para o experimento, ou seja, manipulavam a resposta do parceiro do sujeito experimental de acordo com as estratégias que queriam implantar. No presente estudo, os animais foram estudados em duplas, com ambos os sujeitos da dupla respondendo livremente – e, apesar de apresentarem algumas respostas de cooperação mútua, não houve continuidade desse comportamento.

Com relação à privação, no presente estudo, o teste do Dilema do Prisioneiro começa com condição *ad libitum*, mas a cooperação mútua decresceu até desaparecer ainda nessa condição. As respostas de cooperação não encontraram uma reciprocidade na tentativa seguinte e não foi calculada a probabilidade dessas respostas. A visualização dos gráficos de frequência acumulada, porém, não apontam para uma tendência nesse sentido.

No presente estudo, o comportamento de cooperação mútua diminuiu rapidamente nas primeiras sessões de teste do Dilema do Prisioneiro, mas se manteve o de cooperação unilateral de, pelo menos, um dos sujeitos da dupla, durante a condição que vigorou maior restrição (70%). Resultado semelhante obtiveram Gardner et al. (1984), cujos dados se assemelham bastante com os do presente experimento: curvas de delação mútua altas e respostas de cooperação mútuas e unilaterais com baixa incidência.

Gardner et al. (1984) concluem que os ratos não são animais cooperativos e comportamento de cooperação decresce ainda mais sem comunicação. Alguns comportamentos foram descritos no experimento de Gardner et al. (1984), que os denominou anedóticos, como uma dança de um dos sujeitos tentando influenciar a resposta do parceiro. No presente estudo, o único comportamento notado é de excitação ao ver o parceiro se alimentando e/ou recebendo as pelotas.

Com relação à maior incidência de resposta, proporcional à restrição alimentar sem consequência direta no recebimento de suas próprias pelotas, mas produzindo pelotas para seu parceiro, o barulho do distribuidor de pelota pode ter adquirido a função de reforçador condicionado. Mowrer (1939) também identificou que um sujeito trabalha para satisfazer seu parceiro até que ele mesmo possa se alimentar, numa relação que denominou de trabalhador-parasita.

Nemeth (1970) explica em seu artigo em psicologia social que a falta de cooperação em jogos envolvendo o Dilema do Prisioneiro é a interdependência dos dois jogadores: uma vez que os sujeitos da pesquisa não conseguem prever os resultados de suas ações, para si ou para seu parceiro de jogo, há uma menor probabilidade de surgir um comportamento altruísta. Essa explicação encontra sustentação nas contingências comportamentais entrelaçadas, uma vez que os sujeitos não estabeleceram relação da resposta dos parceiros e recebimento de pelotas (resposta do sujeito sendo estímulo discriminativo para o parceiro), com a resposta simultânea do parceiro sendo consequência reforçadora a longo prazo. Não houve comportamento modelado pelas contingências e, conseqüentemente, não houve cooperação.

## Conclusão

O presente estudo não conseguiu replicar os resultados de Wood et al. (2016). Os resultados de delação mútua foram predominantes, com alguma variação nas respostas de cooperação/delação unilateral, principalmente nas fases de restrição alimentar. Essa resposta unilateral dos sujeitos do experimento não parece encontrar relação nas diferenças já identificadas do treino discriminativo.

Pode haver uma discriminação distinta, não identificada, entre os sujeitos em que um deles trabalha mais recebendo, eventualmente, um ou outro reforço de seu parceiro mais numa relação de contiguidade (comportamento supersticioso) do que numa relação social.

Numa relação translacional para o comportamento social humano, um sujeito pode trabalhar mais devido ao reforçamento intermitente de seu parceiro e aumentar ainda mais esse trabalho quando privado de algum reforçador, sendo a proporção privação/trabalho intimamente relacionada. Um ponto a se considerar é que as sessões em que o sujeito mais trabalhou, o parceiro, em geral, respondeu menos.

Estudos futuros devem investigar diferentes magnitudes do reforço nesse modelo em ratos e verificar se o tempo de ITI poderia produzir alteração nas respostas de cooperação para a espécie. A barra retrátil e o treino discriminativo foram diferenças cruciais que podem ter influenciado os resultados de ambos os experimentos. Alterar o critério, aumentando seu índice, e treinar cada animal da dupla em apenas um lado da caixa operante pode aproximar os resultados com o estudo replicado e avaliar a restrição alimentar com resultados de cooperação mútua.

## Referências

- Alencar, A. I., & Yamamoto, M. E. (2008). A teoria dos jogos como metodologia de investigação científica para a cooperação na perspectiva da psicologia evolucionista. *Psico*, 39(4), 522–529.
- Almeida, H. F., & Chacha, L. A. (2014). Dilema do Prisioneiro Iterado e estratégia evolucionariamente estável: Uma Abordagem Econômica. *Revista de Economia*, 40(1), 35–53. doi:10.5380/re.v40i1.32404
- Axelrod, R. M. (1984). *The evolution of cooperation*. New York: Basic Books Inc.
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489), 1390–1396. doi:10.1126/science.7466396
- Baker, F., & Rachlin, H. (2002). Self-control by pigeons in the Prisoner's Dilemma. *Psychonomic Bulletin & Review*, 9(3), 482–488.  
doi:doi.org/10.3758/BF03196303
- Carlton, P. L. (1962). Effects on deprivation and reinforcement magnitude of response variability. *Journal of the Experimental Analysis of Behavior*, 5, 481–486.  
doi:10.1901/jeab.1962.5-481
- Daniel, W. J. (1942). Cooperative problem solving in rats. *Journal of Comparative Psychology*, 34(3), 361–368. doi:10.1037/h0062264
- Daniel, W. J. (1943). Higher order cooperative problem solving in rats. *Journal of Comparative Psychology*, 35(3), 297–305.  
doi:http://dx.doi.org/10.1037/h0056629
- Dawkins, R. (2007). *O gene egoísta* (R. Rubino, Trad.) São Paulo: Companhia das Letras.

- Echagüe, V. L. (2006). *Explorando algumas relações entre custo de resposta, magnitude de reforço e comportamento cooperativo* (Dissertação de mestrado). Pontifícia Universidade Católica de São Paulo, São Paulo, Brasil.
- Faleiros, P. B. (2009). *Efeitos do tipo de acesso à soma da pontuação do outro jogador na emissão de respostas "cooperativas" no jogo do Dilema do Prisioneiro Repetido* (Tese de doutorado). Universidade de São Paulo, São Paulo, Brasil.
- Fidelis, D. P., & Faleiros, P. B. (2017). Dilema do Prisioneiro na análise experimental do comportamento. *Revista Brasileira de Análise do Comportamento*, 13(1), 42–52.
- Gardner, R. M., Corbin, T. L., & Beltramo, J. S. (1984). The Prisoner's Dilemma game and cooperation in the rat. *Psychological Reports*, 55(3), 687–696.  
doi:10.2466/pr0.1984.55.3.687
- Green, L., Price, P. C., & Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. *Journal of the Experimental Analysis of Behavior*, 64(1), 1–17. doi:10.1901/jeab.1995.64-1
- Hake, D. F., & Vukelich, R. (1972). A classification and review of cooperation procedures. *Journal of the Experimental Analysis of Behavior*, 18(2), 333–343.  
doi:10.1901/jeab.1972.18-333
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859), 1243–1248. doi: 10.1126/science.162.3859.1243
- Kahneman, D., & Tversky, A. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.
- Keller, F. S., & Schoenfeld, W. N. (1950). *Princípios de psicologia* (5. ed.). New York: Editora Pedagógica e Universitária Ltda.

- Killingback, T., & Doebeli, M. (1999 de 09 de 01). 'Raise the stakes' evolves into a defector. *Nature*, *400*, 518. doi:10.1038/22913
- Komorita, S. S., Hilty, J. A., & Parks, C. D. (1991). Reciprocity and cooperation in social dilemmas. *Journal of Conflict Resolution*, *35*(3), 494–518. doi:10.1177/0022002791035003005
- Kraines, D., & Kraines, V. (1989). Pavlov and the Prisoner's Dilemma. *Theory and Decision*, *26*, 47–79. doi:10.1007/BF00134056
- Kropotkin, P. (2009). *Ajuda mútua: um fator de evolução* (W. A. Jr., Trad.). São Sebastião, SP: A Senhora Editora Ltda. (Trabalho original publicado em 1902)
- Kuhn, S. (2014). Prisoner's Dilemma. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Recuperado de <https://plato.stanford.edu/archives/spr2017/entries/prisoner-dilemma>
- Laraway, S., Snyckerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: some further refinements. *Journal of Applied Behavior Analysis*, *36*(3), 407–414. doi:10.1901/jaba.2003.36-407
- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, *37*(1), 149–155. doi:10.1901/jeab.1982.37-149
- Michael, J. (1993). Establishing operations. *The Behavior Analyst*, *16*(2), 191–206.
- Millenson, J. R. (1967). *Princípios de análise do comportamento*. New York: The Macmillan Company.
- Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. New Haven: Yale University Press.
- Mowrer, O. H. (1939). An experimentally produced "social problem" in rats. *Canadian Medical Association Journal*, *80*(8), 684.

- Mowrer, O. H. (1940). Animal studies in the genesis of Personality. *Transactions of the New York Academy of Sciences*, 3(1), 8–11. doi:10.1111/j.2164-0947.1940.tb00091.x
- Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, 54(2), 286–295.  
Recuperado de <http://www.cs.upc.edu/~ia/nash51.pdf>
- Nemeth, C. (1970). Bargaining and reciprocity. *Psychological Bulletin*, 74(5), 297–308.  
doi:10.1037/h0030071
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563. doi:10.1126/science.1133755
- Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, 359, 826–829. doi:10.1038/359826a0
- Pegden, C. D., Sadowski, R. P., & Shannon, R. E. (1995). *Introduction to simulation using SIMAN* (Vol. 2). New York, NY, USA: McGraw-Hill, Inc.
- Portugal, S. J., Hubel, T. Y., Fritz, J., Heese, S., Trobe, D., Voelkl, B., . . . Usherwood, J. (2014). Upwash exploitation and downwash avoidance by flap phasing in ibis formation flight. *Nature*, 505, 399–402. doi:10.1038/nature12939
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 17(1), 15–22.  
doi:<http://doi.org/10.1901/jeab.1972.17-15>
- Rachlin, H., Brown, J., & Baker, F. (2000). Reinforcement and punishment in the Prisoner's Dilemma Game. *The Psychology of Learning*, 40, 327–364.  
doi:10.1016/S0079-7421(00)80024-9
- Rapoport, A., & Chammah, A. M. (1970). *Prisoner's Dilemma: A study in conflict and cooperation* (2. ed.). The University of Michigan Press.



- Rocha, L. M., Meirelles, F., Sérgio, T. M., Micheletto, N., & Bornacina, R. (2010). A restrição de água e de alimento e alguns de seus múltiplos efeitos. *Psicologia: Teoria e Prática*, *12*(1), 2–15.
- Samuelson, L. (2002). Evolution and game theory. *Journal of Economic Perspectives*, *16*, 47–66.
- Sanabria, F., Baker, F., & Rachlin, H. (2003). Learning by pigeons playing against tit-for-tat in an operant prisoner's dilemma. *Learning & Behavior*, *31*(4), 318–331. doi:10.3758/BF03195994
- Senkowski, P. C., Vogel, V. A., & Pozulp, N. C. (1978). Differential reinforcement of lever-press durations: Effects of deprivation level and reward magnitude. *Learning and Motivation*, *9*(4), 446–465. doi:10.1016/0023-9690(78)90005-X
- Siegel, E., & Rachlin, H. (1995). Soft Commitment: Self-control achieved by response persistence. *Journal of the Experimental Analysis of Behavior*, *64*(2), 117–128. doi:10.1901/jeab.1995.64-117
- Skinner, B. F. (2003). *Ciência e Comportamento Humano* (11. ed.). São Paulo: Martins Fontes. (Trabalho original publicado em 1953)
- Smith, J. M. (1982). *Evolution and the theory of games*. Cambridge, United Kingdom: Cambridge University Press.
- Souza, M. R. (2007). *Os efeitos de diferentes níveis de privação sobre a variabilidade comportamental* (Dissertação de mestrado). Pontifícia Universidade Católica de São Paulo, São Paulo, Brasil.
- Stephens, D. W., & Clements, K. C. (1998). Game theory and learning. In L. A. Dugatkin, & H. K. Reeve, *Game theory and animal behavior* (pp. 239-260). New York: Oxford University Press Inc.

- Stephens, D. W., McLinn, C. M., & Stevens, J. R. (2002). Discounting and reciprocity in an Iterated Prisoner's Dilemma. *Science*, 298(5601), 2216–2218.  
doi:10.1126/science.1078498
- Tomanari, G. Y., Pine, A. d., & Silva, M. A. (2003). Ratos Wistar sob regimes rotineiros de restrição hídrica e alimentar. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 5(1), 57–71.
- Viana, D. S., Gordo, I., Sucena, É., & Moita, M. A. (2010). Cognitive and motivational requirements for the emergence of cooperation in a rat social game. *PLoS One*, 5(1), 1–9. doi:10.1016/j.anbehav.2016.01.010
- von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior* (60th Anniversary Commemorative Edition ed.). Princeton: Princeton University Press. (Trabalho original publicado em 1943)
- Wodka, M. B. (2012). *Conditions of cooperation between rats in the Prisoner's Dilemma Model, 154*. Recuperado de [https://digitalcommons.iwu.edu/psych\\_honproj/154](https://digitalcommons.iwu.edu/psych_honproj/154)
- Wood, R. I., Kim, J. Y., & Li, G. R. (2016). Cooperation in rats playing the Iterated Prisoner's Dilemma game. *Animal Behavior*, 114, 27–35.  
doi:10.1016/j.anbehav.2016.01.010