



Libraries and Learning Services

# University of Auckland Research Repository, ResearchSpace

## Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

## General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](#) and [Deposit Licence](#).

Negative discriminative stimulus as punishing consequence:  
A new approach to behavioural change

Vikki Joy Bland

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy  
(Psychology), University of Auckland, 2018.

## Abstract

The study and use of punishment in behavioural treatments has been restrained by ethical concerns. However, there remains a need to decrease harmful behaviour unable to be decreased by alternative methods. This thesis comprises four studies that collectively investigate a new approach to punishment. The first study investigated whether a negative discriminative stimulus will function as a punishing consequence. Pigeons were trained on a multiple schedule to discriminate between a positive discriminative stimulus (S+) signaling the presence of food in one component, and a negative discriminative stimulus (S-), signaling the absence of food in an alternate component. Once learned, every five responses on average to S+ produced S- for a duration of 1.5 s. In addition, responses to S+ produced food. There was a decrease in S+ response rate when responses produced S-, compared to when they did not. In addition, when responses to one of two concurrently available S+ alternatives produced S-, the proportion of choice to that alternative decreased compared with when responses to either alternative did not produce S-. Therefore, the contingent S- stimulus punished response rate and shifted choice. The second study investigated whether a contingent S- stimulus would punish responding for money in humans. There was no reliable effect of the contingent S- stimulus on response rate. The third and fourth studies investigated the ability of a negative discriminative stimulus to shift suboptimal choice processes. The third study replicated and extended a study by Vaughan (1981). Then, using Vaughan's procedure, the fourth study investigated whether a negative discriminative stimulus would shift suboptimal choice in pigeons. Interruption of the suboptimal choice process was observed in five of six pigeons, but was not consistent across repeated cycles.

Together these findings provide a foundation for continued investigation of the use of a negative discriminative stimulus as a punishing consequence. If a contingent S- stimulus is found to function reliably as a punisher, this approach has the potential to assist in the development of safe interventions to reduce or eliminate harmful behaviour in humans.

## Dedication

What I want to do I do not do, and what I hate; I do.

Romans 7:15.

This thesis is dedicated to God: Father, Son and Holy Spirit.

Who best understands what we do, and why we do it.

## Acknowledgements

I would like to thank my primary supervisor, Associate Professor Doug Elliffe for accepting me as his PhD student. I am more than grateful for the opportunity. Doug's vast knowledge of behavioural science – past, present and future – is gold. Doug, your insights are inspired, your knowledge is encyclopedic, and your retrieval time rivals Google.

Thank you to Dr Sarah Cowie, my secondary supervisor, for her positivity and encouragement, not only during my time as a PhD student, but also as an undergraduate student. Thanks for the confidence you instilled in me Sarah, and for your kindness and friendship over the past six years.

My thanks to Associate Professor Chris Podlesnik, of the Florida Institute of Technology, and formerly the University of Auckland. Chris taught me a great deal about experimental design and methodology, and about manuscript writing. He constantly gave of his time, both as my Honours supervisor in Auckland, and as a PhD advisor from Florida. Chris, thanks for your input over these past six years; it's been so much fun.

I would like to thank sponsors of the University of Auckland Doctoral Scholarship Fund. Completing a PhD is a fulltime job, and difficult to justify without an income. It is a privilege to have been a recipient of a University of Auckland Doctoral Scholarship.

I would like to acknowledge my family, and wider family and friends. I cannot name you all here, but you certainly know who you are. My love to our two sons, Matthew and Ethan, who bring me such joy – and joy is a great motivator. Special love and thanks to my mother, Pauline, who provided so much practical support across the PhD years; I couldn't have done this without you mum. Thanks too, to my dear friend Loran Reid, who is always as proud of me as I am of her.

Finally, my very special love and heartfelt thanks go to my husband Miles – my champion not only throughout these university years, but from the first day we met 25 years ago. This achievement is as much his as mine. Thank you LOML.

## TABLE OF CONTENTS

CHAPTER 1.	INTRODUCTION
CHAPTER 2.	PUNISHED BEHAVIOUR: A LITERATURE REVIEW
CHAPTER 3.	THESIS OVERVIEW
CHAPTER 4.	DOES A NEGATIVE DISCRIMINATIVE STIMULUS PUNISH RESPONDING FOR FOOD IN PIGEONS?
CHAPTER 5.	DOES A NEGATIVE DISCRIMINATIVE STIMULUS PUNISH RESPONDING FOR MONEY IN HUMANS?
CHAPTER 6.	SUBOPTIMAL CHOICE AND MELIORATION THEORY: AN OVERVIEW
CHAPTER 7.	REPLICATING AND EXTENDING THE FINDINGS OF VAUGHAN (1981)
CHAPTER 8.	DOES A NEGATIVE DISCRIMINATIVE STIMULUS PUNISH SUBOPTIMAL CHOICE IN PIGEONS?
CHAPTER 9.	GENERAL DISCUSSION
CHAPTER 10.	CONCLUSION
APPENDICES	
REFERENCES	

## LIST OF FIGURES

- Figure 2.1 Percentage of selected basic- (solid line) and applied- (dashed line) punishment studies (y-axis) published in 19-year time periods (x-axis) from 1950 to 2009, and since 2010.
- Figure 4.1. Left green key (G) and right red key (R) stimulus arrangements, and food delivery VI schedule value, for Components 1 and 2 in Baseline (top panel) and Punishment Test (bottom panel) for Conditions 1 to 5. The colour of the 1.5-s contingent S- stimulus (C.S.) presented during Punishment Tests was red for Conditions 1, 2, 3, and 5; and blue for Condition 4 (see Table 4.1).
- Figure 4.2 Left and right key stimulus arrangements for Components 1 and 2 in Baseline (top panel) and Punishment Test (bottom panel) sessions for Condition 5. The 1.5-s contingent S- stimulus (C.S.) presented during Punishment Tests was a red keylight. The C.S. was presented on the left key for the first block of two Punishment Tests, and on the right key for the second block of two punishment tests (see Table 4.1).
- Figure 4.3 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min measured as log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines), for individual pigeons, and group mean, in Condition 1.
- Figure 4.4 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min as a log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines) for individual pigeons, and group mean, in Condition 2.

- Figure 4.5 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min as a log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines), for individual pigeons, and group mean, in Condition 3.
- Figure 4.6 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min measured as log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines), for individual pigeons, and group mean, in Condition 4, the Control condition.
- Figure 4.7 Left-key (L1 and L2) and right-key (R1 and R2) S+ choice (solid black lines, back squares) and food delivery rates (dashed black lines, white squares), measured as a log proportion of mean baseline left- or right-key choice and food delivery rates, for two blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines) for individual pigeons, and group mean, in Condition 5.
- Figure 5.1 Diagram showing stimulus and scheduled money reinforcers arrangements for S+ components (Green screen) and S- components (Red screen) in Training (top panel) and Punishment Test (bottom panel) sessions for all conditions. A response-contingent red screen was superimposed for 1.5-s in S+ components during Punishment Tests according to a VR 5 schedule. Accrued money reinforcers (ACC) are displayed whenever the screen is green or red.
- Figure 5.2 Log proportion of S+ to S- response rate during the last two complete pairs of S+ and S- components in training for individual participants in Condition 1 (left panel), Condition 2 (middle panel), and Condition 3 (right panel). An asterisk indicates log proportion of S+ response rate of 0.20 or higher.

- Figure 5.3 S+ response rate (black line) and S- response rate (black dashed line) for each S+/S- component pair (y-axis) across training and the punishment test for individual learners in Condition 1 (left panel); 2 (middle panel), and 3 (right panel). The red vertical line indicates the end of the training phase.
- Figure 6.1 Illustration from Spetch et al. (1990), showing a concurrent-chains percentage-reinforcement procedure. FT x signifies that terminal links differed in duration by a fixed amount of time. FR signifies a fixed rate of one response required to initiate the terminal link. ‘Reliable’ signifies reinforcers were always available on the terminal link; ‘Unreliable’ signifies reinforcers were available on 50% of trials.
- Figure 7.1 Diagram of programmed hourly rates of reinforcement (y-axis) in Condition A (top panel) and Condition B (bottom panel) for left key (dashed black line, L) and right key (solid black line, R) responses as a function of proportion of time on the right key (x-axis) in a previous preference window, for Vaughan’s (1981) procedure (left) and the procedure for the current study (right). Choice is predicted to shift in a subsequent preference window when proportion of time allocated to the right key in a current window is between 0.26 to 0.74. Equal rates of reinforcement for left and right keys occur when proportion of time allocated to the right key is either 0.125 to 0.25 or 0.75 to 0.875.
- Figure 7.2 Overall obtained reinforcer-rate as the mean of the last five sessions of Condition A1, A2 and A3 (dark grey bars); and the mean of the last five sessions of Condition B1, B2 and B3 (light grey bars), for individual pigeons.
- Figure 7.3 Proportion of time on the right key (y-axis) for individual pigeons, averaged across the last five sessions of successive conditions A1, B1, A2, B2, A3 and B3 (x-axis). Horizontal rectangular boxes indicate optimal (high rate) and suboptimal (low rate) melioration zones.

- Figure 7.4 Proportion of time on the right key (y-axis) within successive preference windows 2-8 (x-axis) for individual pigeons averaged across the last five sessions in A- (left panel) and B conditions (right panel). Horizontal rectangular boxes indicate optimal and suboptimal zones.
- Figure 8.1 Diagram of programmed hourly rates of reinforcement (y-axis) in Condition A (top panel) and Condition B (bottom panel) for left key (dashed black line, L) and right key (solid black line, R) responses as a function of proportion of time on the right key (x-axis) in a previous preference window, for Vaughan's (1981) procedure (left) and the procedure for Cycles 1 and 2 of the current study (right). Choice is predicted to shift in a subsequent preference window when proportion of time allocated to the right key in a current window is between 0.26 to 0.74. Equal rates of reinforcement for left and right keys occur when proportion of time allocated to the right key is either 0.125 to 0.25 or 0.75 to 0.875.
- Figure 8.2 Overall obtained reinforcer-rate as the mean of the last five sessions of Condition A1, A2 and A3 (dark grey bars); the mean of the last five sessions of Condition B1, B2 and B3 (light grey bars), and the mean of the five sessions of Condition PT1, PT2, and PT3, for individual pigeons.
- Figure 8.3 Proportion of time on the right key (y-axis) for individual pigeons, averaged across the last five sessions of conditions A1, B1, PT1, A2, B2, PT2, and A3, B3, PT3 (x-axis). Horizontal rectangular boxes indicate optimal and suboptimal zones.
- Figure 8.4 Proportion of time on the right key (y-axis) within successive preference windows 1-7 (x-axis), for each of three cycles, for individual pigeons as the mean of the last five sessions in A conditions (grey line), B conditions (black line) and PT conditions (black dashed line). Horizontal rectangular boxes indicate optimal and suboptimal zones.

## LIST OF TABLES

- Table 2.1 Total applied punishment studies (N) by behaviour type and condition, with range and mean reduction, as a percentage, and mean post-punishment reduction period in months.
- Table 4.1 Discriminative stimuli for Components 1 and 2 and allocation of VI schedules of food deliveries on left and right keys across all conditions for baseline (top panel) and punishment test (bottom panel) sessions. Total number of sessions, and blocks of sessions, within each condition is shown in the far-right column.
- Table 4.2 Overview of group mean response rates and food delivery rates, in S+ components, with SEM; and group mean response rates in S- components, for Conditions 1 to 4. Condition 5 shows group mean response and food delivery rates for left and right keys in S+ components, with SEM, and response rates for left and right keys in S- components. The Discrimination Index shows group mean proportion of S+ to S- responses.
- Table 5.1 Total number of participants, phase duration, allocation of VI schedules of money reinforcers, or EXT, in each 60-s S+ and S- component, for training and punishment test phases in each condition.
- Table 5.2 Group mean responses and obtained money reinforcers per min in the S+ component, and responses per min in the S- component, with SEM, across the last two pairs of complete S+ and S- components during training, in each condition. The Discrimination Index shows the group mean proportion of S+ to S- response rates.
- Table 5.3 Group mean responses and obtained money reinforcers per min in the S+ component, and responses per min in the S- component, with SEM, across the last two pairs of complete S+ and S- components during training, in each condition, comprised of individual learners with a S+/S- Discrimination Index of 0.61 or higher.

Table 7.1 Allocation of VI schedules of reinforcement on left and right keys for Conditions A and B, as a function of proportion of time on the right key in the previous 4-min preference window.

Table 7.2 Mean overall response rates per minute across the last five sessions of individual A and B conditions, for individual pigeons. The far right-hand column presents the same rates averaged across all conditions.

Table 8.1. Mean overall response rates per minute across the last five sessions of all conditions, in each condition, in each cycle, for individual pigeons. The far right-hand column presents these rates averaged across all conditions.

## CHAPTER 1. INTRODUCTION

Harmful behaviour is an ongoing risk for 32,000 New Zealanders living with intellectual disability<sup>1</sup>, 65,000 with Autism Spectrum Disorder (ASD)<sup>2</sup> and 78,000 who, by 2026, will live with dementia.<sup>3</sup> In addition, between 5% and 10 %, or at least 46,000 typically-developing New Zealand children will emit physically and socially harmful behaviour between the ages of three and 17, causing harm to themselves or others. Left untreated, over time such behaviour results in a higher risk of conduct disorder, a behavioural disorder associated with educational failure and escalating antisocial behaviour in children and young adults (Frick, 2012).<sup>4</sup> Harmful behaviour is also a frequent side-effect of traumatic brain injury (TBI) arising from accidents and assaults. Recent research estimates there are 36,000 new TBI injuries per year in New Zealand alone (Feigin et al., 2013). Examples of harmful behaviour in all the human populations described above include self-injury, verbal abuse, aggression, pica, property destruction, impulsiveness, assault with weapons, stealing, spitting, biting, and screaming. In individuals with developmental disability, self-injury has been found to be disproportionately directed toward the head (Symons & Thompson, 1997), increasing the risk of sufferers sustaining a TBI over time.

In animals, harmful behaviour can include self-injury, such as tail-biting and excessive licking (Bécuwe-Bonnet, Bélanger, Frank, Parent, & Hélie, 2012), stereotypy due to boredom or stress (Mason & Rushen, 2008), predatory attacks on endangered species, and aggression towards humans (Landsberg, 1990) and other animals. From 2004 to 2014 more than 99,000 New Zealanders reported being bitten by a dog. Of these, 5842 required hospitalisation, and more than a third were children. In children younger than four years, 79% of bites were to the head or neck<sup>5</sup>.

---

<sup>1</sup> Ministry of Health, New Zealand (2013). Retrieved from <http://www.health.govt.nz/system/files/documents/publications/innovative-methods-providing-health-services-for-people-with-intellectual-disability-dec13-v2.pdf> on 10 August, 2017.

<sup>2</sup> Autism New Zealand (2017). Retrieved from [http://www.autismnz.org.nz/about\\_autism#4](http://www.autismnz.org.nz/about_autism#4) on 10 August, 2017.

<sup>3</sup> Ministry of Health, New Zealand (2017). Retrieved from <http://www.health.govt.nz/new-zealand-health-system/new-zealand-health-strategy-future-direction/challenges-and-opportunities> on 10 August, 2017.

<sup>4</sup> Ministry of Social Development, New Zealand. Retrieved from <https://www.msd.govt.nz/about-msd-and-our-work/work-programmes/policy-development/anti-social-behaviour/> on 10 August, 2017.

<sup>5</sup> New Zealand Association of Plastic Surgeons (2015). Retrieved from <http://plasticsurgery.org.nz/consumer-information/issues/dog-bite-injuries-a-serious-problem-in-new-zealand/> on 10 August, 2017.

So-called ‘addictive’ behaviour may also be harmful in humans. Examples include problem gambling, drug and alcohol abuse, over-eating, or excessive engagement with online gaming. When these patterns of choice are firmly established, quality of life may be negatively impacted through financial loss, relationship breakdown, loss of employment, criminal conviction, or injury and hospitalisation.<sup>6</sup> While there are many theories as to why some people are more prone to addictive behaviour than others (Netherland, 2012), research suggests the underlying mechanics of addictive behaviour require a series of small, regular choices that become more exclusive across time towards a particular behaviour (Herrnstein & Prelec, 1991; Heyman, 2013; Myers et al., 2016; Rachlin, 1997).

When behaviour does turn harmful, pharmacological interventions are often prescribed by medical and mental health professionals. Psychostimulants such as methylphenidate are prescribed to decrease hyperactive behaviour in people diagnosed with attention deficit hyperactivity disorder (ADHD) (Charach & Fernandez, 2013). Methadone or buprenorphine may be used to decrease opioid drug-taking behaviour (Jones, Campopiano, Baldwin, & McCance-Katz, 2015; Kampman & Jarvis, 2015). Risperidone, an anti-psychotic, is regularly prescribed to decrease aggressive behaviour in people with Autism Spectrum Disorder (ASD) (Sharma & Shaw, 2012), and sedatives are prescribed for individuals with high levels of aggressive or compulsive behaviour (Spencer et al., 2013). Drugs are also used to modify or control animal behaviour (Ogata & Dodman, 2011).

Although pharmacological interventions may decrease harmful behaviour by impacting the physiology and brain function of organisms; rarely do they directly target the behaviour. Therefore, once pharmacological interventions cease, or are compromised, harmful behaviour frequently recovers to pre-medication levels (Devanand et al., 2012; Subotnik et al., 2011; Wang et al., 2010). Drugs designed to decrease behaviour may also have negative side-effects, some of which pose a serious health risk (Deb, Sohanpal, Soni, Lentre, & Unwin, 2007; McKinney & Renk, 2011; Matson, Sipes, Fodstad, & Fitzgerald, 2011; Tyrer et al., 2008). In addition, pharmacological interventions may be over-prescribed through convenience and relative affordability, compared to the cost and availability of alternative approaches to behaviour reduction (Tsiouris, 2010). For these reasons, many individuals with harmful behaviour, and their caregivers, prefer non-pharmacological

---

<sup>6</sup> Ministry of Health, New Zealand (2016). Retrieved from <http://www.health.govt.nz/publication/research-report-new-zealand-drug-harm-index-2016> on 10 August, 2017.

treatment interventions for harmful behaviour (Cohen-Mansfield, 2013; McHugh, Whitton, Peckham, Welge, & Otto, 2013).

Of the range of non-pharmacological treatments available, there is a high rate of efficacy for behavioural interventions that use the principles of applied behaviour analysis (ABA) (Beavers, Iwata, & Lerman, 2013; Doehring, Reichow, Palka, Phillips, & Hagopian, 2014). ABA uses principles of learning to understand the function of behaviour, before attempting to directly target, and modify behaviour. Therefore, ABA asks ‘what is the behaviour *for*’, to understand which environmental variables need to change to see a decrease in behaviour (see Fisher, Piazza, & Roane, 2011, for an overview).

ABA-based interventions have been found to have a high rate of efficacy in reducing harmful behaviour in individuals diagnosed with ASD (Eldevik et al., 2009; Peters-Scheffer, Didden, Korzilius, & Sturmey, 2011; Reichow, 2012; Vismara, & Rogers, 2010), developmental disability (Didden, Korzilius, van Oorsouw, & Sturmey, 2006; Grey & Hastings, 2005; Harvey, Boer, Meyer, & Evans, 2009), and addictive behaviour (Carroll & Onken, 2005; Prendergast, Podus, Finney, Greenwell, & Roll, 2006). Research also suggests ABA interventions may be efficacious in the treatment of behaviour related to ADHD (Antshel & Barkley, 2008; Fabiano et al., 2009; Hodgson, Hutchinson, & Denson, 2014).

A popular ABA approach to the reduction of harmful behaviour is the provision of a high rate of reinforcers for an alternative behaviour, whilst blocking, or reducing, the rate of reinforcers maintaining harmful behaviour. The term ‘reinforcers’ refers to any event that has the effect of increasing the future probability of behaviour. For example, if behaviour produces an event such as food delivery, and the future probability of behaviour increases because of that event, then food delivery is a ‘reinforcer’ for that behaviour. ABA treatments that provide a higher rate of reinforcers for a wanted behaviour and block reinforcers for harmful behaviour are typically referred to as ‘differential reinforcement of alternative behaviour’ or DRA treatments (see Petscher, Rey & Bailey, 2009 for a review).

Health professionals, behaviour analysts, caregivers and animal trainers tend to view DRA interventions positively (Hastings, Boulton, Monzani, & Tombs, 2004; Kazdin, 1980; Lydon, Healy, O’Reilly & McCoy, 2013), primarily because they do not evoke conditional emotional responses such as fear or anger. In this way, DRA interventions are non-harmful, yet able to decrease rates of harmful behaviour associated with a range of developmental and psychological disorders (Doehring et al., 2014). DRA interventions may also be equally effective, or more effective, in reducing harmful behaviour compared with aversive

interventions, such as the use of punishment (Heyvaert, Saenen, Campbell, Maes, & Onghena, 2014).

However, an ongoing problem with DRA and similar treatments is that while harmful behaviour is often decreased during treatment, it is not always decreased to safe levels. For example, if an animal or person engages in self-injurious behaviour such as excessive licking (Eckstein & Hart, 1996), eye-gouging, head banging, or running into traffic, then even a very low level of behaviour is harmful (Symons & Thompson, 1997; Taylor, Oliver, & Murphy, 2011). A further problem with DRA interventions is that behaviour may relapse at higher than pre-treatment levels when treatment ends, or is compromised. In recent years this problem has generated a wave of research investigating the causes of post-DRA treatment relapse (Mace et al., 2010; Podlesnik & Kelley, 2015; Volkert, Lerman, Call, & Trosclair-Lasserre, 2009). However, although researchers have begun to model potential solutions to the problem of post-DRA treatment relapse, these solutions remain a work in progress (Podlesnik & Kelley, 2017).

If DRA and similar treatments fail to decrease harmful behaviour to safe levels, or enhance harmful behaviour post-treatment, the addition of punishment to DRA treatments (Fisher et al., 1998; Hagopian, Fisher, Sullivan, Acquisto & LeBlanc, 1998; Marschark & Baenninger, 2002; Wacker et al., 1990) may be condoned (Axelrod, 2013; 1990; Baumeister & Rollings, 1976; Dura, 1991; Favell & Lovaas, 1987; Gardner, 1969; Gerhardt, Holmes, Alessandri, & Goodman, 1991; Kahng, Iwata & Lewin, 2002; Lohrmann-O'Rourke & Zirkel, 1998; O'Brien, 1989; Picker, Poling, & Parker, 1979; Vollmer, Hagopian, & Bailey, 2011).

In everyday language, the term 'punishment' has become synonymous with terms such as 'pain', 'cruelty', and even 'abuse'. However, behavioural scientists have tried to use terms that better reflect the procedural intention of punishment (Hineline & Rosales-Ruiz, 2013; Skiba & Deno, 1991; Yulevich & Axelrod, 1983.) Punishment is an operant conditioning learning paradigm (Mazur, 2015). In operant conditioning, an organism learns its behaviour makes something happen. If the organism instead learns that an event may occur, but is not contingent on behaviour, then classical conditioning can occur. Therefore, in classical conditioning, organisms learn to associate an event with aspects of the environment, such as time and place. In operant conditioning, organisms learn that behaviour produces an event. Aspects of the environment, such as time and place, still play a role in operant conditioning because the organism learns that the same behaviour can

produce different events when aspects of the environment are different (for example, the consequences of shouting in a playground, compared with shouting in a library).

Within an operant paradigm, punishment is arguably best defined as the process by which an event – hereafter referred to as a *stimulus* – decreases the future probability of the response that produced it (Azrin & Holz, 1966; Church, 1963). The punishing stimulus may be something added to the environment like a loud noise, or it may be something taken away from the environment such as access to food or toys. Although these different processes are sometimes referred to as ‘positive’- or ‘negative’ punishment to denote the addition or subtraction of something from the environment, it is more parsimonious to think of a stimulus as a catalyst for change (Axelrod, 2013). If a stimulus produces a decrease in the future probability of behaviour, the stimulus is a punisher. If a stimulus produces an increase in the future probability of behaviour, the stimulus is not a punisher; it is a reinforcer. Even a stimulus with obvious aversive properties, such as electric shock, can function as a reinforcer, instead of, or in addition to, functioning as a punisher (Estes, 1944). Therefore, deciding what is and is not a punishing stimulus requires attention to the direction of behaviour after the stimulus has been produced.

In applied research, punishment has been found to successfully decrease a wide range of human and animal harmful behaviour, and may be particularly efficacious in reducing self-injurious behaviour (SIB) (e.g., Axelrod & Apsche, 1983; Baumeister & Baumeister, 1978; Beyra & Spinewine, 1973; Brandsma & Stein, 1973; Brantner & Doherty, 1983; Chapman, Smith & Layden, 1971; Charlop-Christy, & Haymes, 1996; Dale, Statham, Podlesnik & Elliffe, 2013; Didden, Duker, & Korzilius, 1997; Fisher et al., 1998; Hanley, Piazza, Fisher & Maglieri, 2005; Matson & DiLorenzo, 1984; Risley, 1968; Salvy, Mulick, Butter, Bartlett, & Linscheid, 2004; Sargisson, Butler & Elliffe, 2011; Smith, Michael & Sundberg, 1996; Tanner & Zeiler, 1975; Wood, 1984). Punishment interventions can permanently decrease harmful behaviour, or decrease behaviour for long periods (Altmeyer, Williams & Sams, 1985; Baroff & Tate, 1968; Church, 1969). A combination of punishing unwanted behaviour whilst simultaneously rewarding wanted behaviour has also been found to aid learning (Ilango, Wetzel, Scheich, & Ohl, 2010).

The success of punishment in decreasing harmful behaviour is perhaps unsurprising given research to suggest that if an organism cannot learn through punishment, it may struggle to socialise or survive (Bolles, 1970; Dugatkin, 2002; Gächter, Renner, & Sefton, 2008; Guzmán, Rodríguez-Sickert, & Rowthorn, 2007; Rockenbach & Milinski, 2006; Smith & Harper, 2003). Indeed, punishment-induced learning tends to be rapid (Azrin &

Holz, 1966; Gershoff, 2002; Skinner, 1953). Few people will intentionally touch a hot stove, approach a snarling dog, or poke a fork into an electrical outlet more than once. Even less physiologically discomforting punishing events, such as peer pressure or financial loss, have the potential to decrease behaviour over time (Balliet, Mulder, & Van Lange, 2011; Engelmann, & Nikiforakis, 2015; Fowler, 2005; Helbing, Szolnoki, Perc, & Szabó, 2010; Jensen, 2010).

Although punishment may lead to a rapid decline in harmful behaviour, the agentic use of punishment by authorities such as governments, schools, health professionals, and caregivers remains controversial, and is legally and ethically problematic (Fisher, 1987; Johnston, 1991; May, McAllister, Risley, Twardosz, & Cox, 1974; Repp & Singh, 1990; Singer, Gert & Koegel, 1999; Sidman, 1989, Van Houten et al., 1988). For example, the improper implementation of punishment procedures, including failure to obtain informed consent, may lead to the punishing agent receiving a criminal conviction – ironically, punishment for the use of punishment (Griffith, 1983). Ethically, punishment is also difficult to justify. Even the social and community benefits found to arise from forms of ‘altruistic’ punishment have been found to depend on the inability of the punished to retaliate (Janssen & Bushman, 2008; van den Berg, Molleman, & Weissing, 2012). In addition, the effects of punishment-induced learning may be more transient than reward-induced learning (Driscoll, 2005; Gershoff, 2002a), possibly because memory operates differently depending on whether punishment or reinforcers are consequences (Nakatani et al., 2009).

The most controversial issue related to the agentic use of punishment are known negative side-effects of aversive punishing stimuli. These include fear, aggression and other conditioned emotional responses (CERs) (Arhant et al., 2010; Axelrod, 2013; Berkowitz, 1993; Casey et al., 2014; Herron, Shofer, & Reisner, 2009; Galbraith, Byrik & Rutledge, 1970; Kazdin, 1980; LaVigna & Donnellan, 1986; Lerman & Vorndran, 2002; Lichstein & Schreibman, 1976; Newsom, Favell, & Rincover, 1983; Polsky, 1994; Roll & Unshelm, 1997; Schilder & van der Borg, 2004; Schalke et al. 2007; Seligman, 1968). Other side-effects include reduction of wanted behaviours (Rooney & Cowan, 2011), the development of new unwanted behaviours (Berkowitz, Cochran, & Embree, 1981; Risley, 1968), and even enhanced harmful or unwanted behaviour (Church, 1963; Flanagan, Goldiamond & Azrin, 1958; Hiby, Rooney & Bradshaw, 2004). It is difficult to determine the extent to which CERs and other negative side effects remain under-reported in punishment procedures. Studies that describe the negative side-effects of punishment tend to be

published to a lesser degree than studies that only report positive outcomes (Gerhardt, Holmes, Alessandri & Goodman, 1991).

Punishment may also be incorrectly applied in an attempt to decrease behaviour caused by unavoidable physical or neurobiological conditions. It may also be administered too late, confusing the learning process (Yamamoto, Kikusui, & Ohta, 2009). Further, automated delivery of punishing stimuli may be prone to technological failure (Sargisson, Butler & Elliffe, 2011). If any one of these scenarios occur, animals and humans may be forced to endure punishment interventions that will fail (Jonckheer-Sheehy, Delesalle, van den Belt, & Van den Boom, 2012; McGreevy & McLean, 2009) or even prove fatal (Grohmann et al., 2013). Alternatively, punishing interventions may succeed, but be impractical to administer over time. For example, the use of overcorrection, which requires the completion of a manually-intensive task contingent on harmful behaviour, is a popular punisher. However, overcorrection is less effective in the reduction of severe self-injurious behaviour (Harris & Ersner-Hershfield, 1978) and can be difficult to administer to adult participants physically able to resist an overcorrection task (Kelly & Drabman, 1977).

Although ethical concerns regarding the study and use of punishment should not be ignored, neither should the impact of harmful behaviour on the safety and well-being of humans and animals. Research shows that even very aversive punishing stimuli such as electric shock can produce positive side-effects including response generalisation, increased sociability, and positive emotional behaviour in some populations (Callias, Carr, Corbett & Jenkins, 1973; Duker, 1976; Hamilton, 1982; Lichstein & Schreibman, 1976; Linscheid et al., 1994; Luckey, Watson & Musick, 1968, Salvy et al., 2004; Whaley & Tough, 1970; Wright, 1973). Indeed, persevering with methods that fail to decrease harmful behaviour to safe levels may itself be considered ethically controversial practice (Balsam & Bondy, 1983; Butterfield, 1990; Gardner, 1969; Kemp, 1996; Murphy et al., 2005; Perone, 2003; Van Houten et al., 1988).

Clearly, there is an urgent need for the development of an approach to behaviour reduction that is non-harmful (i.e. does not produce negative side effects), and that works (Hineline & Rosales-Ruiz, 2013; Horner et al., 2005; Lerman & Vordran, 2002). Any such approach would ideally decrease harmful behaviour without producing CERs, without generating health risks, and without increasing pre-treatment levels of harmful behaviour. Developing an approach that meets these criteria, and investigating how this approach may be applied in behavioural treatments, is the focus of this doctoral thesis. Specifically, this thesis investigates whether it is possible to use simple learning processes to develop a

discriminative stimulus, then use that non-harmful stimulus to punish behaviour without recourse to other consequences.

A discriminative stimulus develops through stimulus discrimination – a process by which organisms learn that behaviour in the presence of one stimulus produces one consequence while the same behaviour in the presence of another stimulus produces a different consequence. For example, in the presence of a green traffic light, driving a car forward allows a driver to enter and exit an intersection, usually without risk of accident. However, driving a car forward in the presence of a red traffic light might result in the driver hitting another car or pedestrian. The behaviour, driving the car forward, remains the same in the presence of each of the different stimuli (the traffic lights). But the consequences for that behaviour, signaled by each stimulus, are different. Using the same process, organisms learn to respond in the presence of a rewarding stimulus, and to not respond in the presence of an unrewarding stimulus. A discriminative stimulus that signals behaviour in its presence will produce something rewarding is called a *positive discriminative stimulus (or S+)*, and a discriminative stimulus that signals behaviour in its presence will produce the absence of something rewarding is called a *negative discriminative stimulus (or S-)*.

This thesis investigates whether behaviour maintained by reinforcers will decrease if S- is made contingent on the same behaviour. If the future probability of behaviour decreases as a result, then the contingent S- stimulus will be a punisher.

If our investigations show that a contingent S- stimulus will punish behaviour without recourse to other consequences, the findings of this thesis will interest translational and applied researchers seeking non-harmful methods of decreasing harmful behaviour in people with developmental delays and neurobiological conditions, as well as typically-developing individuals prone to addictive behaviour. Therefore, the aims of this thesis are to investigate whether a contingent S- stimulus will a) punish the rate of behaviour maintained by positive reinforcers and b) punish choice processes thought to underlie addiction.

## CHAPTER 2.

### PUNISHED BEHAVIOUR: A LITERATURE REVIEW

Before we attempt to develop a contingent S- stimulus as a non-harmful punishing stimulus, it is important to understand how and why stimuli already proven to function as punishers succeed in reducing rates of behaviour or shifting choice. That is, we need to understand how different environmental variables impact the effectiveness of a punishing stimulus. As defined in the previous chapter, punishment does not describe the process of making an aversive stimulus follow the behaviour that produces it; punishment is the process by which the *future probability* of behaviour producing that stimulus decreases or shifts. Therefore, punishment is not about what happens to behaviour the moment a punisher is experienced – it is about what happens to behaviour *after that*. We need to understand which aspects of a punishing stimulus determine the extent to which behaviour decreases in the immediate, mid-, and long-term future.

#### *How does punishment decrease behaviour?*

For more than 100 years, behavioural scientists have used basic and applied research to understand how punishment works. These studies have sought to understand how behaviour initially comes under the control of a punishing stimulus, and how control is then maintained, or not maintained. Punishment research has also produced theories of punishment. For example, one of the earliest punishment studies, by Thorndike (1913), proposes that punishment ‘weakens’ behaviour in the same way that reinforcers ‘strengthen’ behaviour. Later, Skinner (1953) proposed that, rather than punished behaviour being weakened, behaviour likely decreases due to the availability of an alternative behaviour that allows the punishing stimulus to be avoided. This concept developed into the *alternative response assumption* theory of punishment (Dunham, 1971) which spawned *single-process*- and *two-process* (also known as *two-factor*)- theories of punishment. Single process theory assumes that any alternative response is reinforced by escape from the punishing stimulus, and that this reinforced response then competes with the punished response (Dinsmoor, 1954; Miller & Dollard, 1941). Two-process theory assumes the punished response first develops aversive properties by being regularly paired with the punishing stimulus (secondary conditioning). Then, in the second part of the process, the alternative response

disrupts that pairing. Once this occurs, the alternative response competes with the punished response as proposed by single-process theory.

Since 1960, the alternative response assumption has been frequently challenged and refuted. Azrin and Holz (1966) noted that the theory infers an alternative response will always be available, even in the absence of a punished response. Therefore, they argue, the alternative response must exist independently of the punished response. Schuster and Rachlin (1968) further observed that, in two-factor theory, both the secondary aversive properties of a punished response, and the reinforcer for the alternative response (i.e. avoidance), cannot be observed as they are internal to the organism. Therefore, their existence cannot be disproved. Dunham (1971) also suggested that because punished behaviour typically decreases at a more rapid rate than the rate at which alternative behaviour increases (Munavi, 1970), a punished response cannot be systematically disrupted by a reinforced alternative response.

Other studies agree with Dunham (1971) that punished behaviour decreases at a faster rate if an unpunished alternative response is available. However, they observe this fast rate of behaviour decline occurs only if reinforcer deliveries for the alternative response are equal to, or more frequent than, the rate of reinforcer deliveries for the punished response (Spradlin, 2002). When reinforcer deliveries are relatively less frequent for a concurrently available unpunished response, punished responding, or choice, will still decrease, but at a slower rate. This is supported by Todorov (1971), who found that pigeons on a schedule of concurrently available alternatives made fewer changeovers to a lean food schedule when a relatively richer schedule was punished with shock, compared to when the food schedules were equal. However, the pigeons eventually preferred the leaner schedule of reinforcement (also see Deluty, 1976). Negus (2005) showed that monkeys systematically reverse preference between a response-contingent schedule of food, and a self-administered schedule of cocaine, when either food- or cocaine-seeking is punished with histamine injections. Taken together, these findings suggest that the extent to which any stimulus functions as a punisher likely depends on the range of environmental contingencies available to the organism, rather than the presence of a competing alternative response.

One- and two-factor theories of punishment contributed to the development of two general models of punishment. In the *direct suppression* model, punishers directly reduce the strength of behaviour maintained by positive reinforcers (de Villiers, 1980). In the *competitive suppression* model, punishing one behaviour increases the value of reinforcers maintaining alternative behaviour, leading to a decrease in punished behaviour (Deluty,

1976). In support of one-factor theories of punishment, studies investigating punishment and choice with animals and with humans have produced data that better meet the predictions of the direct suppression model (e.g. Critchfield, Paletz, MacAleese & Newland, 2003; Farley, 1980).

Overall, one-factor theories of punishment such as the direct suppression model appear to fit the available data more closely than two-factor theories. If behavior maintained by positive reinforcers is directly suppressed by a punishing stimulus, why does this suppression occur? Recent research suggests the extent to which any stimulus functions as a punisher likely depends upon what that stimulus signals about events occurring in the future. That is, once the delivery of any stimulus functions as a discriminative stimulus for future events, behaviour will change according to the probability of those events occurring (see Baum, 2012; Cowie & Davison, 2016; Killeen & Jacobs, 2017; Shahan, 2017). Further, these changes should occur regardless of whether stimuli are appetitive (e.g., Cowie, Davison & Elliffe, 2011) or aversive (Ayllon & Azrin, 1966; Galbicka & Platt, 1984; Kelleher & Morse, 1968). Therefore, a stimulus previously thought to be excitatory might punish behaviour, and a stimulus thought to be aversive might enhance behaviour. In this way, no stimulus should be assumed to have aversive properties.

As far back as 70 years ago, Estes (1944) found electric shock could sometimes increase the rate of behaviour. Estes reinforced key-pecking for food in pigeons in the presence of non-contingent electric shock. When both food and shock ceased, responding was gradually eliminated. However, when shock was suddenly reintroduced, key-pecking immediately recovered at a high rate, despite the absence of food. The recovery of responding was likely due to shock functioning as a discriminative stimulus signaling the availability of food. Holz and Azrin (1961) found the extent to which response-contingent shock punished key-pecking for food in pigeons depended on whether shock deliveries also signaled reinforcer deliveries for key-pecking. Kelleher and Morse (1968) punished monkeys with shock for lever-pressing for food on a schedule lasting 11 minutes. Responses produced food every two minutes on mean, and produced shock only after the first 10 minutes had elapsed. In the eleventh minute, shock was delivered for every response, followed by timeout. Results showed the monkeys increased response rate across the first 10 minutes of the schedule, with response rate only decreased in the final minute. When shock was removed from the schedule, response rate was no longer enhanced in the first ten minutes or decreased in the final minute. However, when shock was delivered after three minutes instead of after ten minutes, response rate was decreased across each three-minute

period. Taken together, these findings suggest that whether supposedly aversive stimuli decrease or maintain behaviour depends on what those stimuli signal about likely future events. What all this contributes to our understanding of how punishment works is that the delivery of a putative punishing stimulus will only decrease behaviour to the extent its delivery signals that future behaviour will result in something the organism experiences as aversive. If the putative punishing stimulus instead signals an outcome experienced as rewarding, behaviour will likely increase.

If the nature of the stimulus is less important to behaviour than what the stimulus signals about future events, it follows that a successful punishing stimulus does not have to be harmful. This provides an opportunity to investigate the extent to which a non-harmful stimulus (i.e. one that does not produce conditioned emotional responses or other negative side-effects) may be developed as a punishing stimulus.

The remainder of this literature review uses a wide range of punishment literature to achieve three main goals. The first is to consider the extent to which punishment has been found to be efficacious in reducing harmful behaviour in comparison to alternative methods. The second is to discuss and understand how different environmental variables impact the ability of a given stimulus to function as a punisher. The final goal is to discuss literature that suggests discriminative stimuli may be able to function as punishing stimuli without recourse to further consequences for behaviour.

The value of all punishment research, including research with harmful stimuli, is that collectively, it delivers a foundation of understanding for the behavioural process of punishment and its influences. This foundation is valuable because it informs development of new approaches to punishment. Modern punishment research, including the studies within this thesis, is primarily concerned with how non-harmful stimuli may be harnessed to punish behaviour that contributes to negative life outcomes in humans.

## *Method*

### *Selection criteria*

Studies included in this review had to meet the following inclusion criteria. First, use of the term ‘punishment’ had to agree with the definition of punishment used in this review. That is, punishment is a decrease in the future probability of behaviour due to a stimulus produced by behaviour. Second, a putative punishing stimulus could be added to the environment or be a loss of reward previously gained. However, it could not be the unavailability of positive reinforcers going forward (i.e. timeout). This criterion was used

because the purpose of this thesis is to investigate whether a response-contingent negative discriminative stimulus will punish behaviour despite continued access to positive reinforcers. Therefore, the findings of studies that punish behaviour maintained by continued access to positive reinforcers are likely to be more informative than those that investigate the effects of timeout on behaviour.

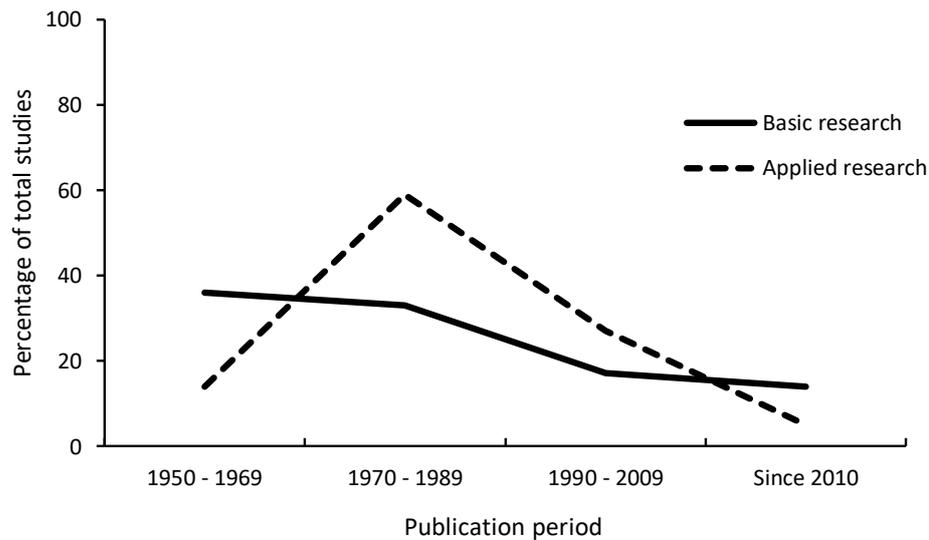
The third inclusion criterion for this review is that the selected studies must either a) investigate the extent to which different environmental variables impact control by a punishing stimulus, or b) investigate the efficacy of a punishing stimulus in decreasing behaviour compared to alternative methods. To meet these criteria, both basic and applied punishment studies were selected. Basic studies maintain a high level of experimental control to investigate the effect of different variables on the efficacy of a punishing stimulus. Applied studies use punishing stimuli to decrease harmful behaviour to improve life outcomes in treatment or training settings. For applied studies with animals, harmful behaviour was defined as behaviour able to cause physical harm to the animal itself, to other animals, or to humans. For applied studies with humans, harmful behaviour was defined as behaviour that directly contributes to loss of quality of life for the self, or others, or that results in physical injury or harm to self or others. Human punished behaviour tended to be associated with one of the following conditions or disorders: developmental disability, degenerative disease, traumatic brain injury, neurobiological disorders, psychopathologies, attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), and addiction.

### *Search methods*

Studies were published in English, in a peer-reviewed journal or book chapter, between 1950 and 2017. The primary databases used were Google Scholar, PubMed Central, ScienceDirect, and Web of Science. Searches using Web of Science were restricted to the fields of 'behavioural sciences' and 'veterinary sciences' to target a wider range of basic punishment research. The success of punishment in behavioural treatments is often measured by the speed and maintenance of behaviour decrease, by maintenance of those decreases across different settings, and by an absence of negative side effects (Schroeder, Mulick, & Schroeder, 1979). Therefore, search terms used in applied research included 'punishment', 'behaviour reduction', 'harmful behaviour', 'stimulus generalisation', and 'side effects'.

Together, these search criteria generated a list of 320 studies which were then subjected to abstract screening. Abstract screening assessed the value of each study to the literature review based on the criteria described in the previous section. Following abstract screening, selected studies or book chapters were read in entirety to assess suitability. Meta-analyses were placed in a separate category, with the individual studies contained within each meta analyses screened, and included if they met the review criteria. On completion of this process, 144 studies were selected for this review. There were 82 basic research studies (see Appendix 1) published between 1952 and 2017, and 62 applied research studies (see Appendix 2) published between 1964 and 2017. For both basic and applied studies, the year of publication, species, sample size, the behaviour targeted for punishment, the punishing stimulus used, dependant measures, and whether the punishing stimulus was used alone or in combination with another stimulus, were recorded. For basic research studies, manipulated independent variables were recorded. For applied studies, punishment efficacy, efficacy in relation to alternative methods, and post-punishment efficacy were recorded.

Figure 2.1 shows when the selected studies were published. Most of the basic research studies were published between 1950 and 1969, and most of the applied studies were published between 1970 and 1989. Of the 82 basic research studies, 48% were reported in the *Journal of the Experimental Analysis of Behaviour*, with the remainder distributed approximately evenly across a total of 24 other journals. Of the 62 applied research studies, 21% were reported in the *Journal of Applied Behaviour Analysis*, with the remainder distributed approximately evenly across a total of 25 other journals. Note the rapid decline in basic punishment studies between 1970 and 2009. The rate of this decline appears to have slowed since 2009.



*Figure 2.1* Percentage of selected basic- (solid line) and applied- (dashed line) punishment studies (y-axis) published in 19-year time periods (x-axis) from 1950 to 2009, and since 2010.

#### *Applied research overview*

Given that the agentic use of punishment is controversial, is important to consider why new approaches to punishment are needed. One justification is that punishment has been shown to decrease harmful behaviour when alternative methods fail. But to what extent is punishment more successful compared with alternative methods? Does one punishing stimulus able to decrease self-injurious or aggressive behaviour also decrease other types of harmful behaviour (Fisher et al., 1994)? Does behaviour decreased by punishment remain at low levels across the post-treatment period, and for how long? As punishing stimuli can generate negative side effects (Hiby, Rooney & Bradshaw, 2004; Linscheid et al., 1994; Rooney & Cowan, 2011; Salvy et al., 2004), we need to understand how long it takes for punishing stimuli to decrease harmful behaviour, and the environmental and procedural conditions that may hasten this process. We use applied research studies within this review to address these questions, and others.

Of 62 applied studies, 94% were with human participants. The most common human behaviour targeted for reduction was self-injurious behaviour (66%), followed by aggression towards others or property (19%), and stereotypy (13%). The most common conditions associated with these behaviours were developmental disability (66%), followed by ASD (16%). Punishing stimuli included shock, ammonia, water-mist, points- or token loss, facial screening (Barrett, Matson, Shapiro & Ollendick, 1981) and over-correction (Azrin et al., 1975; Azrin & Wesolowki, 1975; Matson & Stephens, 1997, 1978).

Overcorrection requires the completion of a manually-intensive task contingent on harmful behaviour.

A punishing stimulus was used alone in 58% of applied studies (n=37). A primary punishing stimulus was presented in combination with a secondary punishing stimulus, or with alternative methods of behaviour reduction, in 42% of studies (n=25). The most common alternative methods of behaviour reduction were DRA or DRO procedures (70%), timeout from positive reinforcement (25%), and physical restraint (45%).

We wanted to know the extent to which punishment was more effective than alternative methods of behaviour reduction. Slightly more than half (n=34) of applied studies used a punishing stimulus only after earlier attempts to decrease behaviour through alternative methods had failed. Of these, 85% found that a punishing stimulus was more effective than alternative methods in reducing harmful behaviour. Decreases in behaviour across these studies ranged from 30% to 100%. Of 40 studies with participants with developmental disability, punishment decreased self-injurious behaviour (SIB) by 93%, aggression by 95%, and stereotypy by 97%, on mean. Of 12 studies with individuals with ASD, punishment decreased SIB by 93%, aggression by 95%, and stereotypy by 94%, on mean. For studies with individuals with either developmental disability or ASD, 35% used shock as a primary punishing stimulus, followed by overcorrection (29%). Other punishing stimuli included facial screening, physical restraint, water mist, verbal reprimands, and unpleasant tastes and smells. Incidence of side-effects was monitored in most studies (99%). No side-effects were found in 87% of studies (but see Gerhardt, Holmes, Alessandri & Goodman, 1991.) Negative side-effects such as aggression, verbal protest, and the development of new unwanted behaviour, were found in 11% of studies. Only 2% of studies reported positive side-effects over and above decreases in harmful behaviour. Positive side-effects included increased communication, social interaction, and improved emotional state. Table 2.1 shows the extent to which harmful behaviour was decreased by punishment according to behaviour type, related condition, and post-treatment maintenance in 52 of the 62 applied studies selected.

This review also asked whether behaviours other than SIB and aggression have been successfully punished, and the extent to which decreases in behaviour are maintained across time. Of the 62 applied studies, 52 targeted SIB, aggression and stereotypy. The remaining 10 targeted behaviour including destruction of property (5), addiction to cigarette smoking (1), food regurgitation (1) predatory aggression (1) and socially problematic behaviour (2). In treatments to decrease property destruction, four of five studies used shock as the

punishing stimulus, with one study using overcorrection. Of these five studies, three decreased incidence of property destruction to low levels. Of those, two used shock in combination with DRA or verbal reprimands (Simmons & Lovaas, 1969), and one used overcorrection with DRA.

Addiction to cigarette smoking was successfully punished for longer than 12 months in 30% of 23 human participants by using shock in combination with DRA and cognitive behavioural therapy (CBT) (Chapman, Smith & Layden, 1971). Shock also punished life-threatening regurgitation of food in a 9-month old infant, resulting in the full elimination of regurgitation across an 8-month post-punishment follow-up period (Toister, Condon, Worley, & Arthur, 1975). Socially problematic behaviour, including severe stuttering, and running around in a crowded classroom, were punished with shock, and overcorrection, respectively. Shock used in combination with visual screening decreased stuttering by up to 60% (Berez, 1972), and overcorrection used alone decreased classroom running by 95%. Both these studies reported a post-treatment maintenance period of more than 12-months (Gordon, Handleman, & Harris, 1986). Similarly, in an applied animal study, Dale, Podlesnik and Elliffe (2017) found that shock decreased predatory approaches to a training stimulus in 78% of 1156 dogs, and that the effects of punishment were maintained over a 12-month post-training period.

The above findings suggest that while punishing stimuli can punish a range of harmful behaviours, the extent of behaviour reduction, and post-punishment maintenance of behaviour reduction, might vary according to behaviour type, or underlying conditions. For example, shock was highly efficacious in achieving short- and long-term decreases in aggression and SIB for those with developmental disability or ASD, but considerably less effective at achieving short- and long-term decreases in rates of cigarette smoking, stuttering and property destruction in typically-developing humans, and predatory approaches in dogs. Therefore, new approaches to punishing stimuli need to consider how a non-harmful punishing stimulus might be developed to be effective in the short- and long term, and across a range of behaviours. Towards this, basic research studies such as those discussed in the next section deliver valuable insight into how punishing stimuli can be impacted by different environmental variables.

Table 2.1

*Total Applied Punishment studies (N) by Behaviour type and Condition, with Range and Mean Reduction, as a Percentage, and Mean Post-punishment Reduction Period in months.*

Behaviour type by related condition	Total studies (N)	Reduction range during punishment period (%)  (Mean %)	Post-punishment reduction period range (months)  Mean (months)
<b>SELF-INJURY</b>			
- Developmental disability	27	50 - 100 (93)	2 - 60 (12.16)
- ASD	3	80-100 (93)	1 – 20 (6.67)
<b>AGGRESSION</b>			
- Developmental disability	10	80 – 100 (95)	3 - 54 (17.83)
- ASD	4	90 - 99 (95)	NA
<b>STEREOTYPY</b>			
- Developmental disability	3	90 - 100 (97)	1 – 6 (3.50)
- ASD	5	80 – 100 (94)	1 – 10 (4.00)

*Basic research overview*

Of the 82 basic punishment studies selected for this review, 84% used animals and 16% used humans. The most commonly used animal subjects were rats (51%), followed by pigeons (29%). Other animal subjects included monkeys (12%), dogs (Steiss et al., 2007), cats (Stokman & Glusman, 1968), bees (Smith, Abramson & Tobin, 1991), and crabs (Abramson & Feinman, 1987). Shock was the most commonly used punishing stimulus

(68%). Other punishing stimuli used include noise or tones (Tolman & Mueller, 1964), ammonia smell (Altman, Haavik & Cook, 1978; Tanner & Zeiler, 1975), strong light (Barker et al., 2010), water mist (Arntzen & Werner, 1999; Dorsey, Iwata, Ong & McSween, 1980), lemon juice (Libet, Sajwaj & Agras, 1973, 1974), tabasco sauce, eye-puffs, and ingested or injected drugs or chemical compounds. Tickling has also been used as a punishing stimulus in basic research (Greene & Hoats, 1971).

Basic research shows that punishing stimuli can decrease even food-seeking and energy-saving behaviour. For example, Wesp, Lattal and Poling (1977) found that high intensities of shock punished auto-shaping of key-peck responses in pigeons. Auto-shaping (Brown & Jenkins, 1968) describes a process in which pigeons immediately receive a reinforcer every time they peck a key signaling the availability of food. Within a short time, pigeons learn to peck keys for food without training. Similarly, Dardano (1968) found that shock could punish attempts to escape rising response cost in pigeons. He arranged a progressive-ratio schedule in which obtaining food reinforcers required a progressively greater number of responses. Pecking an escape key both decreased the PR requirement and delivered shock. Because of punishment, the pigeons made fewer responses to escape the PR schedule.

Azrin (1970) found that, provided response-contingent shock was delivered at high intensity, shock-induced aggression in monkeys could be decreased. This showed that behaviour induced by one stimulus can be punished by the same stimulus, once that stimulus is made response-contingent (Baron, Kauffman & Fazzini 1969; Follick & Knutson, 1978; Laurence, Himeline & Bersch, 1994; Roberts & Blasé, 1971). Poling and Thompson (1977) found that response-contingent shock experienced by rats' lever-pressing for food on one schedule decreased lever pressing for ethanol on another schedule. This likely occurred because lever-presses for ethanol produced response-contingent shocks on the food schedule.

In the late 1950's, and throughout the 1960's, Azrin and colleagues conducted several basic research studies designed to investigate punishment as a behavioural process, and the variables that impact that process (Ayllon & Azrin, 1966; Azrin 1958, 1959, 1960; Azrin & Holz, 1961; Azrin, Holz & Hake, 1963). Although shock was used to punish the behaviour of rats, pigeons and monkeys in many of these studies, noise was used to punish plunger pulls for tokens (Ayllon & Azrin, 1966), and switch-presses for task completion (Azrin, 1958), in humans. Using the combined results of these studies and others, Azrin and Holz (1966) identified a range of variables known to impact the ability of a response-

contingent stimulus to function as a punisher. The following sections discuss some of these variables.

### *Type and intensity of punishing stimulus*

The use of shock in basic punishment studies affords a high level of experimental control because the intensity of the shock is easily regulated. When shock is not used, it can be difficult to measure the effects of changed stimulus intensity on behaviour, to know whether changes in stimulus intensity have been discriminated, or to know if changes in shock intensity are discriminated, but do not differentially impact the rate of behaviour. For example, Friedel, Hart and Odum (2017) found that using tone as a punishing stimulus in rats equally decreased the rate of lever-pressing for food, whether the tone was 1 Khz or 22Khz.

For studies that use shock, although mild levels of shock intensity will decrease behaviour (Rachlin, 1966), punishment is more effective, and responding less likely to recover, when high intensity shock is used (Appel, 1963; Azrin, Holz & Hake, 1963; Dinsmoor, 1952, Hake, Azrin & Oxford, 1967). Shock introduced at high intensity from the beginning of a trial is also more effective as a punisher than shock that gradually increases in intensity across time (Brethower & Reynolds, 1962). Shock introduced at high intensity has also been found to attenuate increases in behaviour caused by a delay in shock delivery (Banks & Vogel-Sprott, 1965; Cohen, 1968). Everly and Perone (2012) found that when high levels of shock intensity were used to punish longer, or shorter, inter-response times (IRTs) between lever-presses in rats, response rate respectively increased or decreased. However, IRT-contingent shocks of mild intensity had no impact on the rate of lever-pressing for food (also see Arbuckle & Lattal, 1992; Galbicka & Platt, 1984).

Increasing the intensity of a punishing stimulus has also been found to decrease rates of drug-seeking in animals and humans. As drug-seeking behaviour is often resistant to punishment (Hopf et al., 2010, Lesscher, van Kerkhof, & Vanderschuren, 2010), using a punishing stimulus of high intensity may decrease addictive behaviour. For example, Grove and Schuster (1974) showed that high levels of shock intensity decreased lever-pressing for cocaine in monkeys, irrespective of the magnitude of cocaine dosage. Bergman and Johanson (1981) also used shock to punish monkeys lever-pressing for cocaine. They found low levels of shock intensity did not disrupt the rate of responding for cocaine, and medium levels only temporarily decreased response rate. However, lever-pressing for cocaine was decreased to low levels when high levels of shock intensity were used.

The extent to which more intense stimuli are more efficacious punishers might also depend on the schedule and magnitude of positive reinforcers maintaining behaviour. For example, Filby and Appel (1966) found that during the post-punishment period, lever-pressing for food by rats was relatively slower to recover if previously punished with high levels of shock intensity, compared with lower levels of intensity. However, if lever-pressing was maintained by a high rate of reinforcers during punishment, although high shock intensity still eliminated lever-pressing, post-punishment behaviour recovered more rapidly compared with lever-pressing maintained by a relatively lower rate of reinforcers. Sierra-Machado et al. (2015) used eye puffs to punish both humans and primates button-pressing for tokens. Punishment probability and reinforcer magnitude were both manipulated. When there was a high probability of punishment, subjects did not respond for smaller reinforcers, but continued to respond for larger reinforcers. Similarly, Azrin (1959, 1960) found that increasing shock intensity decreased the rate of key-pecking for food in pigeons only when food was delivered on a variable-interval (VI) schedule. When food was instead delivered on a fixed-rate (FR) schedule, increasing shock intensity increased the length of the post-reinforcer pause, but did not decrease response rate. A VI schedule arranges reinforcer deliveries every so many seconds or minutes on mean, and a FR schedule arranges reinforcer deliveries following a set number of responses (also see Scobie & Kaufman, 1969).

The effect of more intense punishing stimuli may not only be mitigated by reinforcer magnitude, but by reinforcer type. For example, Geller (1970) found that while low levels of shock intensity decreased lever-pressing for food in rats, high levels of shock intensity did not decrease the rate of lever-pressing for electrical brain stimulation (EBS). Beyra & Spinewine (1973) also tried to punish lever-pressing for EBS in rats. Although shock successfully decreased the rate of lever pressing for low levels of EBS, shock had no effect on responding for mid-level EBS, and actually increased the rate of lever-pressing for optimal-level EBS. Bright light also failed to punish responding for EBS. Taken together, these findings suggest that punishment efficacy might depend less on the type of punishing stimulus, and more on the relationship between the intensity of a punishing stimulus and the magnitude or type of reinforcers.

If the type of punishing stimulus used is less important to punishment efficacy than other environmental variables, this is good news for researchers investigating non-harmful approaches to punishment. However, a disproportionate use of shock in punishment studies has left a gap in the research. As a result, we know little about how we might manipulate

the intensity of a negative discriminative stimulus (S-) used as a punisher. Do we present it for longer, or do we train it to signal the absence of a relatively greater rate of reinforcers? These are important matters because if a contingent S- stimulus is in any way less salient than shock, which is likely, careful manipulation of environmental variables may be required to achieve punishment efficacy.

### *The role of drugs in punishment efficacy*

Any new approach to punishment, including the development of new forms of punishing stimuli, require a sound understanding of how punishment contingencies interact with the administration of drugs. This is particularly important given that harmful behaviour is commonly treated with drugs, and behavioural interventions implemented in conjunction with drug treatment.

Drugs may be used as punishers, reinforcers, or their use may affect the impact of punishing or reinforcing stimuli. Hanson, Witoslawski, and Campbell (1967) found that pentobarbital, chlordiazepoxide, and meprobamate increased lever-pressing for food by monkeys, while scopolamine, d-amphetamine and chlorpromazine either decreased or increased responding depending on other discriminative stimuli available in the environment. Falk and Burnidge (1970) punished water drinking in rats with salt (NaCl). Although the rate of drinking decreased, this reduction was attenuated in rats who were given an injection of chlordiazepoxide. Martin, Moreau, Jenck, and Cumin (1993) punished mice with shock for lever-pressing for food. They found shocked response rate still decreased if the mice were given diazepam but not if they were given buspirone. Podlesnik, Jimenez-Gomez, and Woods (2010) used intravenous histamine injections to punish choice in rats. Histamine was used to continuously punish responding for food on one of two concurrently available levers reinforced at the same rate. Results showed that rats preferred the unpunished lever at higher levels of histamine dosage, while lower levels of dosage did not impact choice (also see Podlesnik & Jimenez-Gomez, 2013). Li Hasio and Li (2013) found that the amount of sucrose solution ingested by rats depended on the magnitude of the dosage of lithium chloride (LiCl) administered to rats following ingestion. Leong et al. (2016) used Lithium Chloride (LiCl) injections to punish rats lever-pressing for cocaine infusions. The length of time in which the rats responded for cocaine infusions prior to punishment was manipulated into short- and long-term access groups. Leong et al. found that LiCl punished responding for short-, but not long-term access to cocaine. This suggests the extent to which punishment decreases drug-seeking behaviour may be attenuated by

longer periods of access to drugs prior to punishment (also see Pelloux, Everitt & Dickinson, 2007). Taken together, these findings suggest that drugs may increase resistance to punishment, or have no impact on punished response rate, and that these differences might depend on the type of drug, or be dose dependent. In addition, the ability of a punishing stimulus to decrease behaviour maintained by drugs, may depend on the amount of prior access to positive drug reinforcers.

### *Learning history*

The efficacy of a punishing stimulus in decreasing behaviour may be compromised by learning history. That is, by the extent of any prior exposure to the punishing stimulus, and to other stimuli in the environment. Myer (1967) manipulated both the intensity of shock as a punishing stimulus, and the prior mice-killing experience of rats, to punish mice-killing by rats. He found that experienced mice-killing rats were more resistant to shock punishment than rats less-experienced with mice-killing. In addition, increasing shock intensity did not decrease the rate of mice-killing in rats with greater mice-killing experience. Holz, Anker, Regier, Claxton & Carroll (2013) investigated the extent to which response-contingent injections of histamine would punish lever pressing for cocaine in rats. They found the extent to which histamine injections decreased responding for cocaine depended upon how long the rats had been able to self-administer cocaine. Responding for cocaine decreased for rats that had limited prior access to cocaine. However, rats with longer prior exposure did not alter their response rate following histamine injections.

Riccio and Hamm (1972) demonstrated how prior exposure to a stimulus could impact the effectiveness of that stimulus as a punisher. They punished two groups rats with cold water immersion for accessing a goal box for food. One group of rats had been previously exposed to cold water immersion with no other contingencies. Punishment tests showed greater speed of goal box access by rats with previous experience of cold water immersion, compared with those who were not familiar with cold water immersion. Taken together, these findings suggest that new approaches to the development of punishing stimuli may be more successful if behaviour maintained by positive reinforcers is punished soon after being established, rather than when behaviour has been established for a long time. In addition, organisms may 'acclimatise' or become habituated to a punishing stimulus if they have extended prior experience of stimulus.

It is possible that using a contingent S- stimulus as a punisher will avoid degradation of the punisher by habituation. Unlike cold water immersion, or shock, the non-harmful

nature of a contingent S- stimulus means organisms will likely experience it as only a signal for the absence of reinforcers. Therefore, any mitigation of its punishing properties would more likely occur due to the accidental pairing of the contingent S- stimulus with a reinforcer delivery, than by habituation.

### *Individual differences*

The extent to which punishing stimuli will decrease rates of harmful behaviour may be mediated by differences in sex, species, age, psychopathology, or phenotype. For example, Ernst and Yee (1975) investigated the interaction between age and punishment by using shock to punish younger and older rats maze-running for food. Despite manipulating training time for both groups of rats, they found younger rats were more resistant to shock punishment than older rats, and vice versa, irrespective of training time. van Oyen, van der Zwan, van de Poll, and Walg (1981) found that lever-pressing for food in female rats, recovered faster compared to male rats when shock punishment ended. A sex difference was also observed by Berez (1972) who found that shock was an effective punisher of cigarette smoking in males, but not females.

Differences in psychopathology may also impact the efficacy of a punishing stimulus. Martin, Cox, Brooks and Savage (2014) found that a group of female smokers were more sensitive to punishment in the form of monetary losses, and less sensitive to monetary gains, than female non-smokers. Similarly, Furukawa et al. (2017) found that choice by children with ADHD was more sensitive to punishment by reward loss than a non-ADHD control group. Sawyer et al. (2016) found adults with PTSD reacted faster on tasks punished by monetary loss than a non-PTSD control group.

Torres et al. (2017) studied two groups of rats which were shock-resistant or shock-sensitive by phenotype. When he used shock to punish responding for methamphetamine intake, response rate decreased only for the shock-sensitive rats. However, *all* rats relapsed during extinction after being presented with drug-seeking cues. This suggests the extent to which individual differences impact punishment efficacy may depend on other discriminative stimuli in the environment.

### *Partial and continuous reinforcement and punishment*

One variable that may impact the effectiveness of either a reinforcing or punishing stimulus is the extent to which that stimulus follows every response. Several studies have found that if every reinforced response is punished (continuous punishment) behaviour

decreases more rapidly, and to a greater degree, than if only a proportion of reinforced responses are punished (i.e. intermittent, or partial punishment) (Hendry & Van-Toler, 1964; Lerman, Iwata, Shore, & DeLeon, 1997; Linden, 1976; Zimmerman & Ferster, 1963). Organisms with prior experience of partial punishment tend to be more resistant to continuous punishment than organisms with no experience of punishment. This phenomenon is called the *partial punishment effect* or PPE (Banks, 1966; Miller, 1960). There is a corollary effect for reinforcement. When responses are only occasionally reinforced, responding tends to be more persistent when all reinforcers are removed in extinction, compared with if all responses were previously reinforced. This phenomenon is called the *partial reinforcement extinction effect* or PREE (see Mackintosh, 1974 for a review).

Some studies suggest that the PPE and PREE share a common mechanism because both feature resistance to change of behaviour producing aversive events (Brown & Wagner, 1964; Wagner, 1966.) That is, with PPE, behaviour produces a punishing stimulus, and with the PREE behaviour produces an absence of reinforcers. To test for interaction between the PPE and PREE, Deur and Parke (1970) investigated the hitting behaviour of 120 school age children. The consequences of hitting an inanimate object were one of: a) continuous reinforcement and no punishment, b) partial reinforcement and no punishment, or c) partial reinforcement and partial punishment. The punisher was a combined stimulus of aversive noise with a verbal rebuke. The reinforcer was access to marbles. When tested in extinction and with continuous punishment, Deur and Parke found that behaviour that was partially reinforced and partially punished in training was both the most resistant to continuous punishment, and the most resistant to extinction. This was attenuated for behaviour that was partially reinforced but not punished in training, and even more attenuated for behaviour that had been continuously reinforced in training. These findings suggest that if partial reinforcement and partial punishment are both used in training they have an additive effect, enhancing both the PREE and the PPE.

Halevy, Feldon and Weiner (1987) also tested the PREE and PPE using three training conditions. Runway runs by rats produced one of: a) continuous food reinforcement and no punishment, b) partial food reinforcement and no punishment or c) continuous reinforcement and partial shock punishment. Half the rats in each condition were then tested in extinction (no reinforcers and no shocks), and the other half were tested with continuous punishment and continuous reinforcement. Therefore, the difference between the procedures of Halevy et al. and Deur and Parke (1970) is that partial punishment and partial

reinforcement did not occur together in a training condition. Halevy et al. found that, compared with continuously reinforced rats, the PREE occurred when rats transitioned from partial reinforcement to extinction, and the PPE occurred when rats transitioned from partial punishment to continuous punishment. However, partially reinforced rats did not show relatively increased resistance to continuous punishment, and partially punished rats did not show relatively increased resistance to extinction. These results suggest that the PREE and PPE are enhanced only when partial reinforcement and partial punishment occur together in training.

As the studies that comprise this thesis will show, reinforcers maintaining behaviour are not withheld during training or punishment test conditions. Therefore, there are no conditions in which behaviour that was previously reinforced, ceases to be reinforced. Without an extinction condition, there is no opportunity for the PREE to occur. In addition, as training sessions in the studies of this thesis will not use punishment, and punishment test sessions use a partial punishment schedule, there is no transition from partial to continuous punishment. Therefore, the PPE cannot occur. However, it is possible that any observed reductions in behaviour in the studies of this thesis may be increased in future studies that use continuous punishment during test sessions.

#### *Punishment and choice: with addiction in mind*

An advantage of studies that investigate punishment and choice is that they deliver an increased understanding of whether harmful choice behaviour can be shifted by punishment, and the extent to which certain variables may compromise those shifts.

In an early study investigating the impact of punishment on choice, Holz (1968) showed that punishing relatively richer and leaner choice alternatives with shock decreased responding to both alternatives at similar rates. Davison (1970) later showed that when rats were given a choice between two schedules that required different behaviour – either a set number of responses for a greater amount of food, or a set period of no responses for a relatively lesser amount of food – rats preferred the behaviour that delivered the greatest overall rate of reinforcers. However, when the last response in the set number of responses was punished with shock, inter-response times (IRTs) increased, indicating a shift in preference towards the unpunished alternate behaviour (also see Dunham, 1972; Sizemore & Maxwell, 1993). Conversely, Fox and Pietras (2013) presented four adult humans with choice between two schedules of monetary reinforcement, each with different response requirements. Participants were instructed regarding the best choice strategy to maximise

rewards across both schedules. When the accuracy of the instructions was manipulated so that *not* following instructions produced greater overall reward, punishment with money loss for rule-breaking only increased rule-following while rule-breaking did not provide a greater overall amount of reinforcement (also see Reed & Yoshino, 2008). The difference in the findings of Davison (1970) and Fox and Pietras (2013) may be accounted for by species, or by the potential for shock to be more aversive than money loss.

Pelloux, Murray and Everitt (2015) used shock to punish cocaine-seeking in rats following long-term unpunished access to cocaine. One group of rats had access to an alternative response reinforced by sucrose, whereas a control group did not. Although punishment decreased the rate of cocaine-seeking in both groups, reduction was enhanced for rats with access to an alternative reinforced response. However, Pelloux et al. found that a sub-group of rats continued to respond for cocaine at high rates despite punishment, and despite access to an alternative reinforced response. This suggests that the extent to which the availability of an alternative reinforced response may enhance punishment of established drug-seeking behaviour will likely depend on individual differences.

It is possible that the findings of Pelloux et al. (2015) may be partially accounted for by consummatory aspects of drug-taking. Rodriguez and Logan (1980) gave rats a choice between two runways that provided access to drinking water. On one runway, the rats were punished with shock prior to drinking, and on the second they were punished with shock *after* drinking. Rodriguez and Logan found the rats preferred to be punished prior to drinking, concluding that rats may prefer punishment of an instrumental response over punishment of a consummatory response. Therefore, it is possible that punishment of drug-seeking behaviour may be more effective than punishment of drug-taking behaviour. This suggestion is supported by Herman and Azrin (1964). Rather than punish cigarette smoking, they used aversive noise to punish humans using an apparatus to access cigarettes. Noise was found to successfully punish the rate of access attempts, and access attempts were further decreased when an equally reinforced unpunished alternative behaviour was made available.

Some studies have investigated the predictions of the Matching Law when choice produces punishing stimuli in addition to positive reinforcers. The Matching Law (Herrnstein 1961, 1970) dictates that the overall proportion of time or responses allocated to any one alternative among competing alternatives will equal the overall proportion of reinforcers obtained from that alternative. Bradshaw, Szabadi and Bevan (1979) demonstrated that when two concurrently available and equally reinforced alternatives were

presented, choice was a function of the relative reinforcer rate between alternatives. However, when choice on one alternative was punished with monetary loss, choice shifted towards the unpunished alternative, violating the Matching Law.

Punishing stimuli have also been used to shift suboptimal choice. Suboptimal choice describes the process by which organisms behave so as to forfeit a higher overall reinforcer rate (e.g., Spetch, Belke, Barnet, Dunn, & Pierce, 1990; Zentall, 2014, 2016a, 2016b). For example, Shimp, Mitchell, Beas, Bizon, and Setlow (2015) showed that rats' choice of high-risk rewards reduced when such choices also produce shocks, and that this effect of shock increases as the magnitude of the shock increases. In contrast, increasing the magnitude of the high-risk reward in the absence of punishment did not increase the proportion of high-risk choice. Similarly, Orsini et al. (2016) showed that shock punishment decreased choice for a smaller, sooner reward and shifted choice toward a larger, later reward in both male and female rats. Therefore, these studies show that punishing stimuli can shift choice, resulting in more optimal choice outcomes. Literature investigating the impact of different environmental variables on suboptimal choice is discussed in more detail in Chapter 6.

#### *Harnessing the power of discriminative stimuli*

Organisms appear to value antecedent discriminative stimuli that signal environmental contingencies. For example, Badia, Coker and Harsh (1973) gave four rats a choice between an unavoidable schedule of shock delivery signalled by a tone, and an unsignalled schedule of shock delivery. Frequency of shock was then systematically increased across conditions. All rats preferred signalled shock when the frequency of signalled shock was the same or double that of unsignalled shock; three rats continued to prefer signalled shock at four times the frequency of unsignalled shock, and one rat maintained preference for signalled shock at eight times the frequency of unsignalled shock (also see MacDonald, 1973). In another study, rats were found to prefer longer and more intense signalled shock over shorter and less intense unsignalled shock (Badia, Culbertson & Harsh, 1973). These findings suggest that organisms not only prefer the presence of discriminative stimuli able to signal future events but are prepared to pay a price to ensure such stimuli are present rather than absent.

If a discriminative stimulus signals an aversive consequence, and a reduction in behaviour occurs in the presence of the discriminative stimulus, then punishment is assumed to be under the control of the discriminative stimulus. However, in a meta-analysis of punishment studies, Doughty, Doughty, O'Donnell, Saunders and Williams (2007) found

few punishment studies published between 1990 and 2007 conclusively demonstrated control by an antecedent discriminative stimulus. Those in which punishment did appear to be under stimulus control tended to use pigeons, shock as the punishing stimulus, and use partial reinforcement and punishment schedules. Doughty et al. found most punishment studies did not show that rates of responding decreased after the onset of an antecedent stimulus and before delivery of a punishing stimulus. Therefore, it was not possible to know the extent to which decreases in response rate occurred due to punisher delivery or due to control by the antecedent discriminative stimulus. As noted by Doughty et al. if behaviour is under the control of the punisher delivery, and not an antecedent stimulus, it is unlikely that post-punishment behavioural reduction will be maintained by presenting the antecedent stimulus in new settings. This is important as studies show that a setting may also function as a discriminative stimulus for likely future outcomes for behaviour occurring in that setting. Bouton and Schepers (2015) reinforced rats with food for lever-pressing in one context and used shock to punish lever-pressing for food in a different context. When both food and shock were removed from relevant contexts, relapse occurred in the first-, but not the second context. This suggests the rats associated punishment with lever-presses made in a particular context, rather than all lever-presses. If lever-pressing for food had been punished with shock in both contexts, we might not expect context to develop as a discriminative stimulus for response-contingent shock. Therefore, the extent to which punishment for the same behaviour is maintained across different contexts will depend on whether context functions as a discriminative stimulus for response-contingent punishment, (see Corte, Wolfe and Locke, 1971, for an applied example), or whether punishment is under the control of a different stimulus.

Balaban, Rhodes and Neuringer (1990) demonstrated how punishing and non-punishing discriminative stimuli can interact. Two groups of human participants were given a learning task in which a tone followed errors (one discriminative stimulus). For one group, partial shock punishment (another discriminative stimulus) was delivered immediately after the error tone. Results showed that participants punished with shock maintained higher error rates compared with those whose errors produced only the tone. Balaban et al. concluded that a punishing stimulus may impact the ability of another discriminative stimulus to change behaviour. Other studies have found a punishing stimulus can decrease errors if the punisher signals a lower overall rate of reinforcers for errors. For example, Fowler, Hochhauser and Wischner (1981) delivered rats with higher, and lower, rates of food reinforcers for correct and incorrect responses, respectively, and shock or no shock for

incorrect responses. When the food differential between correct and incorrect responses was high, fewer errors occurred compared with when the food differential was low. However, punishing errors with shock enhanced the number of correct responses when the food differential was low, compared to no punishment. This finding suggests a punishing stimulus may improve the ability of organisms to discriminate between alternatives which are similar.

If behaviour produces a discriminative stimulus that signals an aversive consequence, and the consequence does not occur, then any subsequent reduction in the rate of behaviour is described as *conditioned punishment*. However, as noted by Doughty et al. (2007), studies that show conditioned punishment often use discriminative stimuli that continue to signal harmful consequences of behaviour. For example, Hake and Azrin (1965) found that a discriminative stimulus signalling non-response contingent shock reduced the rate of pigeons' key-pecking for food when the stimulus was made contingent on key-pecks.

Will a non-harmful discriminative stimulus function as a conditioned punisher? Auge (1977) investigated this question by using three different coloured keylights as discriminative stimuli signaling three equal consecutive periods of a fixed interval schedule of food reinforcement. Auge found the rate of pigeons' key-pecking maintained by food was greatest in the presence of the discriminative stimulus correlated with the shortest time to food. However, when key pecks in the presence of the stimulus correlated with the shortest time to food briefly produced the stimulus correlated with the *longest* time to food, the rate of key-pecking decreased. Auge concluded that a discriminative stimulus signaling a relative absence of positive reinforcement could punish pigeons' responding for food.

In a related study, Thompson (1965) showed that a stimulus signaling a relatively higher response requirement could punish responding maintained by food in pigeons. Two components of a multiple schedule were signaled by different discriminative stimuli and arranged the same overall rate of food presentation. However, key-pecks were reinforced according to a variable-interval (VI) schedule in one component, while reinforcement occurred after a fixed rate (FR) of responses in the other component. The FR value in the other component was manipulated from FR 1 to FR 300 across conditions. During punishment tests, responses in the VI component briefly produced either the discriminative stimulus for the current FR schedule, or a novel stimulus not associated with either schedule as a control. Thompson found that presenting the stimulus for the FR component decreased response rate in the VI component proportionate to the current FR schedule. In contrast,

response rate did not decrease when responses in the VI component produced the novel stimulus.

Orme-Johnson and Yarczower (1974) also made an initially neutral stimulus (a red light) response-contingent. During baseline, partial reinforcement occurred in the presence of a white light. Every five minutes on mean, the white light changed to red for one minute. When the light was red, one group of pigeons received response-contingent shock and either reinforcers or no reinforcers, depending on the condition, while another group received non-contingent shock and reinforcers. The red light was then made briefly contingent on responses in the presence of the white light for both groups during 1-min probe tests. Key-pecks could not produce shock or food during the probe test. A decrease in the rate of key-pecks for food was observed for the group of birds that had experienced non-contingent shock, but not for the birds that had experienced response-contingent shock. Orme-Johnson and Yarczower suggested that the red-light stimulus did not develop aversive properties for response-contingent shock groups, compared with the non-response contingent shock group.

Although the findings of Orme-Johnson and Yarczower (1974) are not generally supported in the punishment literature (e.g. Weisman, 1975), this finding suggests that a negative discriminative stimulus signalling no food will not punish responding for food. However, in the procedure used by Orme-Johnson and Yarczower, pigeons in the response-contingent shock group learned that responses in the presence of a red light could produce food, or shock, or both, at any time. The non-contingent shock group learned that responses could produce food in the presence of any light, but only the red light was correlated with unavoidable shock. It makes sense then, that when the red light was made response-contingent response rate decreased for the non-contingent group, but not the response-contingent group.

The studies by Auge (1977) and Thompson (1965) suggest that a discriminative stimulus that signals a relative absence of positive reinforcers can punish responding maintained by positive reinforcers. Therefore, we might expect that a negative discriminative stimulus signaling a complete absence of reinforcers (i.e. S-) would decrease behaviour to a greater degree than a stimulus that, while unfavourable, is nevertheless correlated with the availability of reinforcers during training (e.g., Thompson, 1965). Supporting this suggestion are studies that have investigated the aversive properties of S-. For example, Honig, Boneau, Bernstein & Pennypacker (1963) trained pigeons to discriminate between a black vertical line on white key that signaled the availability of

reinforcers (S+) and a white keylight which signaled the absence of reinforcers (S-). During a stimulus generalisation test in which the angle of the vertical line was varied, an excitatory gradient formed around S+ and an inhibitory gradient formed around S. This suggested the stimulus that signaled an absence of reinforcers had developed inhibitory properties (also see Farthing & Hearst, 1968).

Terrace (1971) trained pigeons to discriminate between two different stimuli signaling the availability- (S+), and unavailability- (S-), of positive reinforcers on a multiple schedule. Terrace then provided the pigeons with an 'escape' alternative, for which key-pecks during S- removed S- for 5 seconds but did not produce S+. Pigeons that experienced errors during discrimination training responded on the escape key to avoid S- presentations. Although S- was not used as a punisher in this study, these findings suggest S- may be an aversive stimulus (also see Terrace, 1972).

Although these studies suggest that S- develops aversive properties, relatively few studies have investigated whether S- will punish behaviour, once made contingent on behaviour. An exception is Mulvaney, Dinsmoor, Jwaideh & Hughes (1974), who found that a contingent S- stimulus will reduce the rate of observing responses maintained by *conditioned reinforcers* (also see Gaynor & Shull, 2002). However, these studies do not show whether a contingent S- stimulus will reduce the rate of behavior maintained by *positive* reinforcers. That is, we do not know whether a S- stimulus that is contingent on the same response that produces positive reinforcers will function as a punisher. Research in which a S- stimulus is contingent on observing responses is also complicated by the well-documented nonlinear relation between the presentation rate of the contingent S- stimulus and the rate of observing responses (see Shahan & Cunningham, 2015). Given that many behavioural treatments attempt to reduce problem behavior that continues to be maintained by positive reinforcers, there is benefit in considering whether a non-harmful stimulus will function as a punishing consequence for responding maintained by positive reinforcers. Finally, the study by Mulvaney et al uses a contingent S- stimulus that continues to signal an aversive consequence (i.e. extinction) that is still in effect. This leaves the question of whether a contingent S- stimulus that previously signaled an aversive consequence will punish behaviour despite that consequence no longer being in effect.

### *Conclusion*

The most useful methods of punishment are those able to decrease behaviour, or shift choice, across a range of species, irrespective of individual differences. The studies

that comprise this thesis use both pigeons and humans. This provides an opportunity to assess the efficacy of a contingent S- stimulus as a punisher in different species. Studies discussed in this literature review suggest that the type of punishing stimulus used might be less important to punishment efficacy than other environmental variables. Therefore, it is possible that a non-harmful contingent S- stimulus will successfully decrease responding maintained by positive reinforcers, and shift choice.

## CHAPTER 3.

### THESIS OVERVIEW

If effective and non-harmful punishers can be identified, these approaches could be used to effectively decrease problem behaviour, either alone or in combination with other methods (Hineline & Rosales-Ruiz, 2013; Horner et al., 2005; Lerman & Vordran, 2002). In addition, the use of non-harmful punishing stimuli would overcome the moral, ethical and procedural issues that have historically plagued punishment research and application.

As discussed in Chapter 2, research suggests that S-, a stimulus that signals the unavailability of reward, may develop aversive properties (Honig et al., 1963; Terrace, 1971). When this stimulus is made contingent on responding for conditioned reinforcers, the rate of responses may reduce compared to when the stimulus is not produced by responding (Mulvaney et al., 1974). However, there is no known research in which S- is used to punish behaviour maintained by access to primary reinforcers. In addition, in the study by Mulvaney et al., S- continues to signal an aversive outcome that remains in effect. Therefore, production of S- may elicit conditioned emotional responses, and be harmful. The studies that comprise this thesis investigate whether a contingent S- stimulus will punish behaviour maintained by access to positive reinforcers, despite the fact that the consequence signaled by S- is no longer in effect. That is, does S- alone punish behaviour maintained by positive reinforcers without recourse to other consequences? The studies that comprise this thesis also consider the potential for a contingent S- stimulus to shift choice, and suboptimal choice.

Towards our goal of developing, and testing, a contingent S- stimulus as a punisher for behaviour maintained by positive reinforcers, four quantitative basic research studies were conducted. In the first, we use pigeons to investigate whether rate of key-pecking for food in the presence of a green keylight (S+) will decrease if key-pecking for food in the presence of S+ produces a brief contingent S- stimulus signaling absence of food (a red keylight). In addition, we investigate whether choice between two equally reinforced alternatives will shift if responses to one alternative produce a contingent S- stimulus.

In the second study, we use human participants to replicate and extend the findings of the first study with pigeons. Participants were rewarded with money for pressing a key in the presence of S+ (a green computer screen), and did not receive any reward for key-pressing in the presence of S- (a red computer screen). After training, a punishment test

ensured that S+ responses briefly produced S-. We then compare the rate of S+ responding between baseline and punishment test sessions.

The remaining two studies investigate whether a contingent S- stimulus will punish behaviour described as ‘suboptimal’. As discussed in Chapter 2, suboptimal choice describes the process by which a greater long-term reward is forfeited for a lesser reward that seems better, or larger, in the moment (e.g., Spetch, Belke, Barnett, Dunn, & Pierce, 1990; Zentall, 2014, 2016a, 2016b). Suboptimal choice has been linked to the development of maladaptive behaviour in humans. This includes behaviour described as ‘addictive’ such as illicit drug-taking (Heyman & Dunn, 2002) and gambling resulting in loss (Heyman, 1996; Molet et al., 2012).

To investigate whether a contingent S- stimulus will punish suboptimal choice, we first needed to establish a reliable suboptimal choice procedure. We chose to replicate and extend a procedure used by Vaughan (1981). Vaughan developed a suboptimal choice procedure designed to test *melioration theory*. Briefly, melioration theory presumes that, at any given moment, organisms are attempting to equalise local rates of reward between competing choice alternatives (Herrnstein & Vaughan, 1980). The term ‘local rate’ describes the number of rewards obtained for behaviour divided by time spent on behaviour. One way to visualise this is to imagine you are standing between two showers that are raining coins. You only have five minutes until both coin showers stop, and you can only catch coins by standing in one shower or the other. However, the rate at which the coins fall in each shower keeps changing – in any given minute, one shower is faster than the other, and vice-versa. Melioration theory predicts that you will keep swapping between showers depending on which shower you perceive has the highest rate of coins falling *in that moment*. If the showers are very close together, then constantly swapping to the shower with the fastest rate of coin-fall will likely produce the greatest overall reward. But if the showers are 10 metres apart, then constantly swapping between them will result in the forfeiture of the greatest overall amount of reward. That is, it will lead to overall suboptimal choice. If you think of each 10-metre dash between showers as representing a 10% loss in quality of life, relationships, or health, it is possible to see how trying to equalise differences in local reinforcer rates can lead to natural world losses, and negatively impact life outcomes.

In the third study of this thesis, we replicate and extend the study by Vaughan (1981) to test the reliability of Vaughan’s procedure and his suboptimal choice findings. Then, in our fourth and final study, we use the same procedure to investigate whether suboptimal choice occurring through melioration can be punished by a contingent S- stimulus.

All four studies, and the key concepts that underlie this thesis, are extensively informed by a wide range of punishment literature published between 1913 and 2017. A review of this literature was provided in Chapter 2. The remaining chapters of the thesis present results and a discussion of each of the studies described above (Chapters 4 to 8). This is followed by a general discussion of all four studies collectively, including future research directions (Chapter 9). We conclude with the applied implications of our findings (Chapter 10).

## CHAPTER 4.

### DOES A NEGATIVE DISCRIMINATIVE PUNISH RESPONDING FOR FOOD IN PIGEONS?

The study presented in this chapter used pigeons to investigate whether the rate of key pecking maintained by food in the presence of a positive discriminative stimulus (S+) will decrease if key-pecking also produces a brief negative discriminative stimulus signaling absence of food (contingent S- stimulus). Pigeons were trained on a two-component multiple schedule. In one component, responses to a green keylight (S+) were reinforced on a VI 15-s schedule. In the other component, responses to a red keylight (S-) were never reinforced. Once the pigeons learned to discriminate between S+ and S-, two punishment test sessions were interspersed between repeated cycles of baseline. During punishment tests, responses made during S+ resulted in S- being superimposed on S+ for 1.5-s duration, according to a variable-ratio (VR) 5 schedule. To ensure the contingent S- stimulus was the only additional consequence of behaviour during punishment tests, the VI schedule of food was also maintained in the presence of both S+ and S- presentations.

The study arranged five conditions. The first condition investigated the potential for a contingent S- stimulus to punish S+ responding maintained by food when S+ and S- were trained on the same key. The second condition investigated whether training S+ and S- on different keys would impact punishment efficacy. Conditions 1 and 2 used a two-component multiple schedule in training but only a single component during punishment tests. In contrast, Condition 3 arranged a two-component multiple schedule in both training and punishment tests to control for procedural changes. Condition 4 trained S+ and S- as in Conditions 2 and 3, but arranged a novel contingent stimulus during punishment tests (see also Thompson, 1965) to control for potential effects of response-contingent presentation of any stimulus (cf. Podlesnik, Jimenez-Gomez, Ward, & Shahan, 2009). Across the first four conditions, if the rate of behaviour during S+ alone was greater than the rate of behaviour when S+ responses also produced the contingent S- stimulus (with little accompanying change in food delivery rate), then the contingent S- stimulus punished responding.

Finally, we investigated the impact of the contingent S- stimulus on choice in Condition 5. That is, we investigated whether choice would shift from a punished alternative to an unpunished alternative, despite maintaining a 1:1 reinforcer ratio between alternatives.

## METHOD

### Subjects

The subjects were five experimentally naïve pigeons, numbered 131 to 135. Pigeons were housed in individual cages within a colony room containing other pigeons participating in other experiments. Within their individual cages, pigeons had continuous access to water and grit. Access to food was restricted by supplementary post-feeding of mixed grain to maintain pigeons at  $85\% \pm 15$  g of their free-feeding body weight. Lighting was programmed to a 16-h light/8-h dark cycle with the light cycle beginning at 0:00 each day.

### Apparatus

The pigeons' home cages were 375 mm high, 380 mm wide, and 380 mm deep. The side and back walls of the cages were constructed of sheet metal, while the front wall, ceiling and floor were constructed of metal bars. The pigeons' cages contained two wooden perches positioned 60 mm above the floor. One of the perches was positioned parallel to, and 90 mm from, the front wall; the other perch was placed parallel to, and 90 mm from, the right wall. The pigeons' home cages also served as the experimental chambers. All cages were equipped with a response panel, which was placed on the right wall. The response panel contained four translucent response keys that were 18 mm in diameter and centered 260 mm above the wooden perches. Response keys were 65 mm apart, center to center, from neighboring keys. From left to right, the keys were designated Keys 1 to 4, and only Keys 1 and 3 were used in the current experiment. Keys 1 and 3 could be trans-illuminated by LEDs located behind each key with green, red or blue lights.

Key pecks to a lit response key that exceeded a force of 0.1 N closed a microswitch behind the keys, produced a feedback click, and were registered as a response for analyses. When reinforcers were presented for a registered response, a hopper containing wheat was made available for 3 s. The food hopper was located behind a magazine aperture, located 80 mm below the keys, measuring 55 mm in height, 55 mm in width, and 40 mm in depth. As the food hopper was raised to allow pigeons access to food, a light behind the magazine aperture was illuminated and all keylights were extinguished. All experimental events were arranged and recorded by a Windows computer running MED-PC IV® software, which was in a room next to the pigeon-colony room.

Table 4.1

*Discriminative Stimuli for Components 1 and 2 and Allocation of VI schedules of Food Deliveries on Left and Right keys across all Conditions for Baseline (top panel) and Punishment Test (bottom panel) sessions. Total Number of Sessions, and Blocks of Sessions, within each Condition is shown in the far-right column.*

	CONDITION	COMPONENT 1 STIMULI		COMPONENT 2 STIMULI		SESSIONS
		Left green key	Right green key	Left red key	Right red key	
BASELINE	1	VI 15 s	-	EXT	-	3 x 10
	2	VI 15 s	-	-	EXT	3 x 10
	3	VI 15 s	-	-	EXT	3 x 10
	4	VI 15 s	-	-	EXT	3 x 10
	5	VI 20 s	VI 20 s	EXT	EXT	2 x 10
PUNISHMENT TEST	1	VI 15 s	1.5s contingent stimulus Red (left)	-	-	3 x 2
	2	VI 15 s	Red (left)	-	-	3 x 2
	3	VI 15 s	Red (left)	-	EXT	3 x 2
	4	VI 15 s	Blue (left)	-	EXT	3 x 2
	5	VI 20 s VI 20 s	Red (left) Red (right)	VI 20 s VI 20 s	EXT EXT	EXT EXT

sessions were conducted sequentially across pigeons, so that each pigeon had to complete one full session before the next pigeon began a session. The first session started at 1:00 AM each day, and the last session ended at approximately 6:00AM each day. Laboratory staff were not present during experimental sessions.

There were five conditions in total. Each condition consisted of two alternating phases: baseline training followed by a punishment test. Table 4.1 shows the component stimuli and food delivery schedules arranged for responding on the keys across all conditions. A block of two punishment test sessions immediately followed each of three blocks of 10 baseline sessions for Conditions 1 to 4. In Condition 5, a block of two punishment-test sessions followed each of two blocks of ten baseline sessions. Thus, there were six punishment test sessions in Conditions 1 to 4, and four punishment test sessions in Condition 5.

**Baseline.** Each baseline session was 25 min. A two-component multiple schedule was arranged throughout all baseline sessions in Conditions 1 to 4. Condition 5 arranged a two-

component multiple concurrent schedule. All components were 60 s. Each component was followed by a 10-s inter-component interval (ICI) in which all keylights were off and food presentations never occurred. Figures 4.1 and 4.2 show how each component was signaled by discriminative stimuli during baseline and punishment tests across all conditions.

In Conditions 1 to 4, the left green keylight served as the antecedent discriminative stimulus in Component 1, and is referred to as S+. Hereafter, Component 1 is referred to as the *S+ component*. During the S+ component, food reinforcers were delivered according to a VI 15-s schedule that arranged an exponential distribution of reinforcers by interrogating a probability generator every 1 s. The red keylight served as the antecedent discriminative stimulus in Component 2, in which no food deliveries were arranged. Therefore, the red keylight is referred to as S-, and hereafter Component 2 is referred to as the *S- component*. At the beginning of each session, the S+ component always went first, after which the two components alternated. In the S- component, the red keylight appeared on the same key location as for the S+ component in Condition 1 (i.e. the left key), but appeared on the right key during Conditions 2 to 4. The location of S- was switched to the right key for Conditions 2 to 4 to control for the possibility that training S+ and S- at the same location might decrease the extent to which S- will function as a punishing consequence during tests.

Condition 5 investigated the impact of a contingent S- stimulus on choice because choice measures tend to be more sensitive to differential consequences than multiple schedules (see Davison & McCarthy, 1988). We arranged green S+ stimuli on the left and right keys in the S+ component. Reinforcers for left and right key pecks were dependently scheduled on a VI 20-s schedule. We arranged red S- stimuli on the left and right red keys in the S- component. A changeover delay (Herrnstein, 1961) was not arranged following switches between keys.

**Punishment Tests.** The punishment tests differed from baseline only in that responses in the S+ component briefly produced the S- discriminative stimulus (i.e. the red keylight). The bottom panel of Figures 4.1 and 4.2 shows how each component was signaled by different discriminative stimuli during punishment tests for Conditions 1 to 5. Across all conditions, S+ stimuli presented during the S+ component in punishment tests were the same as for the preceding block of baseline sessions (see Table 4.1). However, a variable number of

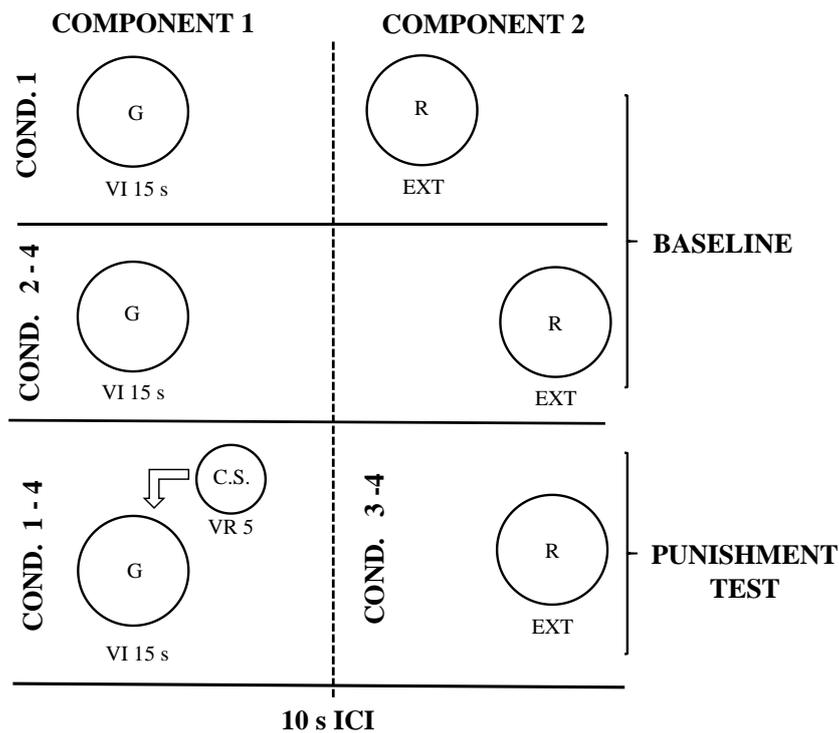


Figure 4.1 Left green key (G) and right red key (R) stimulus arrangements, and food delivery schedule value, for Components 1 and 2 in Baseline (top panel) and Punishment Test (bottom panel) for Conditions 1 to 5. The colour of the 1.5-s contingent S- stimulus (C.S.) presented in Punishment Tests was red for Conditions 1, 2, 3, and 5; and blue for Condition 4 (see Table 4.1).

responses to S+ according to a VR-5 schedule resulted in the replacement of S+ with S- (Conditions 1 to 3), or a novel stimulus (Condition 4), for 1.5 s. Hereafter, replacement of S+ with S- for 1.5 s is referred to as the *contingent S- stimulus*. Replacement of S+ with a novel stimulus for 1.5 s is referred to as the *contingent novel stimulus*. The contingent novel stimulus was a blue keylight not associated with prior training, and hence Condition 4 was a control condition to examine whether contingently presenting any stimulus could affect response rates in S+. In Condition 5, the contingent S- stimulus was presented during the S+ component for 1.5 s only on the left key for the first block of two punishment tests, and only on the right key for the second block of two punishment tests.

During Conditions 1 and 2, the arrangement of components during punishment tests differed from the arrangement of components in baseline. Specifically, S- component was not presented during punishment tests. That is, punishment tests featured only repeated 60-s cycles of the S+ component, with each component followed by a 10-s ICI. As a result,

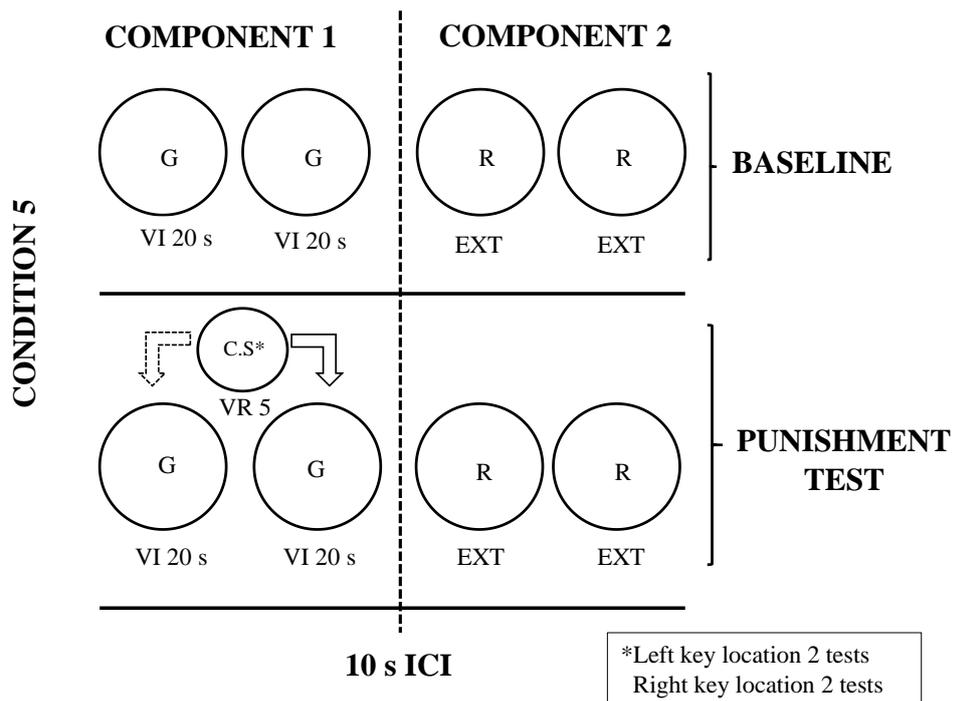


Figure 4.2 Left and right key stimulus arrangements for Components 1 and 2 in Baseline (top panel) and Punishment Test (bottom panel) sessions for Condition 5. The 1.5-s contingent S- stimulus (C.S.) presented during Punishment Tests was a red keylight. The C.S. was presented on the left key for the first block of two Punishment Tests, and on the right key for the second block of two punishment tests (see Table 4.1).

each punishment-test session in Conditions 1 and 2 was 12.5 min duration. As responses made during the S+ component in the present study were reinforced, and responses made during the S- component were not reinforced, the removal of the S- component during punishment tests could have impacted the response rate in the S+ component. To control for this difference, the S- component was added to each punishment test session for Conditions 3 to 5. In Conditions 3 to 5, as the S- component always followed the S+ component, punishment test sessions were 25 min.

During punishment tests, the VI schedule of food for responses in the S+ component was maintained even during S- presentations. Therefore, when responses to S+ produced the contingent S- stimulus for 1.5-s, it was possible for a food delivery to occur. This aspect of the procedure ensured that any punishing effects during the S+ component would be from contingent S- presentations rather than also introducing periods of extinction to the S+ component. This is important as any reduction in S+ response rate can then be attributed to

the contingent S- stimulus, and not to contingent extinction presentations or relatively lower rates of food presented during punishment tests.

**Dependent measures.** In Conditions 1 to 4, for each block of 10 baseline sessions, mean responses and food deliveries per min to S+ stimuli, and mean responses per min to S- stimuli, were calculated for individual pigeons across the last five sessions. During punishment tests, responses per min in the S+ component was calculated by dividing responses that occurred only in the presence of S+ by total min in which only S+ was present. Therefore, responses and time that accumulated during 1.5-s presentations of the contingent S- stimulus, and time accumulated during 3-s food deliveries, were deducted from the total min used in the calculation. Although food rates were held constant between baseline and punishment test sessions, during punishment tests the S+ food rate was calculated using only food obtained in the presence of S+. Food obtained in the presence of the contingent S- stimulus was deducted from the calculation. This allowed us to compare response and food rates in the presence of S+ alone. We used these values to obtain a log proportion of baseline S+ response and food delivery rates, between the preceding block of baseline and each of six punishment tests, for individual pigeons, together with the group mean.

Condition 5 used a two-key concurrent VI VI schedule. Responses to the left key produced the contingent S- stimulus during the first block of two punishment tests, and responses to the right key produced the contingent S- stimulus in the second block of punishment tests. We calculated the mean proportion of left- or right-key S+ choice across the last five sessions of each block of 10 baseline sessions, and in each punishment test, for individual pigeons. This provided a ratio of choice proportions between baseline and punishment tests, which we used to calculate a log proportion of baseline S+ choice. The same measures were used to obtain a log proportion of baseline S+ food rate. If log proportion of baseline S+ choice was negative, choice during punishment tests had shifted away from the punished alternative and towards the unpunished alternative. If log proportion of S+ choice was positive, choice had shifted towards the punished alternative, and away from the unpunished alternative. If log proportion of S+ food rate was positive, or negative, during punishment tests the punished alternative received a relatively higher-, or lower, rate of food, respectively, compared with the same alternative in baseline.

## RESULTS

Table 4.2 shows obtained food deliveries and responses per min in the S+ component, and S- component, across the last five sessions of all blocks of baseline sessions for Conditions 1 to 5, across all pigeons. Group mean S+ response rate was relatively greater during baseline in Conditions 1 and 2 than in Conditions 3 to 5 because Pigeon 135's response rate decreased considerably during Conditions 3 to 5. Food delivery rates across Conditions 1 to 4 closely followed the arranged baseline reinforcer contingencies for S+ component ( $M = 4.00$ ). Left and right key food delivery rates for Condition 5 closely followed the arranged 1:1 ratio for concurrent choice in S+ component. In all conditions, rates of responding in the S+ component were substantially higher than in the S- component, as demonstrated by all discrimination indexes being .99 or greater. This demonstrates strong discrimination between the two components and that S- was established as a signal for the absence of reinforcers.

**Condition 1.** During baseline, S+ and S- featured in alternating components, but were trained on the same key. During punishment tests, the S- component was absent and every five responses on mean made during S+ component resulted in the contingent S- stimulus replacing S+ for 1.5 s. Figure 4.3 shows log proportion of baseline S+ response and food rates in each test within three blocks of two punishment tests, for individual pigeons, and as a group mean. Log proportion of baseline S+ response rate was negative in all six punishment tests for Pigeon 132, for five of six tests for Pigeon 133, for four of six tests for Pigeon 135, for three of six tests for Pigeon 134, and in two of six tests for Pigeon 131. Therefore, there were 20 of 30 tests in which log proportion of baseline S+ response rate was negative. Of these, log proportion of baseline S+ food rate was greater than log proportion of baseline S+ response rate in 16 of 20 tests.

Overall, these results suggest the contingent S- stimulus tended to punish S+ response rate in most tests for Pigeons 132, 133 and 135 but not for Pigeons 131 and 134. It is possible that training S+ and S- at the same location might have decreased the extent to which the contingent S- stimulus functioned as a punishing consequence for these two pigeons. Condition 2 controlled for this possibility by presenting S+ and S- at different locations in training.

Table 4.2

*Overview of Group Mean response rates and Food Delivery rates, in S+ components, with SEM; and Group Mean response rates in S- components, for Conditions 1 to 4. Condition 5 shows Group Mean response and Food Delivery rates for Left and Right keys in S+ components, with SEM, and response rates for Left and Right keys in S- components. The Discrimination Index shows Group Mean Proportion of S+ to S- responses.*

CONDITION	Group mean baseline response rates per min			Discrim. Index S+	Group mean baseline food S+ Comp.		
	S+ Component	SEM (S+)	S- Component		Schedule	Food/min	SEM
1	114.80	2.90	0.87	0.99	VI 15 s	3.92	0.07
2	110.20	5.67	0.65	0.99	VI 15 s	3.88	0.07
3	98.92	7.41	0.85	0.99	VI 15 s	3.89	0.08
4	99.87	7.61	0.90	0.99	VI 15 s	3.92	0.08
5	46.35/50.72	1.85/1.70	0.37/0.20	0.99/1.00	VI 20s/VI 20s	2.80/2.60	0.07/0.07

**Condition 2.** In Condition 2, S+ and S- were trained on different keys (see Table 4.1). Figure 4.4 shows log proportion of baseline S+ response and food rates in each test within three blocks of two punishment tests, for individual pigeons, and as a group mean. Log proportion of baseline S+ response rate was negative in four of six punishment tests for Pigeons 132, 133, and 135, and for one test for Pigeons 131 and 134. Therefore, there were 14 of 30 tests in which log proportion of baseline S+ response rate was negative. Of these, log proportion of baseline S+ food rate was greater than log proportion of baseline S+ response rate in 13 of 14 tests.

Overall, these results suggest the contingent S- stimulus tended to punish S+ response rate in most tests for Pigeons 132, 133 and 135. As observed in Condition 1, the contingent S- stimulus tended not to punish S+ response rate for Pigeons 131 and 134.

**Condition 3.** In Condition 3, the S- component was added to punishment test sessions to ensure training and punishment test procedures were as similar as possible with the exception of contingent S- presentations in the S+ component. Figure 4.5 shows log proportion of baseline S+ response and food rates in each test within three blocks of two punishment tests, for individual pigeons, and as a group mean. Log proportion of baseline S+ response rate was negative for all six punishment tests for Pigeon 132, for five of six tests for Pigeon 135, for four of six tests for Pigeons 131 and Pigeon 134, and for one test for Pigeon 133. Therefore, there were 20 of 30 tests in which log proportion of baseline S+ response rate was negative. Of these, log proportion of baseline S+ food rate was greater

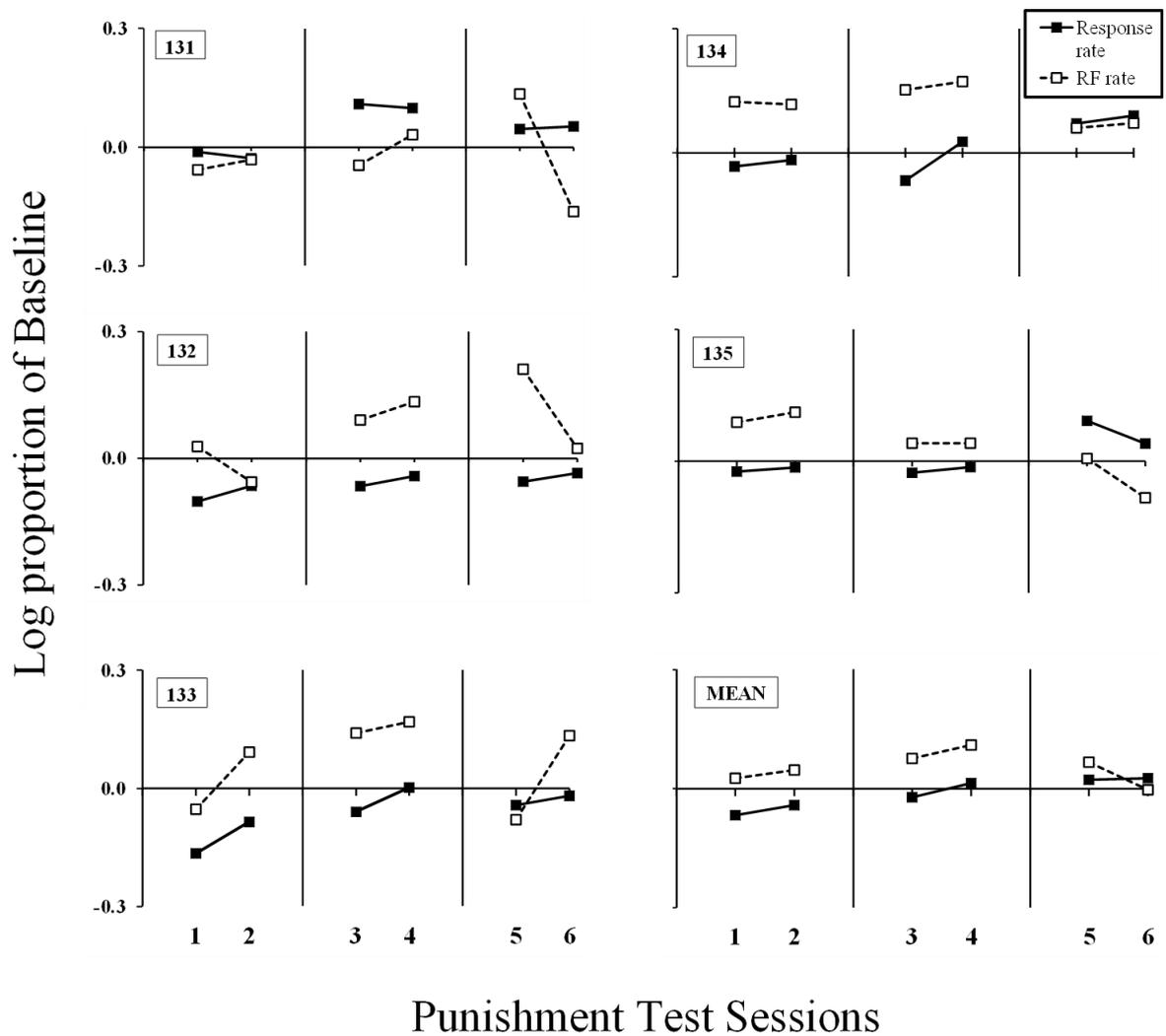
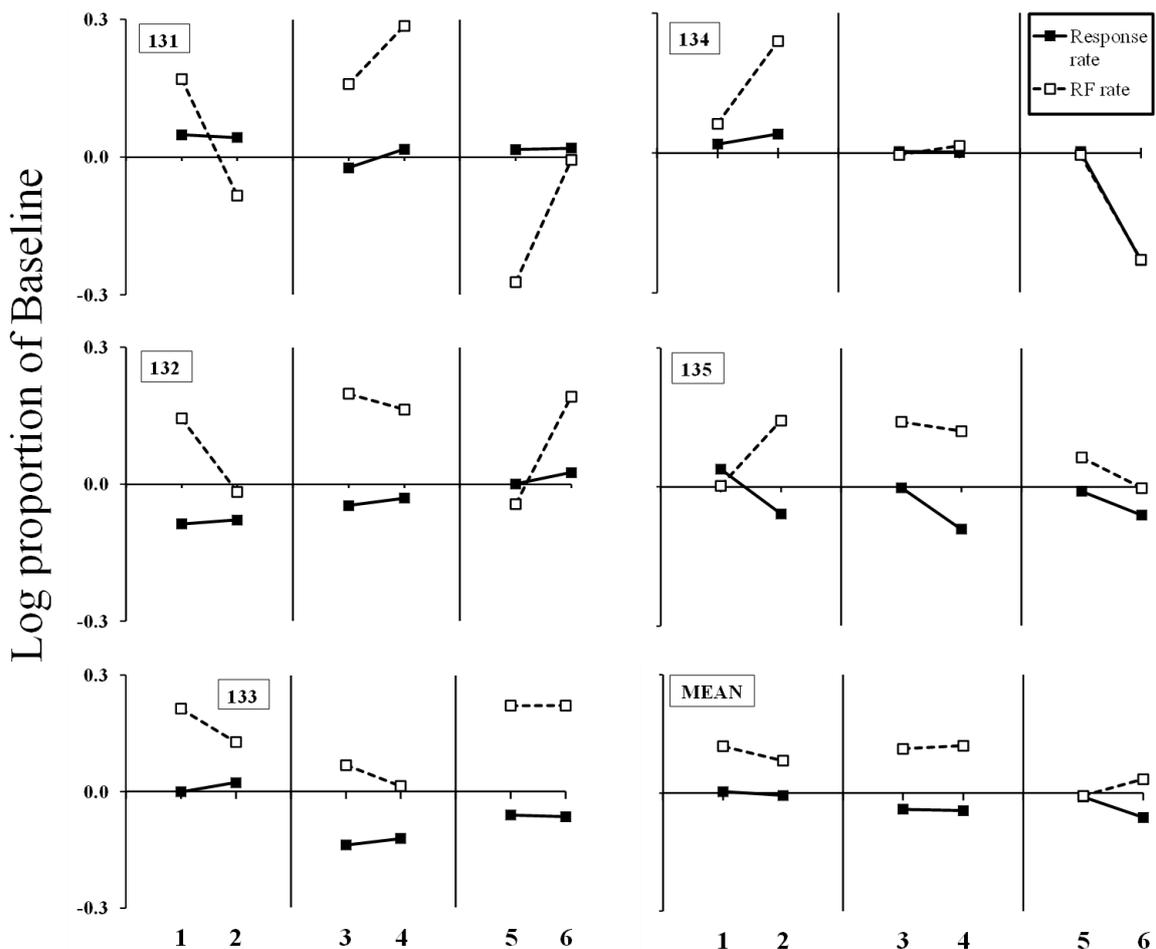


Figure 4.3 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min measured as log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines), for individual pigeons, and group mean, in Condition 1.

than log proportion of baseline S+ response rate in 12 of 20 tests. Overall, these results suggest the contingent S- stimulus tended to punish S+ response rate in most tests for Pigeons 131, 132, 134 and 135, but not for Pigeon 133.

**Condition 4.** In Condition 4, baseline S+ and S- stimuli were the same as for Condition 3. In addition, every five responses on mean made in S+ component during punishment tests



### Punishment Test Sessions

Figure 4.4 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min as a log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines) for individual pigeons, and group mean, in Condition 2.

resulted in a contingent novel stimulus replacing S+ for 1.5 s. Therefore, Condition 4 was a control condition. Figure 4.6 shows log proportion of baseline S+ response and food rates in each test within three blocks of two punishment tests, for individual pigeons, and as a group mean. Log proportion of baseline S+ response rate was negative in no punishment tests for Pigeons 131 and 133, for one of six tests for Pigeons 134 and 135, and for two of six tests for Pigeon 132. Therefore, there were 4 of 30 tests in which log proportion of

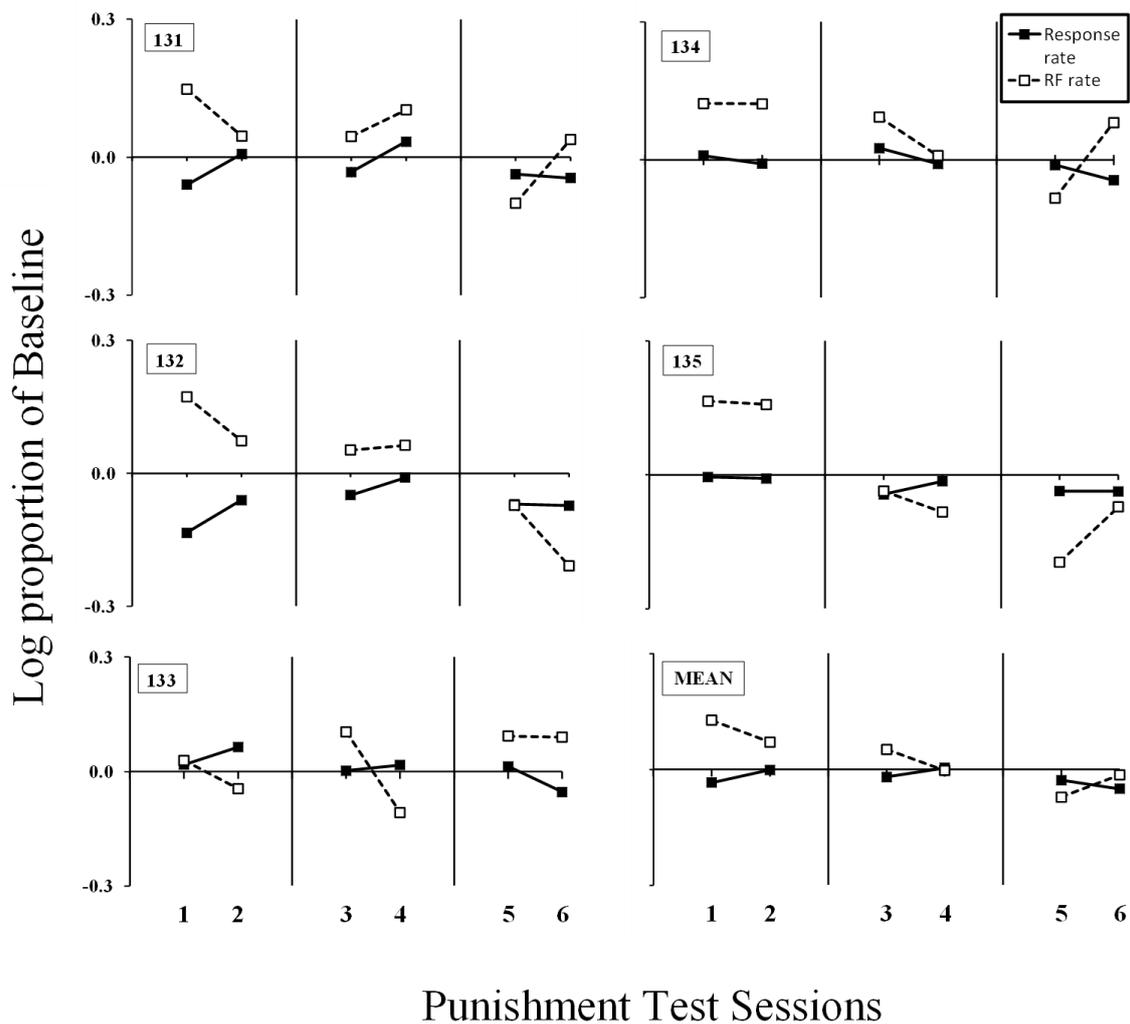


Figure 4.5 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min as a log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines), for individual pigeons, and group mean, in Condition 3.

baseline S+ response rate was negative. Of these, log proportion of baseline S+ food rate was greater than log proportion of baseline S+ response rate in all four tests. Overall, these results suggest the contingent novel stimulus tended not to punish S+ response rate in most tests for all pigeons in the Control condition.

**Condition 5.** In Condition 5, baseline sessions featured concurrently presented left- and right-key S+ stimuli in the S+ component, and concurrently presented left- and right-key S-

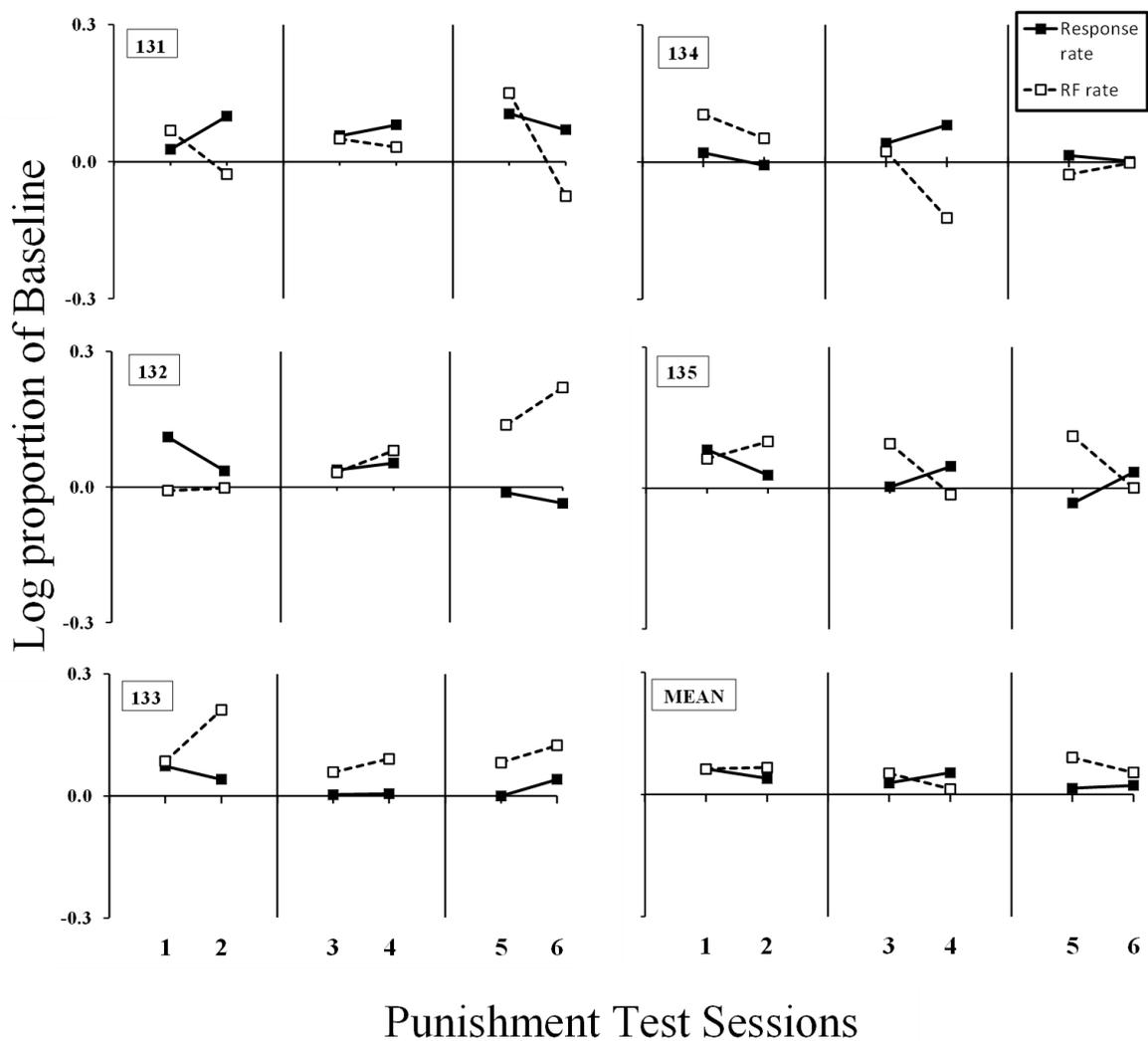
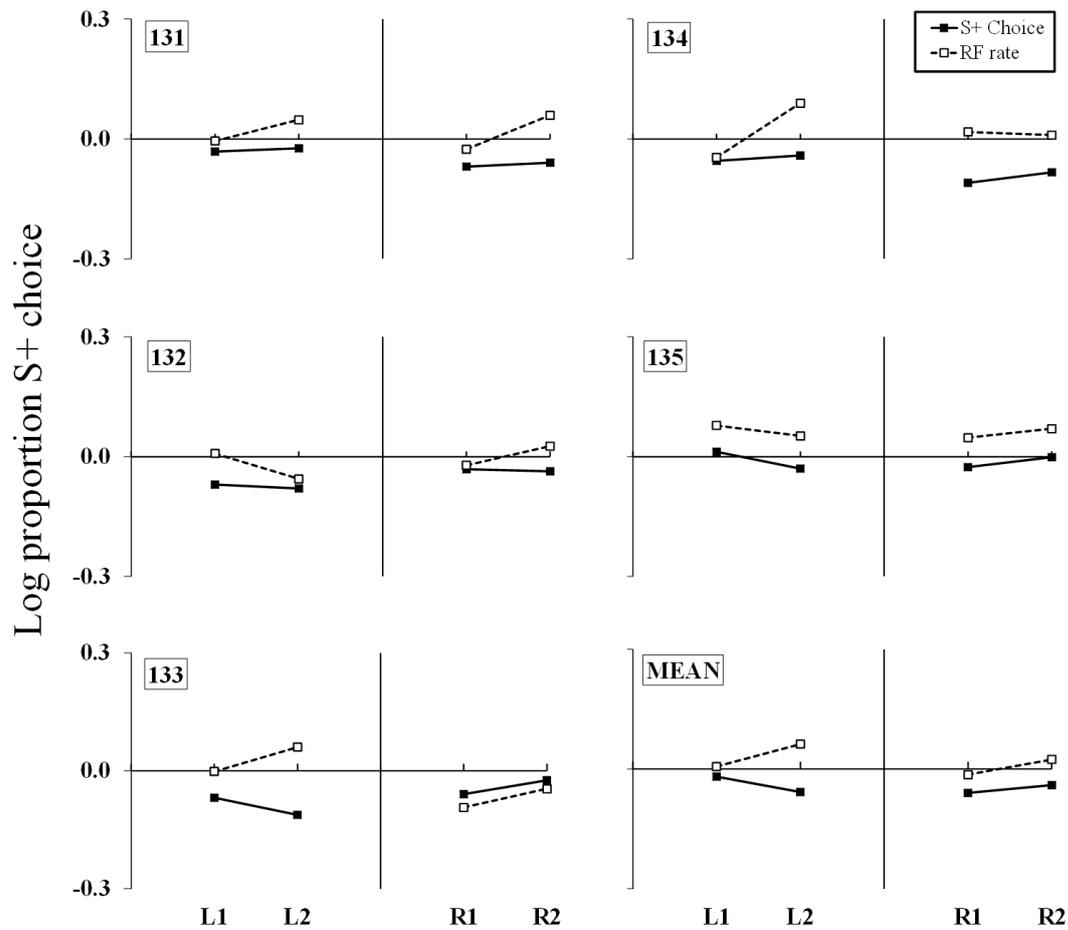


Figure 4.6 S+ response- (solid black lines, black squares) and food delivery rates (dashed black lines, white squares) per min measured as log proportion of mean baseline S+ response and food delivery rates (y-axis) for three blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines), for individual pigeons, and group mean, in Condition 4, the Control condition.

stimuli in the S- component (see Table 4.1). During baseline, group mean proportion of left-key S+ choice was  $M=0.49$  ( $SEM = 0.01$ ) and group mean proportion of left-key reinforcers was  $M=0.49$  ( $SEM = 0.01$ ). Figure 4.7 shows log proportion of baseline left- or right-key S+ choice in each punishment test, for individual pigeons and as a group mean. Log proportion of baseline S+ choice was negative in all four punishment tests for Pigeons 131 to 134, and for two of four punishment tests for Pigeon 135. That is, choice moved away from the punished alternative, and towards the unpunished alternative. Of these 18 cases in



### Punishment Test Sessions

Figure 4.7 Left-key (L1 and L2) and right-key (R1 and R2) S+ choice (solid black lines, back squares) and food delivery rates (dashed black lines, white squares), measured as a log proportion of mean baseline left- or right-key choice and food delivery rates, for two blocks of two punishment tests (PT) (x-axis) following each block of 10 baseline sessions (BL) (vertical black lines) for individual pigeons, and group mean, in Condition 5.

which choice became less extreme following the introduction of the punisher, log proportion of baseline S+ food rate was positive in 10 of 18 tests. Thus, despite the arranged 1:1 reinforcer ratio, the contingent S- stimulus tended to shift choice away from the punished alternative and toward the unpunished alternative on all tests for Pigeons 131 to 134, but not for Pigeon 135.

## DISCUSSION

We investigated the extent to which a response-contingent 1.5-s negative discriminative stimulus (contingent S- stimulus) which previously signaled an absence of food reinforcers would decrease the rate of key-pecking by pigeons in the presence of a different stimulus signaling availability of food (S+). In punishment tests for Conditions 1 to 3, when responses during S+ produced the contingent S- stimulus, the rate of S+ responding decreased for most pigeons on most tests compared to baseline. That is, S+ response rate decreased in at least four of six punishment tests, for at least three of five pigeons. This occurred despite the fact responses during both S+ and the contingent S- stimulus could produce food. Reductions in S+ response rate also occurred when the S+ food rate during punishment tests had increased relative to baseline. In Condition 5, the contingent S- stimulus shifted choice away from the punished alternative and toward the unpunished alternative despite a 1:1 food delivery ratio between those alternatives. Therefore, reduction in S+ response rate during punishment tests in Conditions 1 to 3 cannot be attributed to decreased food rates.

Condition 4 was a control condition that replaced the contingent S- stimulus with a contingent novel stimulus. The results of this condition show that the effects of the contingent S- stimulus in Conditions 1 to 3 cannot be attributed to the mere presence of a contingent stimulus. Log proportion of baseline S+ response rate was positive in most tests across pigeons. This suggests the contingent novel stimulus tended not to punish S+ response rate. Instead, these findings suggest response-contingent presentations of stimuli not correlated with food reinforcement can increase response rates. Others have found auditory stimuli uncorrelated with food reinforcement with rats increases both response rate and resistance to disruption (see Reed & Doughty, 2005). Therefore, the findings of Condition 4 of the present study contributes to findings revealing stimuli uncorrelated with positive reinforcement can nevertheless influence operant behaviour. One interpretation of the results of the first three conditions in this study is that response rate during S+ was punished by the removal of S+ for 1.5-s periods, rather than by 1.5-s presentations of the contingent S- stimulus. However, the results of the Control condition do not support this suggestion. Specifically, when responses produced the neutral contingent stimulus during punishment tests, S+ response rate was greater than S+ baseline response rate on 25 of 30 tests, across all pigeons. This suggests that reductions in S+ response rate observed during punishment tests across Conditions 1 to 3 did not occur due to removal of the S+ stimulus.

Many applied studies have shown harmful stimuli that produce conditioned emotional responses effectively punish operant behaviour in humans (Duker & Seys, 2000; Eckstein & Hart, 1996; Linscheid & Richenbach, 2002; Salvy et al., 2004) and animals (Dale, Statham, Podlesnik & Elliffe, 2013). However, behavioural practitioners prefer non-harmful methods of punishment that decrease problem behaviour (Lerman & Vorndran, 2002; Himeline & Rosales-Ruiz, 2013.) Auge (1977), Davidson (1970), and Thompson (1965) demonstrated that when reinforced responses produce a discriminative stimulus previously associated with an unfavorable outcome, response rate decreases relative to when those responses do not produce a contingent discriminative stimulus. The results of the present study suggest that a contingent S- stimulus associated with an absence of reinforcers could also decrease S+ response rate and shift choice away from a punished alternative and towards an unpunished alternative.

A similarity between the studies by Thompson (1965) and Davidson (1970) and the present study is that responses punished only by a contingent discriminative stimulus were still able to produce positive reinforcers. This will be of interest to applied researchers, given that methods of treatment to reduce harmful behaviour in humans contend with the continued availability of positive reinforcers for behaviour.

While Thompson and Davidson controlled the overall number of reinforcers obtained in baseline and punishment tests, they did not report whether differences in reinforcer *rates* occurred between baseline training and punishment test conditions. It is therefore possible that relatively lower reinforcer rates during punishment tests impacted punished response rate in their studies. In contrast, the present study shows that a contingent S- stimulus will punish responding maintained by positive reinforcers, even when the reinforcer rate during punishment tests is greater than the baseline reinforcer rate. An additional difference between the studies of Thompson and Davidson and the present study is that not every response made during punishment tests in the present study produced the contingent discriminative stimulus. As continuous punishment has been found to be more effective than intermittent punishment in decreasing responding maintained by positive reinforcers (Hendry & Van-Toler, 1964; Lerman, Iwata, Shore, & DeLeon, 1997; Linden, 1976; Zimmerman & Ferster, 1963), it is promising that a non-harmful contingent stimulus will punish responding for positive reinforcers, despite not being produced by every response. Nevertheless, future research should examine approaches to making contingent S- presentations more effective, including the schedule of S- presentations.

Our findings suggest it is possible that contingent S- presentations may be more effective punishers if S+ and S- are trained at the same location. For example, we changed the baseline training location of S- from the left key in Condition 1, to the right key in Condition 2, without making additional changes to procedure. Results show log proportion of baseline S+ response rate was negative in 20 of 30 punishment tests in Condition 1, compared with 14 of 30 tests in Condition 2. As S+ and S- were trained at the same location in Condition 1, these results suggest training S+ and S- in different locations might reduce the extent to which a contingent S- stimulus will punish behavior.

Future research should also examine the longevity with which a contingent S- stimulus will continue to punish responding maintained by positive reinforcement. In the present study, log proportion of baseline S+ response rate was measured across a block of two 25-min punishment test sessions. This was followed by a return to a block of 10 baseline training sessions. However, even when successful in the short term, punishment interventions do not always reduce behaviour for the long-term (Vorndran and Lerman, 2006). Therefore, future studies should examine how long a contingent S- stimulus will continue to punish responding maintained by positive reinforcement, without a return to training.

A limitation of the present study is the use of a blue keylight as a neutral contingent stimulus in Condition 4, the Control condition. The pigeons used in this study were experimentally naïve and had no prior learning history associated with keylight colours. However, the colours of the blue keylight used as a neutral contingent stimulus, and the green keylight used to signal the availability of food (S+) are located along the same dimension. Given this, the findings of the Control condition would have been strengthened if the neutral contingent stimulus had been from another dimension than that of S+. Alternatively, keylight colours could have been counterbalanced across the pigeons, so that for some pigeons a red keylight signaled access to food, and a green keylight signaled an absence of food; and for some pigeons, vice versa. However, this limitation falls short of a being a confound for the results of the first three conditions of this study. In the Control condition, S+ response rate during the punishment test was *greater* than baseline S+ response rate in 21 of 30 tests, and the same as baseline S+ response rate in 5 of 30 tests, across pigeons. As S+ response rate was relatively greater for 70% of punishment tests, it appears the contingent blue keylight signaled a value greater than that of the green keylight on those tests. Although the findings of Condition 4 are not without precedent (e.g. Reed & Doughty, 2005), it is also possible that, due to S+/S- discrimination training with green and

red keylights, the blue keylight was discriminated as a more salient version of the green keylight during punishment tests in Condition 4. If so, the findings of Condition 4 further support the suggestion that in Conditions 1 to 3, S+ response rate was reduced in most punishment tests for most pigeons, compared with baseline, because the contingent red keylight signaled a value that was less than that of the green keylight. That is, the results of the Control condition support the suggestion that a contingent discriminative stimulus associated with previous training will change behaviour.

A further limitation of the present study is that punishment by the contingent S- stimulus tended to occur for different pigeons in Condition 3, compared to Conditions 1 and 2. For example, when the baseline training procedure differed from the procedure used in punishment tests (i.e. Conditions 1 and 2), the contingent S- stimulus tended to punish S+ response rate in most tests for Pigeons 132, 133 and 135, but not Pigeons 131 and 134. When training and punishment test procedures were then aligned (i.e. Condition 3), the contingent S- stimulus tended not to punish S+ response rate for Pigeon 133, but punished S+ response rate in most tests for all the other pigeons. This suggests procedural changes between discrimination training and punishment tests could impact the punishing efficacy of a contingent S- stimulus, and that the direction of that impact might differ between individuals. Alternatively, punishment for different pigeons in different procedures might indicate general variability in the extent to which a contingent S- stimulus will punish responding maintained by positive reinforcers. Further research is required in which experimental conditions are repeated across cycles to analyze the extent to which punishment by a contingent S- stimulus is maintained for the same subjects. Ongoing research could also determine whether procedural changes between training and the application of a contingent S- stimulus during punishment tests either enhance or diminish the punishing effects of the stimulus, or have little effect. These investigations are important given applied research that shows punishing stimuli do not usually require previous discrimination training to be effective (e.g. Hanley, Piazza, Fisher & Maglieri, 2005).

Condition 5 investigated the effect of the contingent S- stimulus on S+ choice. Choice has been found to be more sensitive to consequences compared with simple schedule responding on multiple schedules (see Davison & McCarthy, 1988). For example, Catania (1963) found that unequal reinforcer magnitudes strongly impacted choice in pigeons, but had little impact on response rates on a single-VI schedule. The suggestion that choice is more sensitive to consequences than response rate on simple schedules is supported by the results of the present study. The contingent S- stimulus tended to punish S+ response rate

for three of five pigeons on most tests with multiple schedules arranged in Conditions 1 to 3. However, in Condition 5 the contingent S- stimulus shifted choice away from the punished alternative, and toward the unpunished alternative, on all tests for four of five pigeons.

Although decreases in behavior were small for some subjects in some conditions in this study, these findings invite further research. Future studies might investigate whether procedural changes could enhance the punishing effects of a negative discriminative stimulus. For example, we might investigate the impact of changes to punisher and reinforcer schedules in simple schedules, and in studies of choice. For example, if a contingent S- stimulus will shift choice without changes to reinforcer ratios, future studies could investigate the extent to which a contingent S- stimulus could prevent post-DRA treatment relapse when made contingent on harmful behaviour. A contingent S- stimulus could also alter choice processes thought to contribute to overall suboptimal outcomes. Suboptimal choice has been linked with the development of maladaptive behaviour in humans, including behaviour described as ‘addictive’ such as illicit drug-taking (Heyman & Dunn, 2002) and gambling resulting in loss (Heyman, 1996; Molet et al., 2012).

The extent to which any discriminative stimulus functions as a punisher will depend upon what that stimulus predicts about the nature and probability of future events (see Baum, 2012; Cowie & Davison, 2016; Killeen & Jacobs, 2017; Shahan, 2017). If we accept that behaviour changes as a function of events that are likely to follow, these changes should occur regardless of whether discriminative stimuli are appetitive (e.g., Cowie, Davison & Elliffe, 2011), aversive (Ayllon & Azrin, 1966; Galbicka & Platt, 1984; Kelleher & Morse, 1968) or have no phylogenetic importance (Boutros, Davison & Elliffe, 2011). In support of this research, the findings of the present study suggest that, once made response contingent, a discriminative stimulus previously correlated with an absence of reinforcers can decrease responding maintained by positive reinforcers and shift choice.

## CHAPTER 5.

### DOES A NEGATIVE DISCRIMINATIVE STIMULUS PUNISH RESPONDING FOR MONEY IN HUMANS?

The findings described in Chapter 4 contribute to the literature on punishment by showing that a contingent S- stimulus has the potential to punish operant behaviour maintained by access to positive reinforcers in pigeons. In addition, Study 1 showed that punishment by a contingent S- stimulus can occur despite S- no longer signaling an absence of reinforcers. Given the importance of reducing harmful behaviour in applied settings, it is now useful to investigate whether a contingent S- stimulus might also punish responding for positive reinforcers in humans.

Harmful human behaviour, including self-injurious behaviour, is problematic for behavioural practitioners (Doehring, Reichow, Palka, Phillips & Hagopian, 2014; Harvey, Boer, Meyer, & Evans, 2009; Symons & Thompson, 1997; Vismara, & Rogers, 2010). One approach to the reduction of harmful behaviour is delivery of a high rate of reinforcers for an alternative behaviour, while reducing the rate of reinforcers for harmful behaviour (i.e. DRA; see Petscher, Rey & Bailey, 2009 for a review). However, while harmful behaviour is often reduced during DRA treatment, it is not always reduced to safe levels (Taylor, Oliver, & Murphy, 2011). In addition, when DRA interventions end, behaviour can relapse at higher than pre-treatment levels (Mace et al., 2010; Podlesnik & Kelley, 2015; Volkert, Lerman, Call, & Trosclair-Lasserre, 2009).

If DRA and similar treatments sometimes fail to reduce harmful behaviour to safe levels, then punishment can be added to those treatments (Hanley, Piazza, Fisher & Maglieri, 2005; Hagopian, Fisher, Sullivan, Acquistio & LeBlanc, 1998). As defined in Chapter 1, punishment occurs when a stimulus reduces the subsequent probability of the behaviour it follows (Azrin & Holz, 1966; Yulevich & Axelrod, 1983). However, many types of punishing stimuli have negative side effects, including fear, aggression, and other conditioned emotional responses (Axelrod, 2013; Hamilton, 1982; Newsom, Favell, & Rincover, 1983).

In the previous chapter we showed when key-pecks to S+ briefly produced a S- stimulus, S+ response rate decreased for most pigeons on most tests compared with when responses did not produce the S- stimulus ( $M = -0.05$ ,  $SEM = 0.01$ , measured as a log proportion of training S+ response rate). This decrease occurred irrespective of whether the

rate of food during punishment tests had increased or decreased compared with training. The present study uses a procedure adapted from the procedure used with pigeons to investigate whether a contingent S- stimulus will punish key-pressing maintained by money reinforcers in humans. If a contingent S- stimulus will reduce behaviour maintained by positive reinforcers in humans, this type of punishing stimulus could be developed for use in combination with differential-reinforcement methods to decrease levels of harmful behaviour and maladaptive choice in applied treatment settings.

In the present study, 30 human participants pressed keys on the keyboard of a laptop computer. During training, a two-component multiple schedule was arranged so that key presses in the presence of a green computer screen (S+ component) produced a money reinforcers every 15-s on mean (VI 15 s), while key presses in the presence of a red computer screen (S- component) never produced money reinforcers. A punishment test immediately followed the training phase. During the punishment test, the training procedure was repeated except that every five key presses on mean (VR 5) during the S+ component produced the super-imposed red screen (S-) for 1.5-s. If the rate of key-pressing in the presence of S+ decreases during the punishment test, we can conclude that the S- contingent stimulus punished responding for money reinforcers in humans.

## **METHOD**

### **Participants**

The participants were 30 undergraduate students at the University of Auckland, and Auckland University of Technology, recruited between March and October, 2017. There were 5 males and 25 females; aged between 18 and 30 years. There were three conditions in total. As each participant completed only one session, the 30 participants were divided into three groups, with 10 participants in each condition. Participants were labelled with two numbers. The first number indicated which condition the participant was in, and the second indicated the order of the participant within that condition. Therefore, the first participant in Condition 1 is labelled 1-1, and the first participants in Conditions 2 and 3 labelled 2-1, and 3-1, respectively. Participants were seated alone in a quiet room to complete a computer task within one session.

### **Pre-experiment information**

In pre-experiment advertisements, potential participants were advised that the purpose of the research was to study the impact of simple environmental cues on the rate of

responding for reward. They were told they would be completing a computer task, which would result in money reinforcers. Recruited participants were then provided with the following information: “For the computer task, you will press the keyboard space bar throughout the task, whilst looking at a computer screen. The computer screen will either a) be one of two colours, or b) remain blank, or c) turn white for a brief period. Presses to the space bar throughout the task will result in accrual of a variable amount of money reinforcers according to a contingency known only to the researcher. However, money cannot be obtained when the screen is blank or white. Total accrued money reinforcers will be visible in the right-hand corner of coloured screens throughout the task. This amount will be redeemed as a cash payment at the end of the session. You will be required to sign a receipt when you receive the cash.”

Participants were remunerated with a guaranteed payment of \$NZ10 if the computer task was completed within the session. Up to \$NZ15 in additional payment was possible through money reinforcers accrued during the computer task. Therefore, total maximum payment for participation in the study was between \$NZ10 and \$NZ25 for each participant. Participants were advised that data generated from their session would be used in conference papers, in the PhD thesis of the first author, and might be used in journal publications. Data remained confidential, with identifying information including the participant’s name and email address not retained following initial communication. Each participant was assigned a code which, along with the participant’s age and gender, was the only identifying information placed upon the computer-based task.

### **Apparatus**

Participants were seated at a desk, on which there was a Hewlett-Packard (HP) EliteBook laptop computer running the Windows 7 Enterprise operating system. The computer-based task was programmed using the Python 3 programming language. All experimental events were recorded in Python 3, which was then converted to a Microsoft Excel spreadsheet file for data analysis.

### **Procedure**

Each participant completed a computer task during one session of approximately 25 min duration. The task consisted of a training phase and a punishment test phase. Each phase was either 10 or 15 min in duration, depending on condition. Table 5.1 provides an overview

of training and punishment test duration, and arranged reinforcer schedules, for each condition.

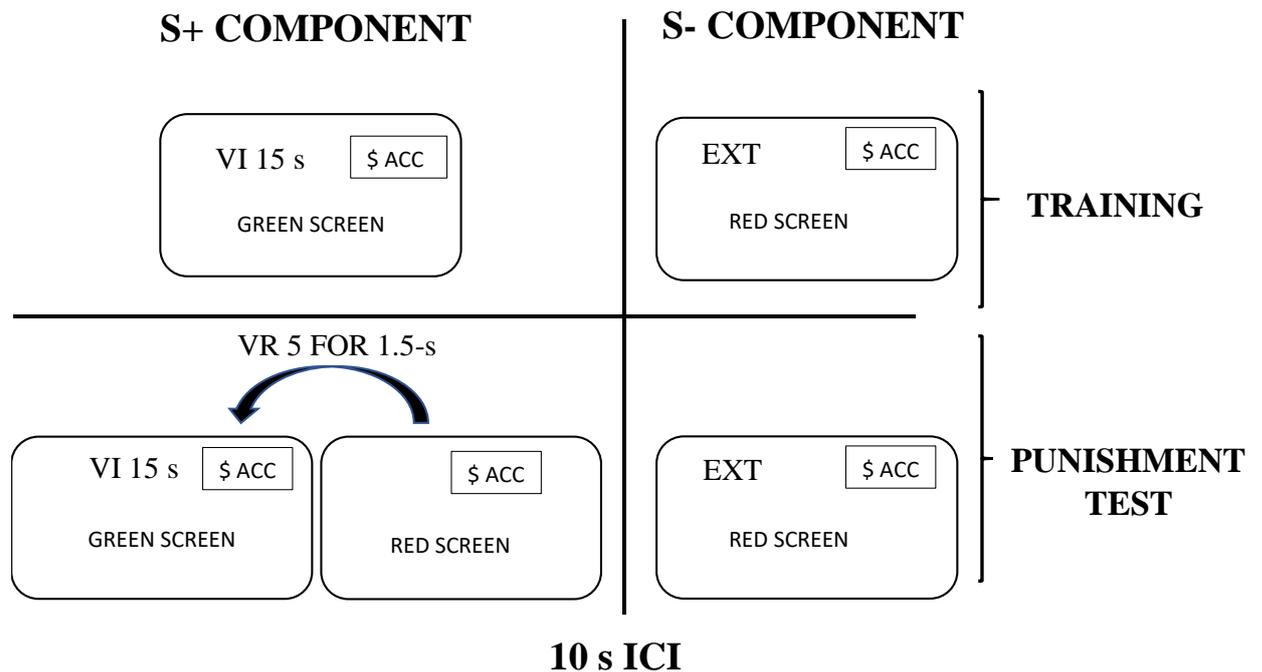
In Condition 1, training and punishment test phases were 10- and 15 min in duration, respectively. In Condition 2, the duration of training and punishment test phases was reversed to investigate the impact of longer training duration on the effectiveness of the contingent S- stimulus during punishment tests. Therefore, training duration was 15 min, and the punishment test was 10 min. As the number of participants that showed evidence of S+/S- discrimination learning decreased in Condition 2 compared to Condition 1, in Condition 3 we arranged another training duration of 15 min to investigate whether the learning results of Condition 2 were replicable. In addition, the punishment phase was extended to 15 min in Condition 3.

Table 5.1

*Total Number of Participants, Phase Duration, Allocation of VI schedules of Money reinforcers, or EXT, in each 60-s S+ and S- component, for Training and Punishment Test phases in each Condition.*

CONDITION	N	TRAINING			PUNISHMENT TEST		
			S+ COMPONENT	S- COMPONENT		S+ COMPONENT	S- COMPONENT
		Duration	Schedule value	Schedule value	Duration	Schedule value	Schedule value
1	9	10 min	VI 15 s	EXT	15 min	VI 15 s	EXT
2	10	15 min	VI 15 s	EXT	10 min	VI 15 s	EXT
3	10	15 min	VI 15 s	EXT	15 min	VI 15 s	EXT

**Training phase.** In all three conditions, participants responded by key-pressing on a two-component multiple schedule. The two components were strictly alternated, with Component 1 always occurring first. The training phase ended after a pre-arranged duration of 10 or 15 mins, depending on condition, and irrespective of current component duration. Each component was 60-s in duration, with the exception of the final component, which could be interrupted by the end of the training phase. Each 60-s iteration of either component was followed by a 10-s inter-component interval (ICI) in which the laptop screen was completely blank. Key-presses during the ICI were not recorded, and could not produce money reinforcers. Figure 5.1 shows how each component was signaled by discriminative stimuli during training and punishment test phases across all conditions.



*Figure 5.1* Diagram showing stimulus and scheduled money reinforcers arrangements for S+ components (Green screen) and S- components (Red screen) in Training (top panel) and Punishment Test (bottom panel) sessions for all conditions. A response-contingent red screen was superimposed for 1.5-s in S+ components during Punishment Tests according to a VR 5 schedule. Accrued money reinforcers (ACC) are displayed whenever the screen is green or red.

In Component 1, the computer screen was entirely green, with the exception of a small white rectangular box in the top right-hand corner of the screen. The box showed the total accrued amount of money reinforcers. Accruals could only occur in the presence of the green screen. Therefore, Component 1 is referred to as the *S+ component*. Each accrual was worth \$NZ 0.20 (20 cents) in money reinforcers. Reinforcers accrued according to a VI schedule that arranged an exponential distribution of reinforcers by interrogating a probability generator every 1 s. This ensured that key-presses produced money reinforcers every 15-s on mean. Whenever a money reinforcer occurred, the screen changed from green to white for 3-s and participants heard the sound of a cash register ringing. After 3-s, the screen returned to green, with the newly accrued total visible in the white box in the top right-hand corner of the screen. In Component 2, the computer screen was entirely red, with the exception of the white rectangular accrual box in the top right-hand corner of the screen. The box displayed the total amount of reinforcers accrued across iterations of Component

1. However, key-presses during Component 2 could not produce money reinforcers. Therefore, Component 2 is referred to as the *S-* component.

**Punishment Tests.** When the training phase ended, the punishment test phase was immediately initiated. During the punishment test, S+ and S- components alternated as in baseline, but key presses in S+ could produce a brief (1.5-s) presentation of the red S- screen, as well as reinforcers on a VI 15-s schedule. The super-imposition of the red screen on the green screen for 1.5-s in the S+ component is hereafter referred to as the *contingent S-* stimulus. The VI 15-s schedule of money reinforcers was maintained throughout the S+ component during punishment tests, irrespective of whether key-presses produced the contingent S- stimulus or not. That is, any key-press made in the S+ component could produce money reinforcers, whether the screen was green or red. Key-presses during the S- component could not produce money reinforcers. The punishment test phase ended after a pre-arranged duration, depending on condition, and irrespective of the duration of the final component.

**Dependent measures.** During the training phase, for each participant, mean responses per min were calculated across each complete pair of S+ and S- components. To assess whether participants had learned that S+ signaled the availability of money reinforcers and S- signaled an absence of money reinforcers, we calculated the log proportion of S+ to S- response rate across the last two complete pairs of S+ and S- components. If log proportion of S+ response rate was positive, participants responded at a greater rate in S+ components than S- components. If log proportion of S+ response rate was negative, participants responded at a greater rate during S- components than S+ components. If log proportion of S+ response rate was 0.20 or higher, we concluded that participants had learned the different reinforcer contingencies signaled by S+ and S-. This criterion required participants to make at least 1.6 responses during S+ components for every 1 response made during S- components (i.e. a minimum discrimination index of 0.61) Although this index value is low compared to discrimination learning in animal studies (e.g. Chapter 4), the criterion was considered appropriate due to there being only one training phase of relatively short duration.

During punishment tests, in the S+ component we calculated responses made in the presence of the green screen (S+), and divided these by total minutes in which only the S+ stimulus was present. Therefore, responses, money reinforcers, and time that accumulated

during 1.5-s presentations of the contingent S- stimulus, and time accumulated during 3-s money reinforcer deliveries, were not included in the calculation. This allowed us to compare S+ response rates between training and punishment test phases. In the S- component, all responses made in the presence of the red screen (S-), were divided by the total minutes of that component.

## RESULTS

Of 30 participants, we excluded data for one participant due to the failure of that participant to make any responses during the punishment test phase. Therefore, we included data for 29 participants who completed the computer-based task correctly. There were nine participants in Condition 1, and 10 in each of Conditions 2 and 3.

**Training phase.** Table 5.2 shows group mean responses and obtained money reinforcers per min in the S+ component, and responses per min in the S- component, for each condition. The Group mean S+ response rate was lower in Condition 1 compared to the other conditions. This occurred due to two participants (1-1 and 1-2) responding at a considerably lower rate ( $M = 48.00$  responses per min) than the other seven participants ( $M = 300.43$ ) in that condition. Money reinforcers per min across Conditions 1 to 3 closely followed the arranged reinforcer schedule for the S+ component ( $M = 4.00$ ). However, group mean rate of money reinforcers was greater in Condition 1 compared to the other conditions. This occurred due to three participants obtaining considerably more money reinforcers ( $M = 9.33$ ) than the other six participants in that condition.

Figure 5.2 shows the log proportion of S+ to S- response rate during training for individual participants in Condition 1 (left panel), Condition 2 (middle panel), and Condition 3 (right panel). Log proportion of S+ to S- response rate was greater than 0.20 for five of nine participants in Condition 1, and three of 10 participants in both Conditions 2 and 3. Therefore, more participants learned to discriminate between S+ and S- in Condition 1, compared with Conditions 2 and 3.

Table 5.2

Group Mean responses and Obtained Money reinforcers per min in the S+ component, and Responses per min in the S- component, with SEM, across the last two Pairs of Complete S+ and S- Components during Training, in each Condition. The Discrimination Index shows the Group Mean Proportion of S+ to S- response rates.

CONDITION	N	Group mean baseline responses per min				Discrim. Index S+	Group mean baseline RF		
		S+ Component	SEM (S+)	S- Component	SEM (S-)		Schedule	RF/min	SEM
1	9	149.36	31.49	76.44	20.44	0.66	VI 15 s	4.25	0.55
2	10	211.66	40.09	173.73	27.77	0.55	VI 15 s	3.35	0.42
3	10	277.18	18.29	201.50	33.37	0.58	VI 15 s	4.03	0.52

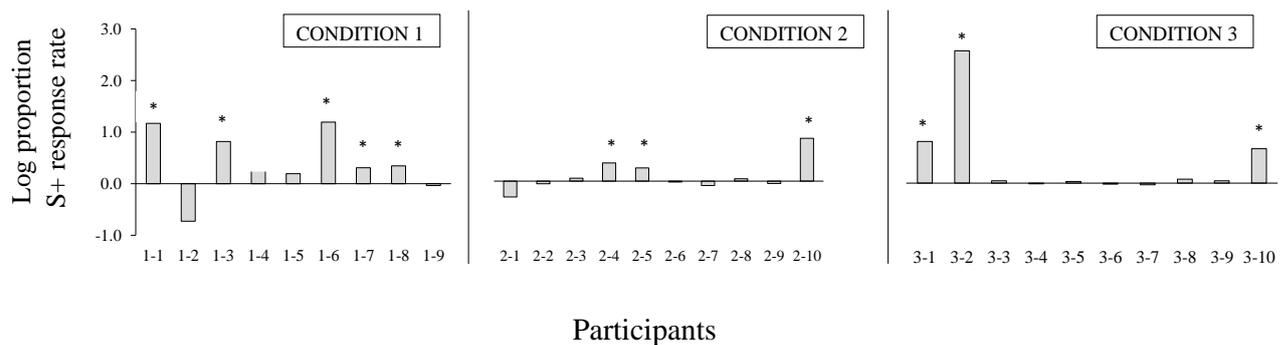


Figure 5.2 Log proportion of S+ to S- response rate during the last two complete pairs of S+ and S- components in training for individual participants in Condition 1 (left panel), Condition 2 (middle panel), and Condition 3 (right panel). An asterisk indicates log proportion of S+ response rate of 0.20 or higher.

As the only difference between Condition 1 and the other two conditions was duration of the training phase, these results suggest that shorter training duration might enhance discrimination learning. Overall, log proportion of S+ to S- response rates suggest that S- was established as a signal for the absence of money reinforcers for 11 of 29 participants, across all three conditions. Hereafter, these 11 participants are referred to as 'the learners'. Table 5.3 shows group mean responses and obtained money reinforcers per min in the S+ component, and responses per min in the S- component, across the last two pairs of complete S+ and S- components during training, in each condition, for the learners.

Compared with all participants, group mean S+ response rate was greater, S- response rate lower, and the discrimination index of S+ to S- response rate greater, for the learners.

Table 5.3

*Group Mean responses and Obtained Money Reinforcers per min in the S+ component, and Responses per min in the S- component, with SEM, across the last two Pairs of Complete S+ and S- components during Training, in each Condition, comprised of Individual Learners with a S+/S- Discrimination Index of 0.61 or higher.*

CONDITION	N	Group mean baseline responses per min				Discrim. Index S+	Group mean baseline RF		
		S+ Component	SEM (S+)	S- Component	SEM (S-)		Schedule	RF/min	SEM
1	5	163.47	43.60	47.10	19.73	0.78	VI 15 s	4.75	0.55
2	3	312.79	52.65	139.17	51.92	0.69	VI 15 s	4.24	1.66
3	3	305.06	16.46	36.50	18.17	0.89	VI 15 s	4.00	0.24

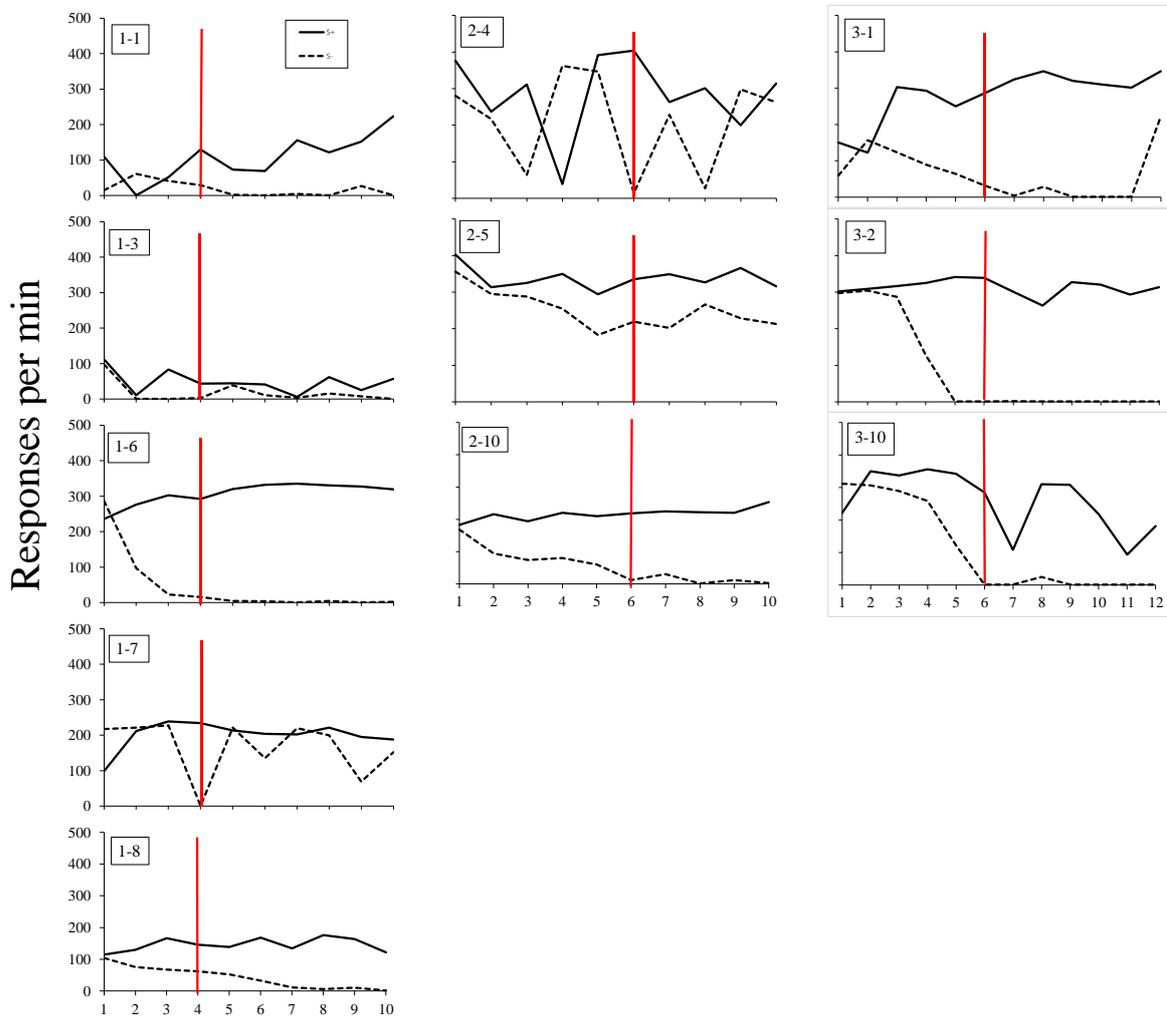
### **Punishment test phase.**

Figure 5.3 shows S+ and S- response rates for each complete pair of S+ and S- components across training and punishment test phases, for each of the 11 learners. Across all three conditions, S+ response rate did not decrease during the punishment test phase, compared with the training phase, for Participants 1-3, 1-6, 1-7, 1-8, 2-5, 2-10 and 3-1. That is, there was no obvious effect of the contingent S- stimulus for 7 of 11 learners. S+ response rate decreased in the first S+ component of the punishment test phase, compared with the last S+ component of the training phase, for Participants 1-1, 2-4, 3-2, and 3-10. This decrease was not maintained across subsequent S+ components in the punishment test phase.

## **DISCUSSION**

We investigated the extent to which a response-contingent 1.5-s negative discriminative stimulus (contingent S- stimulus) which previously signaled an absence of money reinforcers would decrease the rate of key-pressing by humans in the presence of a different stimulus signaling availability of money reinforcers (S+). Of 29 participants, 11 learned the different reinforcer contingencies signaled by S+ and S- during training (the 'learners'). For a majority of learners, the rate of S+ responding did not decrease when S+

responses sometimes produced S-. For a minority of learners, the rate of S+ responding briefly decreased in the first S+ component of the punishment test phase, compared with the last S+ component in training.



### S+/S- component pairs

Figure 5.3 S+ response rate (black line) and S- response rate (black dashed line) for each S+/S- component pair (y-axis) across training and punishment test phases for individual learners in Condition 1 (left panel), Condition 2 (middle panel), and Condition 3 (right panel). The vertical line indicates the end of the training phase.

## DISCUSSION

We investigated the extent to which a response-contingent 1.5-s negative discriminative stimulus (contingent S- stimulus) which previously signaled an absence of

money reinforcers would decrease the rate of key-pressing by humans in the presence of a different stimulus signaling availability of money reinforcers (S+). Of 29 participants, 11 learned the different reinforcer contingencies signaled by S+ and S- during training (the 'learners'). The rate of S+ responding did not decrease when S+ responses sometimes produced S- in a majority of learners. The rate of S+ responding briefly decreased in the first S+ component of the punishment test phase, compared with the training, in a minority of learners.

The results of the present study do not reliably support the findings of the previous study with pigeons (see Chapter 4). More research is required to investigate whether a contingent S- stimulus will function as a punisher across species, and across different types of responses, and reinforcers. However, a brief effect of the contingent S- stimulus on S+ response rate in a minority of learners suggests it is possible that a contingent S- stimulus *might* reduce behaviour maintained by access to positive reinforcers in humans. Applied research shows that the same punishing stimulus delivered in the same way will successfully punish behaviour in some humans, but not others (e.g. Chapman, Smith & Layden, 1971). Therefore, further research with humans is needed to investigate whether changes to procedural variables will enhance any aversive properties of a contingent S- stimulus, producing a reliable effect of the S- contingent stimulus. For example, the effectiveness of a punisher can vary depending on the frequency of magnitude of reinforcers maintaining behaviour. Sierra-Machado et al. (2015) used eye puffs to punish button-pressing for tokens in both humans and primates. They manipulated both the probability of punishment and reinforcer magnitude. When there was a high probability of punishment, response rate maintained by smaller magnitude reinforcers decreased for both humans and primates. However, response rate maintained by larger magnitude reinforcers did not decrease. Similarly, Geller (1970) found that while low levels of shock intensity reduced lever-pressing for food pellets in rats, high levels of shock intensity did not reduce lever-pressing for electrical brain stimulation (also see Beyra & Spinewine, 1973). Given these findings, it is possible that a contingent S- stimulus that previously signaled an absence of food reinforcers has a greater negative value for food-deprived pigeons, than a contingent S- stimulus that previously signaled an absence of money reinforcers has for humans. Further research is needed to investigate whether a contingent S- stimulus that previously signaled the absence of a greater magnitude reward will reliably punish human behaviour.

Another difference between the findings of the previous study with pigeons, and the present study, is that all five pigeons learned the different contingencies signaled by S+ and

S- stimuli in training. In contrast, only 11 of 29 human participants showed evidence of discrimination learning. Research suggests there is little difference between the ability of animals and humans to learn to discriminate between stimuli signaling different environmental contingencies (Harlow, 1949; Mackintosh, 1965). However, the two studies differ not only in species, but also in aspects of procedure, training amount, and training duration.

A limitation of the present study was that, due to resource and participant retention constraints, we arranged only one training phase, which was immediately followed by a single punishment test phase. In contrast, in Chapter 4 we arranged three blocks of 10 training sessions, with each block interspersed by two punishment test sessions. Therefore, the design of the present study is AB, compared to ABABAB for the pigeon study. The previous study with pigeons also used a Control condition to show that decreases in S+ response rate were due to the contingent S- stimulus, and not simply the production of a novel contingent stimulus. We were unable to include a Control condition in the present study due to resource constraints.

Research suggests that amount of training is positively correlated with increased discrimination of stimuli in animals (Farthing & Hearst, 1968; Greenway, 1975; Margolius, 1955; Sewell, Nickel, Johnson, & Williamson, 1979; Vyazovska, Navarro, & Wasserman, 2016; but see Selekman, 1973). It is also possible that an interval between multiple training sessions contributes to an increase in the rate of acquisition in discrimination learning (Holland & Morrell, 1996). Therefore, it is possible more participants would have become learners in the present study if the training phase had been repeated more than once prior to the punishment test, and with an interval between training phases. Future research is needed to discover whether these procedural changes would increase the number of learners in a replication of the present study.

It is surprising there were more learners in the condition with the shortest training duration (i.e. 10-min, Condition 1) than Conditions 2 and 3 (15-min) in the present study. Although it is possible participants attended less to discriminative stimuli during the slightly longer training phase of Conditions 2 and 3, this suggestion is not supported by animal research. Instead, research suggests persistence of attending to stimuli is mainly governed by positive reinforcement rates in the training context (Shahan and Podlesnik, 2008). As obtained positive reinforcement rates were similar across all three conditions during the training phase (see Table 5.2), it is unlikely that participants attended less to stimuli presented across a 15-min training phase, compared with a 10-min training phase.

To investigate whether training duration *did* contribute to the probability of learning the S+/S- discrimination, we compared training data across all three conditions using the last two complete pairs of S+ and S- components that were presented *before 10 mins had elapsed* for Conditions 2 and 3. As the training phase in Condition 1 was 10-min, the reanalysis allowed us to compare directly log proportion of S+ to S- response rate across conditions, using components that occurred at the same point in training. Results of this analysis showed there were fewer learners after 10 min of training in Conditions 2 and 3 than there were after 15 min had elapsed. Therefore, it is unlikely that the relatively longer training duration of Conditions 2 and 3 contributed to fewer learners, compared to Condition 1. This makes sense, given research that shows more training enhances, rather than reduces, discrimination learning (e.g. Vyazovska, Navarro & Wasserman, 2016; Wright, 1997).

As we used a between-subjects design across conditions, we might conclude that relatively fewer learners were recorded in Conditions 2 and 3 due to individual learning variability. We did not ask participants to complete an exit question to subjectively assess the extent to which each participant discriminated the arranged S+ and S- reinforcer contingencies. Even so, anecdotally, and unprompted, some participants offered versions of the comment “I figured out you could only earn money in green” while others said they “really had no idea what was going on”. Replications of this study would benefit from an exit question that contributes to the process of identifying learners.

It is possible that changes to procedural variables would enhance any aversive properties of a contingent S- stimulus, leading to more reliable findings than those of the present study. For example, in addition to changes to reinforcer magnitude or frequency, continuous punishment has been found to be more effective than intermittent punishment in decreasing responding maintained by positive reinforcers (Lerman, Iwata, Shore, & DeLeon, 1997; Linden, 1976). It would also be interesting to investigate whether a contingent S- stimulus will reliably shift choice in humans, as occurred for the pigeon subjects in Study 1. As noted in Chapter 4, choice is more sensitive to consequences compared with simple schedule responding on multiple schedules (see Catania, 1963; Davison & McCarthy, 1988).

Overall, punishment by the contingent S- stimulus did not occur for most learners in the present study. However, there was a brief effect of the contingent S- stimulus for a minority of learners in the first S+ component of the punishment test phase. Given the findings of Study 1 with pigeons, there is merit in continuing to investigate the potential for a contingent S- stimulus to punish human behaviour.

## CHAPTER 6.

### SUBOPTIMAL CHOICE RESEARCH AND MELIORATION THEORY: AN OVERVIEW

So far, we have shown that a contingent S- stimulus will reliably shift choice in pigeons, and might have the potential to punish response rate maintained by positive reinforcers in pigeons and humans. The remaining studies of this thesis combine to investigate how we might apply a contingent S- stimulus to punish suboptimal choice processes associated with addictive behaviour in humans. The purpose of this chapter is to provide a brief overview of suboptimal choice research, including variables that impact suboptimal choice. We also introduce melioration theory as one explanation of suboptimal choice.

Evolutionary or economic theories of behaviour often posit a normative account in which organisms behave so as to achieve the highest possible overall rate of reinforcement. These are maximizing, or optimizing, accounts (e.g., Einhorn & Hogarth, 1981). However, there are many experimental arrangements in which organisms behave in such a way as to forfeit a higher overall reinforcer rate (e.g., Spetch, Belke, Barnet, Dunn, & Pierce, 1990; Zentall, 2014, 2016a, 2016b). For example, Kendall (1974; 1985) found that, under specific conditions, pigeons preferred a key that led to a lower overall rate of reinforcers, to a key that led to a higher overall rate of reinforcers. Patterns of choice that lead to fewer reinforcers overall are typically termed *suboptimal*.

Suboptimal choice remains an important research area because of its possible link with maladaptive behaviour in humans. Maladaptive behaviour often appears in human pathologies such as gambling, substance abuse, and eating disorders, all of which result in a net loss of resources (Heyman & Dunn, 2002; Heyman, 1996; Molet et al., 2012). Many theories suggest genetic inheritance, physiological factors and psychopathology contribute to the development and maintenance of addictive behaviour. However, both human and animal studies show that despite the physiological impact of harmful substances, choice remains sensitive to the consequences of behaviour, and to competing sources of reinforcement (Grant & Dawson, 2006; Berridge & Robinson, 2003). That is, so-called ‘addictive’ behaviour is choice (Heyman, 1996; 2013; Heather, 2017; also see Read & Roelofsma, 1999). For example, Lenoir, Serre, Centin & Ahmed (2007) studied the addictive effects of heavy doses of self-administered cocaine in rats. Prior to cocaine

binging, the rats were preference tested for cocaine compared to saccharin and were shown to prefer saccharin 75 percent of the time. After cocaine binging, the rats were tested for cocaine addiction as measured by changes in motor function and mid-brain dopamine levels. Once addiction was established through physiological testing, the rats were again preference tested for cocaine and saccharin. Lenoir et al. found the rats' preference for saccharin had not changed. That is, they still preferred saccharin 75 percent of the time. When the saccharin doses were made contingent on delays and higher task requirements, the rats still only preferred cocaine 50 percent of the time. Lenoir et al. concluded that while cocaine use had altered the rats' brain and motor function, tests of *preference* for cocaine were negative. Therefore, environmental contingencies, not an innate vulnerability to addictive behaviour, determined the extent to which the rats preferred cocaine.

#### *What drives suboptimal choice?*

If addictive behaviour is suboptimal choice, what drives it? Spetch et al. (1990) used a concurrent chains procedure to investigate the role of different variables in suboptimal choice in pigeons. In a concurrent chains procedure, organisms choose between two concurrently available alternatives called *initial links*. After one response to either alternative, the alternative that is not chosen becomes unavailable, and the schedule associated with the chosen alternative, one of the *terminal links*, is initiated. Once the terminal link is completed, the initial links become available again. Usually, the same discriminative stimuli are used to signal the initial links, and different discriminative stimuli are used to signal terminal links, or sub-alternatives within a terminal link.

Spetch et al. (1990) investigated the effects of signaled reinforcer probabilities on suboptimal choice. Figure 6.1 provides an overview of the procedure they used. In the signaled condition, the initial links were signaled by left and right white keylights. A peck to the left key produced either a blue or a green keylight. If the keylight was green, there was a 100% probability of reinforcers. If blue, there was a 100% probability of no reinforcers. In comparison, a peck to the right key produced a red keylight which signaled a 100% probability of reinforcers. Therefore, in the signalled condition the left key was considered 'unreliable' and the right key 'reliable' in terms of the arranged reinforcer contingencies. In the unsignaled condition, the initial links were the same as for the signaled condition. During terminal links, a peck to the reliable right key also produced the same stimulus and reinforcer outcome as for the signaled condition. However, a peck to the unreliable left key led to only a 50% probability of reinforcers irrespective of whether the

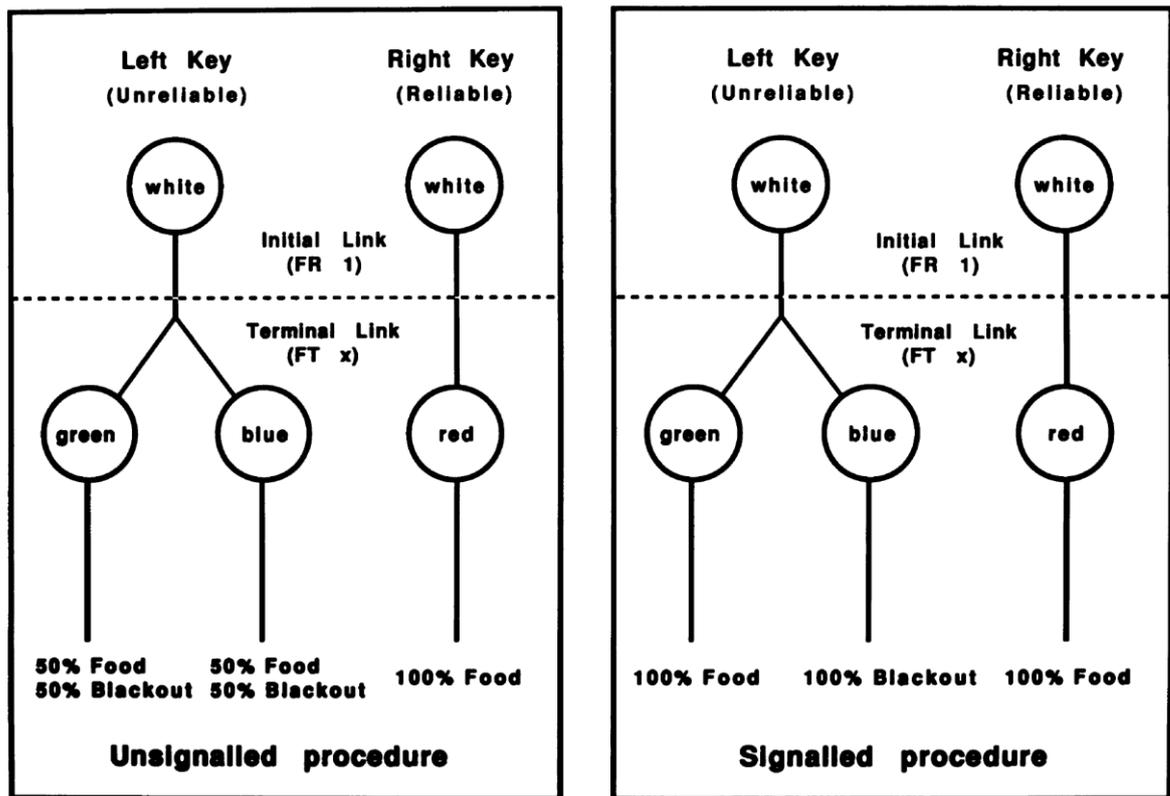


Figure 6.1 Illustration from Spetch et al. (1990), showing a concurrent-chains percentage-reinforcement procedure. FT x signifies that terminal links differed in duration by a fixed amount of time. FR signifies a fixed rate of one response required to initiate the terminal link. ‘Reliable’ signifies reinforcers were always available on the terminal link; ‘Unreliable’ signifies reinforcers were available on 50% of trials.

green or the blue keylight were produced. That is, the green and blue keylights no longer signaled different outcomes.

Spetch et al. (1990) found that when the terminal links were very short in duration (10-s), preference for the reliable key was high in both signaled and unsignalled conditions. However, when the terminal links were longer in duration (30-s), preference for the reliable key *decreased* in the signaled condition compared with the unsignalled condition, even though this led to the pigeons obtaining a lower overall rate of reinforcers. Spetch et al. attributed this to the likelihood that responses to the unreliable key produced conditioned reinforcers in the signaled condition, whereas responses to the reliable key did not. Briefly, positive reinforcers are final outcomes of behaviour, such as food. Conditioned reinforcers are interim outcomes of behaviour that provide information about positive reinforcers (see Shahan, 2010). Spetch et al. concluded that, as delays to reinforcers lengthen, pigeons prefer

the alternative that will produce an immediate *conditioned* reinforcer, even if this reduces the overall number of positive reinforcers they receive. These findings have since been replicated and extended (Smith & Zentall, 2016; Zentall, Laude, Stagner & Smith, 2015).

Since 2010, Zentall and colleagues have investigated a wide range of variables found to impact the extent to which suboptimal choice occurs in pigeons and humans. Their research shows that suboptimal choice is more likely to occur in pigeons when food deprivation is greater (Laude, Pattison and Zentall, 2012), and less likely to occur if pigeons are housed in a socially enriched environment (Pattison, Laude and Zentall, 2013). This suggests that motivation for reinforcers and social environmental factors might both impact suboptimal choice. A suboptimal alternative may also be preferred if it provides greater certainty of reinforcers (Kendall, 1974; Zentall & Laude, 2013; Zentall & Stagner, 2011), or less delay to reinforcers, even though it produces a lower overall rate of reinforcers (Herrnstein, Prelec, & Vaughan, 1986). Other studies show that suboptimal alternatives may be preferred if suboptimal choice across a block of trials determines the probability of reinforcers in the following block of trials (Tunney & Shanks, 2002). Recently, Hinnenkamp, Shahan, and Madden (2017) showed that a suboptimal alternative can be preferred even if the response cost is higher than that of an optimal alternative.

### *Pigeons, rats and people*

Some studies have investigated whether the extent to which suboptimal choice occurs is species-dependent. For example, Trujano and Orduna (2015) found that unlike pigeons (Gipson, Alessandri, Miller & Zentall, 2009), and humans (Molet et al., 2012; but see Fortes, Case, Jacob & Zentall, 2017), rats preferred an unsigned alternative associated with 50% probability of reinforcers over a signaled alternative associated with 20% probability of reinforcers (also see Roper & Baldwin, 2004). That is, rats chose optimally. However, Chow, Smith, Wilson, Zentall & Beckmann (2017) found that rats choose suboptimally if a stimulus signaling the suboptimal alternative has incentive value. ‘Incentive value’ is ascribed to a stimulus that supports species-specific sign- or goal-tracking behaviour. For example, a pigeon will visually sign-track towards a keylight, even if doing so moves it progressively further from reinforcers, thereby forfeiting reinforcers across time (Hearst & Jenkins, 1974; Silva, Silva, & Pear, 1992). Similarly, rats will sign-track to levers, but not lights (Holland et al., 2014). Chow et al. showed that when a discriminative stimulus signaling a lower probability of food requires a lever press, and a different stimulus signaling a higher probability of food requires a nose poke, rats will

choose suboptimally. These findings suggest the extent to which suboptimal choice occurs in different species might depend on whether the discriminative stimuli in the environment have incentive value (but see Martinez et al., 2017).

### *Suboptimal choice and gambling behaviour*

The above studies suggest the reasons for suboptimal choice are not reducible to a single mechanism. However, if addictive behaviour in humans is choice, continued suboptimal choice research with animals is useful for understanding the behavioural mechanisms that underlie such behaviour. Stagner and Zentall (2011) suggested the contingencies maintaining suboptimal choice in pigeons and rats might be similar to those that maintain gambling behaviour in humans. In their study, pigeons chose between two alternatives. The suboptimal alternative produced two different stimuli, one signaling 10 food pellets on 20% of trials, and the other signaling no food pellets on 80% of trials. The optimal alternative also produced two different stimuli, each signaling three food pellets on all trials. The suboptimal alternative led to the mean of two food pellets across all trials, compared to three food pellets for the optimal alternative. Yet Stagner and Zentall found the pigeons preferred the suboptimal alternative. Therefore, as occurs in human gambling behaviour, preference for an infrequent stimulus associated with a larger gain is not offset by a frequent stimulus associated with certain loss.

Molet et al. (2012) investigated whether humans would behave in a similar way to pigeons when exposed to a procedure similar to that of Stagner and Zentall (2011). Participants were selected and grouped according to self-reported gambling behaviour. Then, using screens taken from video games, participants chose between two alternatives. One alternative led to either a red planet, which appeared on 20% of trials, or a green planet, which appeared on 80% of trials. The participant could only shoot at enemy spaceships and kill 10 enemy generals on the red planet. The second alternative led to a blue planet, which appeared on 20% of trials, or a yellow planet, which appeared on 80% of trials. The participant could shoot at enemy spaceships and kill three enemy generals on either the blue or the yellow planet. Molet et al. found that participants with a background of gambling behaviour preferred the first, suboptimal, alternative to a greater extent than participants that did not have a background in gambling behaviour. Molet et al. concluded that suboptimal choice tasks involving animals are suitable for the investigation of choice processes that underlie suboptimal decision-making in humans (also see Laude, Beckmann, Daniels & Zentall, 2014).

### *Suboptimal choice – a moment by moment thing?*

The findings of the studies described thus far suggest suboptimal choice might be driven by a preference for alternatives that lead to conditioned reinforcers, and alternatives that deliver greater return in the moment, despite greater overall loss. However, it is also possible suboptimal choice occurs for no other reason than animals and humans prefer to continually choose whatever seems 'better' in the moment. Tunney and Shanks (2002) tested this hypothesis by using money to reinforce choices of left and right buttons on a computer screen by human participants. The probability of reinforcement for consistently selecting the left button across 10 consecutive trials was 0.66, compared with 0.33 for consistently selecting the right button across 10 consecutive trials. However, probability of reinforcement for choosing the right button immediately after consistently selecting the left button was 0.99, compared with 0.00 for choosing the left button immediately after consistently selecting the right button. Although exclusively choosing the left button yielded the highest cumulative payoff, Tunney and Shanks found that 11 of 12 participants consistently chose the right button, leading to suboptimal choice.

Herrnstein, Prelec and Vaughan (1986) also used money to reinforce choice between two alternatives by human participants. In their task, choice produced a delay to monetary reinforcement of between 2 and 8 s. The value of the delay depended on how choice was allocated across the previous block of 10 trials. Consistently choosing alternative 1, which initially produced a 2-s delay, increased the delay for alternative 1 in subsequent trial blocks. Consistently choosing alternative 2, which initially produced an 8-s delay, *decreased* the delay in subsequent trial blocks. Yet although exclusively choosing alternative 2 produced the greatest overall amount of reinforcement, most participants preferred alternative 1, with some preferring alternative 1 exclusively.

Although these findings suggest that suboptimal choice outcomes occur because organisms prefer the option that seems better in the moment, they do not explain the underlying behavioural mechanism by which organisms monitor and regulate choice in order to choose the option that seems better in the moment. However, the studies by Tunney and Shanks (2002) and Herrnstein, Prelec and Vaughan (1986) do support the predictions of melioration theory (Herrnstein & Vaughan, 1980).

*Melioration: a theory of suboptimal choice*

As discussed in Chapter 3, melioration theory presumes that, at any given moment, organisms are attempting to equalise local rates of reward between competing choice alternatives. The term ‘local rate’ describes the number of rewards obtained for behaviour divided by time spent on behaviour. When the local rate of reinforcement between alternatives is not equal, melioration theory predicts that choice will shift so as to minimize the difference in local reinforcer rates. During this process, a greater proportion of time is allocated to the alternative that produces the higher local rate of reinforcement in the moment. Once local rates of reinforcement are equal between concurrently available alternatives, the melioration equation may be written as:

$$\frac{R_1}{T_1} = \frac{R_2}{T_2} \quad (1)$$

with  $R_1$  and  $R_2$  representing obtained reinforcement for alternatives 1 and 2, respectively;  $T_1$  and  $T_2$  representing time allocated to alternatives 1 and 2, respectively. For standard concurrent-choice scenarios, melioration theory accommodates choice data able to be described by the strict matching law (Herrnstein 1961, 1970). Strict matching dictates that the overall proportion of time or responses allocated to any one alternative among competing alternatives will equal the overall proportion of reinforcers obtained from that alternative. Therefore, Equation 1 may be rearranged to give strict matching. The time-based strict matching equation is written as:

$$\frac{T_1}{T_2} = \frac{R_1}{R_2} \quad (2)$$

with terms being identical as those in Equation 1. However, for as long as the local rates of reinforcement between concurrently available alternatives are *not* equal, the melioration equation may be written as:

$$R_d = \frac{R_1}{T_1} - \frac{R_2}{T_2} \quad (3)$$

where  $R_d$  represents the signed difference (positive or negative) between local rates of reinforcement for alternatives 1 and 2. Once  $R_d$  is sufficiently close to zero, melioration theory predicts variability of choice will reduce to maintain a low  $R_d$  value, even if this leads to suboptimal choice. Therefore, melioration theory offers one explanation for how in-the-

moment choice processes contribute to overall suboptimal choice outcomes, including suboptimal decision-making in humans (Herrnstein & Prelec, 1991; Heyman, 2013; Rachlin, 1997.)

*Establishing a reliable suboptimal choice platform*

If environmental variables impact the extent to which suboptimal choice occurs (Zentall, 2014; 2015; 2016), then it is reasonable to assume a contingent S- stimulus might punish suboptimal choice. However, we first need a choice procedure that allows us to test this hypothesis. In the next chapter we replicate and extend the study by Vaughan (1981) which uses a procedure in which optimal and suboptimal choice outcomes are either compatible, or incompatible, with the process of melioration. As we have demonstrated that a contingent S- stimulus functions as a punisher in pigeons and humans (see Chapters 4 and 5), we now investigate whether a contingent S- stimulus might punish suboptimal choice processes. Our aim is to establish the extent to which Vaughan's (1981) procedure is reliable, so that we may then use the same procedure to investigate whether suboptimal choice is able to be punished with a contingent S- stimulus.

## CHAPTER 7.

### REPLICATING AND EXTENDING THE FINDINGS OF VAUGHAN (1981)

Vaughan (1981) designed a study with pigeons to test the predictions of melioration theory. In his study, maximizing the overall rate of reinforcement required behaviour that was incompatible with melioration. In each 30-min session, Vaughan (1981) arranged eight successive time periods called 'preference windows'. These were used to calculate allocation of time to left and right keys, which determined the local reinforcer rates for each key in the next preference window.

The left panel of Figure 7.1 shows Vaughan's hourly programmed reinforcer rates for responses to left and right keys in Condition A (top panel) and Condition B (bottom panel), as a function of proportion of time allocated to the right key in the previous preference window. Proportional time allocations of between 0.125 and 0.25 to the right key have been designated the 'optimal zone', because the local rates of reinforcement on the two keys are equal (melioration), and the overall reinforcer rate from both keys combined is at a maximum. Similarly, time allocations between 0.75 and 0.875 to the right key are designated as the 'suboptimal zone', because the local reinforcer rates are equal, but the overall reinforcer rate is at a minimum. Therefore, for pigeons to gain the maximum amount of reinforcers, choice needs to be between 0.125 and 0.25 in both Conditions A and B, because that produces the highest possible overall reinforcer rate.

In Condition A, if choice falls at any point outside the two melioration zones, the local reinforcer rates on the two keys are unequal. Melioration then predicts that more time will be spent on the key with the higher local rate (i.e. the option that is 'better in the moment'). For time allocations greater than 0.25, the key with the higher local rate is the left key. Therefore, melioration drives choice towards the left key, so that proportional time allocation to the right key decreases. This pushes choice back to the optimal zone. In Condition B, however, the local rate on the right key exceeds that on the left for all intermediate time allocations. Melioration therefore predicts spending more time on the right key, pushing choice towards the suboptimal zone of 0.75 to 0.875, even though such choice produces the *lowest* possible overall reinforcer rate. Thus, both melioration and maximization predict choice in the optimal zone in Condition A, but their predictions for Condition B differ.

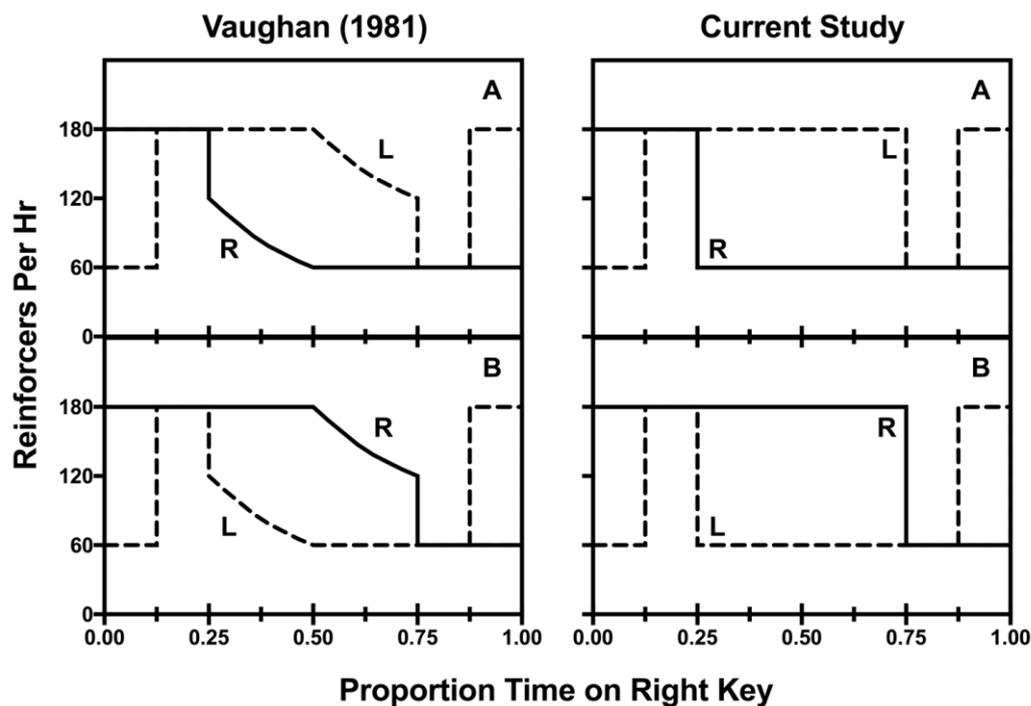


Figure 7.1 Diagram of programmed hourly rates of reinforcement (y-axis) in Condition A (top panel) and Condition B (bottom panel) for left key (dashed black line, L) and right key (solid black line, R) responses as a function of proportion of time on the right key (x-axis) in a previous preference window, for Vaughan's (1981) procedure (left panel) and the procedure for the current study (right panel). Choice is predicted to shift in a subsequent preference window when proportion of time allocated to the right key in a current window is between 0.26 to 0.74. Equal rates of reinforcement for left and right keys occur when proportion of time allocated to the right key is either 0.125 to 0.25 or 0.75 to 0.875.

In Vaughan's (1981) study, three pigeons responded on a two-key concurrent variable-interval (VI) schedule. A response to either the left key or the right key initiated a 2-s timer for that key and paused the VI-schedule timer for the other key. The timer for any one key stopped when either 2 s had elapsed without a response to that key, or when the other key was pecked. This procedure was used to ensure a precise measure of accumulated time on both keys. A 1-s changeover delay (Herrnstein, 1961) operated for switches between keys.

Vaughan (1981) found that, in the last five sessions of Condition A, all three pigeons allocated time so as to remain in the optimal zone, but had shifted to the suboptimal zone in the last five sessions of Condition B. This led to the pigeons obtaining 17% fewer reinforcers overall in Condition B relative to Condition A. Vaughan concluded that if maximization of overall reinforcer rates requires organisms to deviate from arranged local contingencies by allocating more time to a locally leaner alternative, maximization of overall reinforcer rates

is less likely to occur. Instead, pigeons will allocate more time to a locally richer alternative in support of melioration theory.

Vaughan's (1981) study had three main limitations. First, he only used three pigeons. Of these, Vaughan increased the difference in local reinforcer rates between competing alternatives for one pigeon during Condition B. This occurred due to the apparent inability of that pigeon to discriminate the difference in local reinforcer rates between alternatives. Second, Vaughan tested Condition A followed by Condition B with no replication of either condition. Therefore, Vaughan did not demonstrate whether repeated changes to local reinforcer contingencies between A and B conditions would result in repeated reversals of choice between A and B conditions, and relatedly, did not rule out order effects of conditions or mere exposure to these procedures influencing choice. Third, Vaughan's dependent measure, proportion of time on the right key, was reported across an entire session, rather than within each of the preference windows comprising one session. Thus, Vaughan's study did not provide within-session data to show that choice across short time spans is under the control of differences in local reinforcer rates.

The study presented in this chapter addresses the limitations of Vaughan's (1981) study to assess the reliability of his findings. We investigated systematic shifts in overall choice between repeated alternating conditions, and shifts of within-session choice according to arranged local contingencies.

Six pigeons were trained to respond on left and right keys on dependent concurrent VI VI schedules. We arranged three alternating cycles of Vaughan's Condition A and Condition B (i.e., ABABAB). In an across sessions choice analysis, the dependent measure of proportion of time on the right key was calculated across each session, and for each condition and cycle. To discover the extent to which within-session choice contributes to overall suboptimal choice and supports melioration theory, we also measured proportion of time allocated to the right key within each of seven of the eight successive preference windows comprising one session. Window 1 was omitted from this analysis because no reinforcement contingencies were arranged in a previous window within that session.

If melioration theory is used to explain suboptimal choice, it is necessary to demonstrate that choosing a locally richer alternative may lead to relatively lower overall reinforcer rates. Vaughan (1981) used a procedure in which local contingencies either led to, or were in opposition to, the maximization of overall reinforcer rates. He found overall choice followed arranged local contingencies whether those contingencies lead to suboptimal choice or not. The present study extends this finding by investigating whether

(1) within-session choice systematically follows arranged local contingencies across short time periods, (2) overall choice systematically changes when local contingencies change across alternating conditions and (3) overall suboptimal choice occurs through choice systematically following local contingencies. The present study provides an opportunity to assess the validity of using Vaughan's procedure and further investigate the role of melioration in overall suboptimal choice.

## **METHOD**

### **Subjects**

Subjects were six experimentally experienced homing pigeons, numbered 61 to 66. The pigeons had previous experience with two-alternative concurrent-choice schedules. Pigeons were individually housed within a pigeon colony room and maintained at  $85\% \pm 10\%$  of their free-feeding body weights using post-session supplementary feeding of mixed grain. Water and grit were available at all times. The colony room lighting was switched on at 00:00 and off at 16:00 daily. Experimental sessions began at 01:00 with a 60 min blackout. No personnel entered the room during sessions.

### **Apparatus**

The pigeons' home cages served as the experimental chambers. The cages measured 380 mm high, 380 mm wide and 380 mm deep. Three walls were constructed of sheet metal while the floor, ceiling, and front wall were metal bars. In each chamber, two wooden perches were mounted 60 mm above the floor. One perch was parallel to, and 90 mm from, the front wall while the other was parallel to, and 90 mm from, the right wall. The operant panel was located on the right wall and consisted of three 20 mm diameter translucent keys set 220 mm above the wooden perches and 85 mm apart, center to center. Only the left and right keys were used in this experiment. The center key was not used. The left and right keys could be trans-illuminated by LEDs located behind each key with red or green lights. Pecks to a lit key exceeding 0.1 N were registered as responses. The magazine aperture was located 100 mm below the center key and measured 50 mm high, 50 mm wide, and 40 mm deep. A hopper containing wheat was situated behind the magazine. During hopper presentations, all key lights were turned off and the hopper was raised and illuminated for 2 s. We excluded hopper time from all analyses. All experimental events were arranged and recorded by an IBM® PC-compatible computer running MED-PC IV® software which was located in an adjacent room.

## Procedure

As the pigeons in this study were experimentally experienced, no pretraining was required. We arranged two types of conditions: A and B (see Table 7.1). We assessed Condition A followed by Condition B three times for all pigeons. Conditions were labelled A1, B1, A2, B2, A3, and B3, with letters representing type and number representing cycle. Both Condition A and Condition B used a two-key dependent concurrent VI VI schedule (Fleshler & Hoffman, 1962). In all sessions, the left key was lit green and the right key was lit red. As in Vaughan's (1981) procedure, each response to a key initiated a timer for that key, which ran for 2 s. Time spent on a key accumulated only while the timer was running. The timer stopped if 2 s elapsed without another response, or if a reinforcer was delivered, or if a response was made to the other key. If a response was made to the other key, a timer was instead initiated for that key, but no reinforcer could be obtained for 1 s following the changeover (a 1-s changeover delay). The VI schedule associated with a key advanced only while the timer was active. If a reinforcer was arranged, no further reinforcers could be arranged for either response until that reinforcer had been obtained. Each reinforcer delivery consisted of 2 s of access to food.

Table 7.1 shows VI schedule reinforcer rates for left and right keys in Condition A and Condition B. The number of 30-min sessions per condition varied until the primary dependent measure of proportion of time allocation was stable across five sessions for all six pigeons. The number of sessions per cycle was 60, 26, and 29 for Conditions A1, A2 and A3, respectively; and 25, 26, and 34 for Conditions B1, B2 and B3, respectively. The value of the VI schedule of reinforcement for each key in any 4-min period was determined by the proportion of time allocated to the right key in the previous 4-min period. Hereafter, these four-min 'time windows' (Jones & Davison, 1997; Tonneau, 2005) are referred to as 'preference windows'. There were eight separate and consecutive preference windows in each session. In keeping with Vaughan's (1981) procedure, Windows 1 to 7 were each of 240-s duration; Window 8 was 120 s in duration. The start and end of each window was not marked by any stimulus change.

The right panel of Figure 7.1 shows programmed hourly rates of reinforcement in Condition A and Condition B for left key and right key responses as a function of proportion of time on the right key in a previous preference window. (Vaughan's 1981 procedure is shown on the left panel.) Access to the maximum rate of reinforcement within any one preference window (i.e., VI 20 s versus VI 20 s) required proportion of time on the right key

Table 7.1

*Allocation of VI schedules of Reinforcement on Left and Right keys for Conditions A and B, as a function of Proportion of Time on the Right key in the Previous 4-min Preference Window*

Relative time on right key in previous 4 min preference window*	CONDITION A		CONDITION B	
	Left key	Right key	Left key	Right key
0.0 - 0.125	VI 60	VI 20	VI 60	VI 20
0.125 - 0.25	VI 20	VI 20	VI 20	VI 20
0.25 - 0.75	VI 20	VI 60	VI 60	VI 20
0.75 - 0.875	VI 60	VI 60	VI 60	VI 60
0.875 - 1.0	VI 20	VI 60	VI 20	VI 60
* Window 1 left/right key VI values randomly assigned from VI 20 and VI 60				

for the previous preference window to be within the range 0.125 – 0.25; *the optimal zone*. Access to the minimum rate of reinforcement (i.e., VI 60 s versus VI 60 s) required proportion of time on the right key to be within the range 0.75 – 0.875 for the previous preference window; *the suboptimal zone*. The zones relate to optimal and suboptimal-rate melioration zones because if proportion of time on the right fell within either of these zones, the programmed difference in local reinforcer rates between left and right keys was zero ( $R_d = 0$  from Equation 3).

According to these contingencies, in Condition A, choice should be within, or transition toward, the optimal zone across successive preference windows, and away from the suboptimal zone. In Condition B, by comparison, choice remaining within, or transitioning toward the optimal zone required a departure from behaviour supporting melioration. However, in Condition B, choice remaining within or transitioning toward the suboptimal zone did not require a departure from melioration. For example, if the proportion of time on the right key was between 0.26 – 0.74 in any given preference window in Condition B, in the next preference window the right key would pay off at 180 reinforcers per hour, and the left key would pay off at 60 reinforcers per hour. These contingencies were reversed for the range 0.26 – 0.74 in Condition A. Therefore, if choice fell within this range in Condition B, it was not possible to return to the optimal zone unless the pigeons

consistently allocated more time to the locally leaner left key. Doing so would require a temporary departure from melioration.

Note that left- and right-key reinforcer rates for the range 0.26 – 0.74 in the present study (Figure 7.1, right panel) differed slightly from those arranged by Vaughan (1981) (Figure 7.1, left panel). In Vaughan's study, reinforcer rates within the range of 0.26 – 0.74 gradually decreased from 180 to 60 reinforcers per hr. For example, if the proportion of time on the right in a previous preference window was 0.30 or 0.40 in Condition A, while the left key then paid off at 180 reinforcers per hour, the right key paid off at 90 and 70 reinforcers per hour, respectively. If proportion of time on the right reached 0.50, the right key paid off at 60 reinforcers per hour, as per the present study. The decision to remove the curve function from programming for the present study was made to simplify programming of the contingencies. Because the arranged contingencies operate under the same maximization and local contingencies arranged by Vaughan, we did not expect that arranging a more abrupt change in rates of reinforcement between alternatives would impact the primary findings.

## RESULTS

**Response and Reinforcer rates.** Table 7.2 shows mean overall response rates per minute across the last five sessions of individual A and B conditions, for individual pigeons. The far right-hand column presents these rates averaged across all conditions. Mean overall response rates for Pigeons 61 and 64 were consistently lower across conditions compared with the other four pigeons.

Figure 7.2 shows mean overall reinforcer rates from the last five sessions of Condition A1, A2 and A3, and then averaged for individual pigeons. We used the same method to calculate mean overall reinforcer rates across replications of Condition B. We found no reliable difference in overall reinforcer rates between conditions. However, overall reinforcer rates were higher for Condition A than Condition B for four of six pigeons, with exceptions being Pigeons 61 and 64.

**Choice across sessions.** Figure 7.3 shows proportion of time on the right key for individual pigeons, averaged from the last five sessions from each condition. Choice tended to be toward the optimal zone in A conditions, and toward the suboptimal zone in B conditions

Table 7.2

*Mean Overall Response rates per minute across the Last Five Sessions of Individual A and B Conditions, for Individual Pigeons. The furthestmost right-hand Column presents the same Rates Averaged across all Conditions.*

PIGEON	CONDITION						MEAN
	A1	B1	A2	B2	A3	B3	
61	15.53	44.27	47.47	49.15	59.86	57.14	45.57
62	69.10	63.61	54.93	62.00	63.63	60.15	62.24
63	70.64	84.95	76.41	86.09	97.74	85.04	83.48
64	47.59	49.69	46.13	60.82	48.41	46.41	49.84
65	96.77	66.15	101.97	76.27	89.60	60.64	81.90
66	83.96	65.46	73.31	74.62	70.55	79.35	74.54

for all pigeons. However, there were individual differences in the degree to which choice shifted between the optimal- and suboptimal zones between conditions. For example, choice by Pigeons 61 and 64 consistently remained closer to the suboptimal zone across all conditions compared with the other pigeons, while choice by Pigeon 63 remained relatively closer to the optimal zone across all conditions compared with other pigeons.

**Within-session choice analysis.** Figure 7.4 shows mean proportion of time on the right key across successive preference windows 2-8 for individual pigeons as a mean of the last five sessions in A conditions (left column) and B conditions (right column). Window 1 was omitted from this analysis because no reinforcement contingencies were arranged in a previous window within that session. The arranged contingencies (see Figure 7.1, right panel) predict that during A conditions, the proportion of time allocated to the right key should decrease across successive 4-min preference windows toward the optimal zone. During B conditions, the proportion of time allocated to the right key should increase across successive preference windows toward the suboptimal zone.

Despite difference levels of choice variability across individual pigeons, choice was generally toward the optimal zone across successive preference windows in most A conditions for four of six pigeons (Pigeons 61 and 64 were an exception), and toward the suboptimal zone across successive preference windows in most B conditions, for all six pigeons.

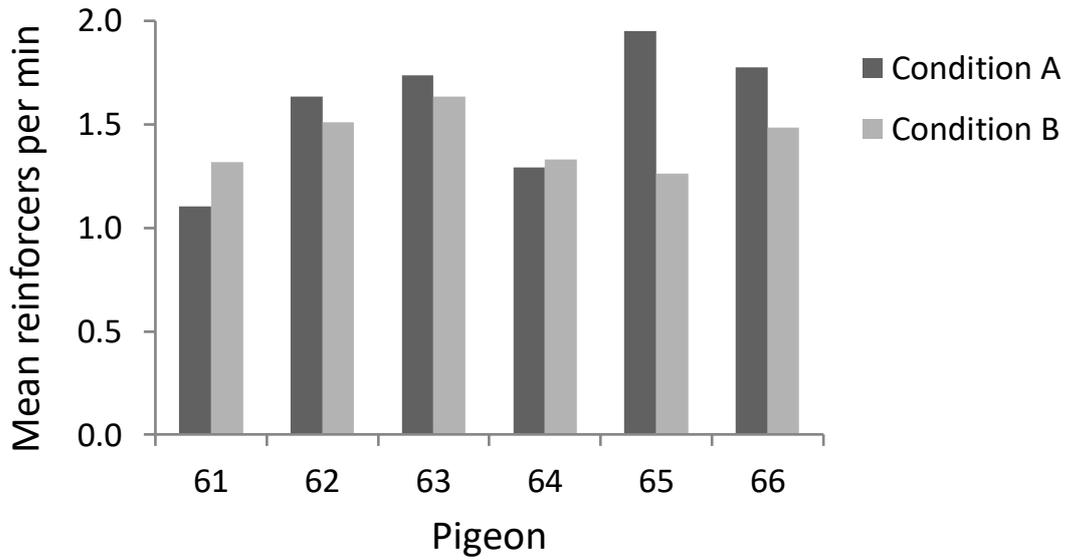


Figure 7.2 Overall obtained reinforcer-rate as the mean of the last five sessions of Condition A1, A2 and A3 (dark grey bars); and the mean of the last five sessions of Condition B1, B2 and B3 (light grey bars), for individual pigeons.

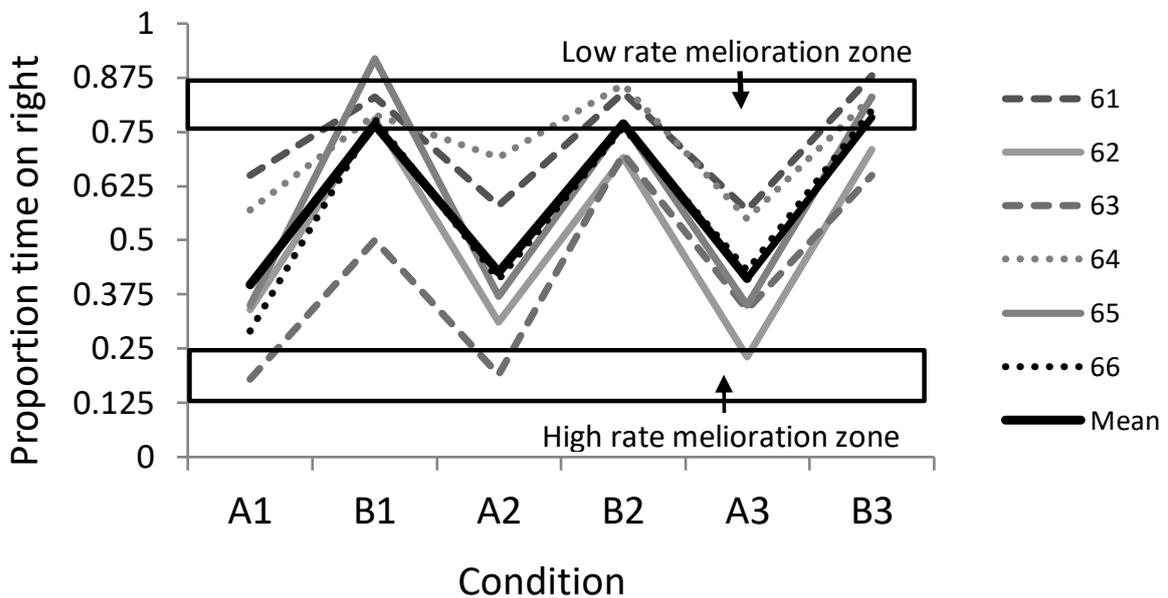


Figure 7.3 Proportion of time on the right key (y-axis) for individual pigeons, averaged across the last five sessions of successive conditions A1, B1, A2, B2, A3 and B3 (x-axis). Horizontal rectangular boxes indicate optimal (high-rate) and suboptimal (low-rate) zones.

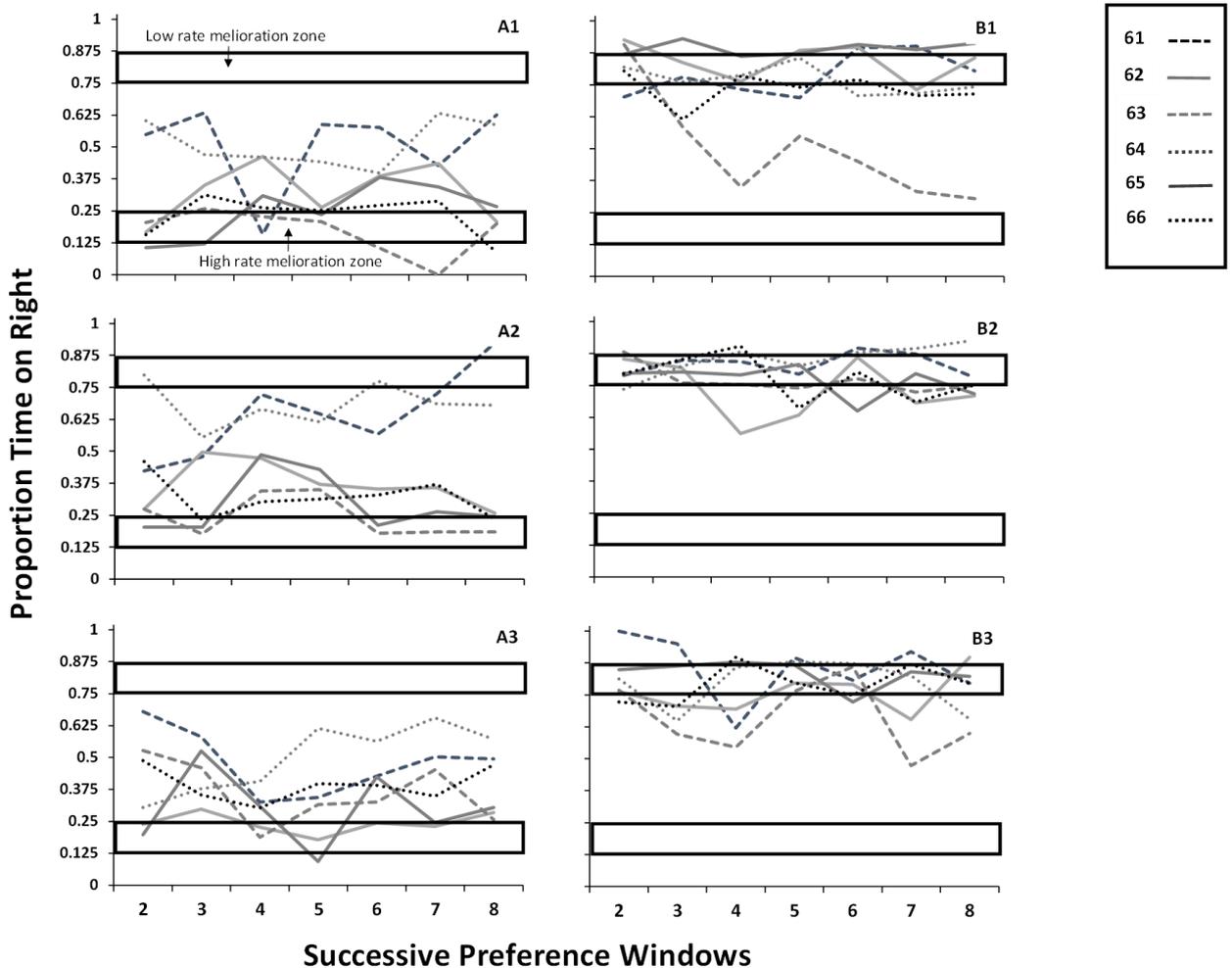


Figure 7.4 Mean proportion of time on the right key (y-axis) within successive preference windows 2-8 (x-axis) for individual pigeons as the mean of the last five sessions in A conditions (left panel) and B conditions (right panel). Horizontal rectangular boxes indicate optimal (high rate) and suboptimal (low rate) zones.

## DISCUSSION

We tested melioration theory by investigating whether within-session and across sessions choice systematically follows arranged local contingencies. Specifically, this study sought to replicate and extend the findings of Vaughan (1981) by (1) using six pigeons (2) analyzing both across sessions (global analyses) and within-session dependent measures (within-session analyses) and (3) repeating cycles of alternating conditions in which maximized reinforcer rates favored melioration (Condition A) or did not (Condition B). Overall, the results of this study partially support melioration theory (Herrnstein & Vaughan, 1980) and extend the suboptimal choice findings of Vaughan (1981).

The across sessions choice analyses (Figure 7.3) shows that all six pigeons shifted choice toward the optimal zone and away from the suboptimal zone in A conditions, and toward the suboptimal zone and away from the optimal zone in B conditions. Changes in allocation between Conditions A and B in these analyses may be expected given we calculated mean proportion of time on the right across the last five sessions of each block of 15 baseline sessions. Therefore, shifts in choice direction between A and B conditions suggest the pigeons had learned the local contingencies of each condition.

Although choice changed systematically across conditions for all pigeons, global choice fell *exactly* within optimal- or suboptimal zones in respective A and B conditions only for two of six pigeons in A conditions (Pigeon 63 in A1 and A2, and Pigeon 62 in A3) and three of six pigeons in B conditions (Pigeon 65 in B1 and B3, and Pigeons 61 and 64 in all B conditions). This may be compared with all pigeons in A conditions, and two of three pigeons in B conditions, for Vaughan's (1981) study. Therefore, the across sessions results of the present study generally support those of Vaughan. However, questions remain as to why all pigeons' behaviour was not precisely controlled by the prevailing contingencies.

The extent to which choice fell exactly within the optimal or suboptimal zone likely depended on the extent to which the pigeons could maintain learned local contingencies. For example, the within-session analyses of the present study (see Figure 7.4) show that when choice in any one preference window fell within either the optimal or suboptimal zone, choice remained toward that zone across subsequent windows for most pigeons in accordance the local contingencies. This suggests most pigeons could discriminate local reinforcer-rate differences. However, as observed in the across sessions analyses, choice across successive preference windows did not always remain exactly *within* a zone. This suggests choice did not always follow the local contingencies that led to equal local reinforcer rates.

The extent to which choice followed arranged local contingencies for optimal or suboptimal zones had an impact on overall obtained reinforcer rates in A and B conditions. For example, the across sessions analyses show that four of six pigeons (62, 63, 65 and 66) obtained a lower overall rate of reinforcement in B conditions compared with A conditions (Figure 7.2). The within-session analyses (Figure 7.4) shows this occurred because most pigeons (except Pigeon 63) allocated a greater proportion of time to the locally richer right key across successive preference windows in B conditions, despite the fact this decreased the overall rate of reinforcement in B conditions. Although this reduction was small, it was likely not larger due to optimal and suboptimal zones being narrow, or constrained. For

example, in Condition A, to access VI 20 VI 20 schedules of reinforcement, proportion of time on the right must fall precisely between 0.125 and 0.25 within any 4-min window. If the proportion of choice is instead between 0.26 and 0.74, the pigeons are exposed to VI 20 VI 60 schedules of reinforcement. The same applies in Condition B: if proportion of choice is not exactly within 0.75 to 0.875, the schedule of reinforcement will be VI 60 VI 20, not VI 60 VI 60. Therefore, although the difference in overall obtained reinforcer rates between A and B conditions is small, the lower overall obtained reinforcer rates in B conditions is evidence of a suboptimal choice outcome.

Although most of the pigeons obtained a higher overall reinforcer-rate in A conditions compared with B conditions, this did not occur for Pigeons 61 and 64 (see Figure 7.2). These pigeons obtained a lower overall rate of reinforcement across both conditions compared to the other pigeons. In addition, the across-sessions analysis (Figure 7.3) shows that reversals in choice direction between A and B conditions were less pronounced for Pigeons 61 and 64 compared with the other pigeons. The within-session analysis (Figure 7.4) shows this occurred due to Pigeons 61 and 64 allocating a greater proportion of time to the locally leaner right key across preference windows in A conditions compared with other pigeons, and subsequently to the locally richer right key in B conditions, in accordance with other pigeons. Therefore, a greater proportion of time was allocated to the right key for these two pigeons, irrespective either of condition or whether the right key was locally richer or leaner across successive preference windows. As this behaviour is not in accordance with the local contingencies of A conditions, the overall suboptimal choice observed for Pigeons 61 and 64 cannot be directly attributed to melioration.

It is possible choice by Pigeons 61 and 64 followed contingencies other than the arranged local contingencies when local reinforcer rates were not equal, but followed arranged local contingencies when local reinforcer rates were equal. For example, these pigeons may have learned the more general rule that ‘choosing right more often’ eventually resulted in equal local reinforcer rates in any condition. If so, it could be argued that behaviour remained under the control of discriminated equal local reinforcer rates, neither supporting nor fully opposing melioration theory. While this interpretation remains provisional, it offers one explanation for why choice remained close to the suboptimal zone in both A and B conditions for these pigeons.

It is also possible that a right-key bias overrode the arranged local contingencies in A conditions for Pigeons 61 and 64. The within-session analyses of Figure 7.4 shows relative time on the right key across successive 4-min preference windows. In A conditions,

the arranged local contingencies required the pigeons to allocate more time to the locally richer left key across preference windows to maximize overall reinforcer rates. Pigeons 61 and 64 allocated more time to the locally leaner *right* key in A conditions than the other pigeons. However, choice was only consistently toward the right key across preference windows for both pigeons in Condition A2. In Condition A1, choice across preference windows varied for Pigeon 61, and did not shift toward the right key until preference window 6 for Pigeon 64. In Condition A3, choice was more consistently toward the right key across preference windows for Pigeon 64. However, for Pigeon 61 this did not occur until preference window 5, after which proportion of time on the right remained below 0.60. Although the evidence of a right key bias is inconclusive, it is possible. This raises the question of whether the melioration equation could be improved by the inclusion of a bias parameter.

There is research to support the possibility that Pigeons 61 and 64 had difficulty discriminating unequal differences in local reinforcer rates between alternatives in both A and B conditions. Davison (1990) investigated the role of discriminated local reinforcer rates within the framework of melioration theory. Davison proposed ‘melioration threshold theory’ to explain studies in which choice by pigeons did not closely follow a locally richer alternative (Baum, 1979; Myers & Myers, 1977; Taylor & Davison, 1983; Wearden & Burgess, 1982). Davison modeled exponential schedules of concurrent VI VI reinforcement that ranged from VI 56 s VI 56 s to VI 29 s VI 556 s. He found that as the difference in local reinforcer rates between alternatives became progressively smaller, feedback functions became flatter. That is, simulated choice ratios progressively failed to follow modeled reinforcer ratios. In referring to data from studies with pigeons, Davison showed these flat feedback functions would present as increasingly indifferent or random choice. Davison concluded that when small differences in local reinforcer rates become difficult to discriminate, the ‘threshold’ required for melioration is breached, increasing the likelihood of indifferent choice.

Supporting the findings of Davison (1990), Vaughan (1981) attributed the random choice demonstrated by one pigeon in his study to a failure to discriminate differences in local reinforcer rates. Specifically, while mean global choice ultimately fell exclusively within optimal- and suboptimal zones in those conditions, this initially was observed for only two of three pigeons in Condition B of Vaughan’s study. Shifting choice to the suboptimal zone in Condition B for that one pigeon required increasing the difference in local reinforcer rates between alternatives. As a result, that pigeon began to allocate a greater

proportion of time to the right key during Condition B in accordance with the local contingencies. This suggests Vaughan made the arranged local contingencies more discriminable for this one pigeon. As suggested by Davison, the extent to which local reinforcer rate differences are discriminated likely depends upon the extent to which the melioration threshold is maintained and not breached. Vaughan's results suggest there are likely to be individual differences in where that threshold lies. Similarly, the findings of the present study suggest individual differences in the ability of Pigeons 61 and 64 to discriminate local reinforcer-rate differences compared with the other four pigeons.

The attenuated control by arranged local contingencies for Pigeons 61 and 64 relative to the other pigeons might have resulted in lower overall response rates for these pigeons compared with the other four pigeons (see Table 7.2). For example, relatively lower overall response rates for Pigeons 61 and 64 may be due to a right-key bias, or due to difficulty in discriminating unequal local reinforcer rates. It is also possible that lower overall response rate impacted the ability of these pigeons to discriminate local contingencies. We are not aware of any research that suggests difficulty with contingency discriminability impacts overall response rates, or vice versa. Nevertheless, the finding that overall response rate was consistently lower for Pigeon 61 and 64 than the other pigeons across all conditions is of interest and might play some role in their patterns of behaviour.

At present, melioration theory does not provide a parameter that measures the extent to which differences in local reinforcer rates between alternatives are discriminated. This may be addressed in development of quantitative analyses designed to integrate melioration theory and contingency discriminability theory (Davison & Jenkins, 1985). Taken together, the findings of Davison (1990) and Vaughan (1981) suggest that if differences between local reinforcer rates had been increased in the present study, Pigeons 61 and 64 may have allocated progressively more time to the locally richer left key during A conditions in accordance with the local contingencies.

Although we did not increase the size of local reinforcer rate differences for Pigeons 61 and 64, the findings of Vaughan (1981) and Davison (1990) suggest that, for some individuals, discriminating local contingencies may require larger differences in local reinforcer rates. Moreover, if melioration does contribute to overall suboptimal choice there is benefit in investigating whether progressively smaller local reinforcer-rate differences disrupt choice already following local contingencies. Future studies might also investigate the extent to which suboptimal choice may be subverted if local reinforcer rates become difficult to discriminate, and overall reinforcer rates are simultaneously signaled. For

example, in a study with pigeons Heyman and Tanz (1995) made greater overall reinforcer rates contingent on choice not following local reinforcer rates. In Experiment 1, if choice closely followed local reinforcer rates, responses were not eligible for reinforcement. Therefore, overall reinforcer-rate decreased, resulting in suboptimal choice. In Experiment 2, a blue light was presented whenever responses were not immediately eligible for reinforcement and a white light was presented whenever responses were eligible for reinforcement. Heyman and Tanz found when the lights were absent, choice followed local-reinforcer rates, leading to suboptimal choice. When the lights were present, the extent to which choice followed local reinforcer rates decreased as a function of increased overall reinforcer rate. Therefore, the results of the studies by Davison (1990) and Heyman and Tanz suggest that the extent to which melioration occurs may depend on the ability of organisms to both discriminate local reinforcer rates and *not discriminate* overall reinforcer rates.

Alsop (2018) suggests that a limitation of the present study is that rather than arrange independent concurrent schedules of reinforcement as Vaughan (1981) did, we arranged dependent concurrent schedules of reinforcement. Alsop (2018) suggests that melioration is not possible if reinforcers are arranged across concurrent schedules dependently. However, the operation of dependent scheduling in a procedure that schedules *local* reinforcer rates differs from that of standard concurrent schedules, which schedule *overall* reinforcer rates. In the present study, once one VI schedule set up a reinforcer, no further reinforcers could be set up until the first reinforcer had been obtained. This differs from the form of dependent scheduling common to standard concurrent schedules which ensures a prearranged distribution of reinforcers between alternatives. In the present study, because the VI schedule on the alternate key is already stopped when the pigeon is responding on one key, the only time the dependency constraint can operate is if a reinforcer sets up on one key in the middle of a switch – that is, after the last response to that key and before the first response to the other key. If that happens, responses after the switch will not produce reinforcers. If melioration is correct, the pigeon will switch back to the original key, as the local reinforcer rate on the switched-to key is now zero. The pigeon will collect the reinforcer, and the schedule that controls the local rate on the original key will restart. As pigeons typically repeat responses on the same key more often than they switch between keys, such an event will be relatively rare. Therefore, local reinforcer rates on both keys may be slightly less than those programmed, but will not result in a wholesale distortion of local rates that precludes the possibility of melioration occurring.

It has been suggested that melioration may underlie suboptimal, and maladaptive, choice in humans (Herrnstein & Prelec, 1991; Heyman, 1996; 2013; Rachlin, 1997). However, although these studies support the overall predictions of melioration theory, they do not demonstrate the extent to which choice remains systematically under the control of local contingencies. Vaughan's (1981) seminal study attempted to do just that, but his results are not fully compelling due to a lack of demonstrated experimental control across repeated condition changes. The present study addresses these limitations, and extends Vaughan's study by showing overall choice shifts systematically with repeated changes in local contingencies. The results of the present study also suggest that proportioning choice across successive preference windows will be maintained if the proportion of choice results in equal local reinforcer rates. These proportions will be maintained even if local choice leads to overall suboptimal choice outcomes.

This study facilitates the more confident use of this and similar procedures in future studies designed to further investigate suboptimal choice. If melioration contributes to suboptimal choice, there is much to be gained from further investigation of the extent to which increased differences in local reinforcer rates lead to better discrimination of those rates. If larger differences in local reinforcer rates result in choice closely following arranged local contingencies, this offers an opportunity to investigate ways in which a stable melioration process may be able to be subverted, resulting in more optimal choice outcomes. Such findings could provide insight into addressing suboptimal choice as it relates to addiction.

## CHAPTER 8.

### DOES A NEGATIVE DISCRIMINATIVE STIMULUS PUNISH SUBOPTIMAL CHOICE IN PIGEONS?

In this chapter we use the procedure developed by Vaughan (1981), and further developed in the study described in Chapter 7, to investigate whether a negative discriminative stimulus signaling an absence of food reinforcers will subvert suboptimal choice in pigeons, leading to more optimal outcomes.

As discussed in Chapter 6, it has been suggested that melioration may underlie suboptimal choice in humans, and therefore may account for behaviour described as addictive (Herrnstein & Prelec, 1991; Heyman, 1996; 2013; Rachlin, 1997). However, while many studies have investigated the association between suboptimal, or maladaptive, choice, and *sensitivity* to reward and punishment in humans (e.g. Jonker, Ostafin, Glashouwer, van Hemel-Ruiter, & de Jong, 2014; Morris et al., 2016; Mestre-Bach et al., 2016), relatively few studies have attempted to directly punish suboptimal choice in humans or animals. Those that have been conducted suggest that punishment may be a powerful tool for reducing suboptimal choice. For example, Shimp, Mitchell, Beas, Bizon, and Setlow (2015) showed that rats' choice of high-risk rewards is reduced when such choices also produce shocks, and that this effect of shock increases as the magnitude of the shock increases. In contrast, increasing the magnitude of the high-risk reward in the absence of punishment did not increase the proportion of high-risk choice. Similarly, Orsini et al. (2016) showed that shock punishment decreased choice for a smaller, sooner reward and shifted choice toward a larger, later reward in both male and female rats. Therefore, these studies show that when sub-optimal choice is punished, choice can become more optimal.

Vanderschuren, Minnaard, Smeets, and Lesscher (2017) reviewed a range of studies in which aversive stimuli were used to decrease rates of substance-seeking behaviour in animals. The aversive stimuli included quinine, lithium chloride, histamine, and shock, used to punish behaviour including cocaine- and alcohol self-administration. Of the punishers, quinine, lithium chloride, and histamine were found to punish immediate, but not long-term aspects of substance-seeking behaviour. Shock was found to generate fear and other conditioned emotional responses, but unlike quinine or lithium chloride, several studies found animals will endure mild shock to obtain cocaine or alcohol. Vanderschuren et al. noted that in the natural world, substance abuse in humans is usually not punished

immediately, and that delayed punishment in animals has been found to be less effective in decreasing cocaine self-administration. Therefore, they suggested the immediacy of timing between behaviour and punisher should be incorporated into the design of animal models of addictive behaviour.

Powell and Azrin (1968) attempted to punish cigarette smoking in men using a cigarette case that delivered a shock when opened. Although the rate of smoking decreased as a function of the intensity of the shock, smoking returned to previous levels when the shock was discontinued. This might be expected, given that punishment was contingent on cigarette-case opening, rather than smoking. Similarly, Pelloux, Murray, and Everitt (2015) used rats to show that substance-*seeking* behaviour may be easier to punish than substance-*taking* behaviour. More research is required to determine whether punishing substance-seeking behaviour will lead to a decrease in substance-taking in humans or animals.

The studies discussed above attempted to punish suboptimal choice arising from substance-seeking and substance-taking behaviour in animals and humans. The punishers used were drugs, chemical compounds, or shock. Clearly, there is a gap in the research regarding the use of non-harmful punishing stimuli to decrease suboptimal choice. That is, the use of stimuli that do not produce unpleasant physical side-effects, or conditioned emotional responses. As demonstrated in the studies presented throughout this thesis, it might be possible to punish suboptimal choice using a contingent S- stimulus previously associated with an absence of reinforcers. In this fourth, and final study, we are interested in whether suboptimal choice following arranged local contingencies will shift direction if that choice also produces a contingent S- stimulus. We also want to know if previously suboptimal choice that shifts through punishment by a contingent S- stimulus will become more optimal as a result. That is, will pigeons receive a relatively greater reinforcer rate overall if the melioration process is punished?

## METHOD

### Subjects

Subjects were six experimentally experienced homing pigeons, numbered 61 to 66. The pigeons had previous experience with two-alternative concurrent-choice schedules. In addition, the six pigeons were previously exposed to the procedure developed by Vaughan (1981) and conducted in Chapter 6. Pigeons were individually housed within a pigeon colony room and maintained at  $85\% \pm 10g$  of their free-feeding body weights using post-session supplementary feeding of mixed grain. Water and grit were available at all times.

The colony room lighting was switched on at 00:00 and off at 16:00 daily. Experimental sessions began at 01:00 with a 60 min blackout. No personnel entered the room during sessions.

## **Apparatus**

The pigeons' home cages served as the experimental chambers. The cages measured 380 mm high, 380 mm wide and 380 mm deep. Three walls were constructed of sheet metal while the floor, ceiling, and front wall were metal bars. In each chamber, two wooden perches were mounted 60 mm above the floor. One perch was parallel to, and 90 mm from, the front wall while the other was parallel to, and 90 mm from, the right wall. The operant panel was located on the right wall and consisted of three 20 mm diameter translucent keys set 220 mm above the wooden perches and 85 mm apart, center to center. Only the left and right keys were used in this experiment. The center key was not used. The left and right keys could be trans-illuminated by LEDs located behind each key with red, green or white lights. Pecks to a lit key exceeding 0.1 N were registered as responses. The magazine aperture was located 100 mm below the center key and measured 50 mm high, 50 mm wide, and 40 mm deep. A hopper containing wheat was situated behind the magazine. During hopper presentations, all key lights were turned off and the hopper was raised and illuminated for 2 s. We excluded hopper time from all analyses. All experimental events were arranged and recorded by an IBM® PC-compatible computer running MED-PC IV® software which was located in an adjacent room.

## **Procedure**

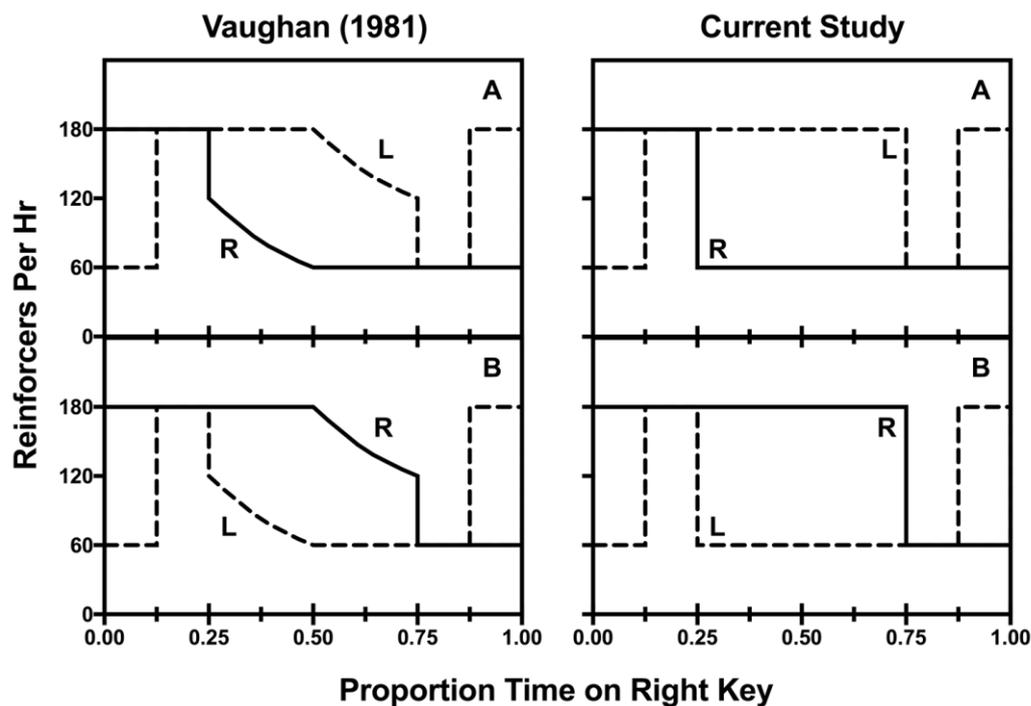
As the pigeons in this study were experimentally experienced, experimental conditions were initiated from the outset of the experiment. The procedure combined the S+ and S- discrimination training (DT) procedure used in the punishment studies of Chapters 4 and 5 of this thesis with Conditions A and B of the suboptimal choice procedure developed by Vaughan (1981). Therefore, we arranged four consecutive conditions: Conditions A, B, DT, and a punishment test (PT). As local choice contingencies led to suboptimal choice only in B conditions in Vaughan's procedure, in punishment tests for the present study we tested the effect of the punisher within a repetition of Condition B.

We repeated the four conditions three times, comprising three complete cycles. Each cycle contained the same conditions, and the conditions ran in the same order for all three cycles. The order of the conditions within each cycle was: A, B, DT, PT. Maintaining this

order of conditions in each cycle allowed us to a) compare choice between Condition A and B (as per Vaughan; 1981, and the previous study; see Chapter 7); and b) assess the effect of the punisher by comparing punished suboptimal choice (Condition PT) with non-punished suboptimal choice (Condition B). As there were three cycles, each condition was labelled with the letter of the condition first, followed by the number of the cycle. Therefore, the conditions are: A1, B1, DT1, PT1, A2, B2, DT2, PT2, A3, B3, DT3, and PT3.

**A and B conditions.** Each session in Conditions A and B was 30 min in duration. The number of 30-min sessions per condition varied until the primary dependent measure of proportion of time allocation on the right key was stable across five sessions for all six pigeons. The number of sessions per cycle was 30, 40, and 35 for Conditions A1, A2 and A3, respectively; and 25, 18, and 25 for Conditions B1, B2 and B3, respectively. The procedure for Conditions A and B was identical to that used in the study reported in Chapter 7. A full explanation of this procedure, including an explanation of *optimal and suboptimal zones*, is also provided in Chapter 7. As a reminder, Figure 8.1 shows programmed hourly rates of reinforcement in Condition A and Condition B for left key and right key responses as a function of proportion of time on the right key in a previous preference window. Vaughan's (1981) procedure is shown on the left, and the procedure for A and B conditions in the present study is shown on the right. The ratio of left and right key reinforcers per min for choice that fell outside of optimal or suboptimal zones was 1:3 (A conditions) or 3:1 (B conditions). In comparison, the ratio of left and right key reinforcers per min for choice that fell within the optimal zone (0.125 to 0.25) or suboptimal zone (0.75 to 0.875), was 3:3 and 1:1, respectively.

**DT conditions.** In the DT condition, each session was 30-min in duration. There were 20 sessions in total. We arranged a two-component multiple concurrent schedule. Each component was 60-s in duration, followed by a 10-s inter-component interval (ICI). Component 1 always ran first, followed by Component 2, after which the components alternated. In Component 1, left and right keys were lit green. Key pecks produced food reinforcers according to concurrent VI 20 VI 20 schedules. In Component 2, left and right keys were lit red. Key pecks to either key were never reinforced (see Chapter 4).



*Figure 8.1* Diagram of programmed hourly rates of reinforcement (y-axis) in Condition A (top panel) and Condition B (bottom panel) for left key (dashed black line, L) and right key (solid black line, R) responses as a function of proportion of time on the right key (x-axis) in a previous preference window, for Vaughan's (1981) procedure (left) and the procedure for Cycles 1 and 2 of the current study (right). Choice is predicted to shift in a subsequent preference window when proportion of time allocated to the right key in a current window is between 0.26 to 0.74. Equal rates of reinforcement for left and right keys occur when proportion of time allocated to the right key is either 0.125 to 0.25 or 0.75 to 0.875.

**PT conditions.** In the PT condition, each session was 30-min in duration. There were five sessions in total. The procedure for PT conditions was exactly the same as for B conditions, with the exception that every time choice across a previous 30-s time span fell within the suboptimal zone (see Figure 8.1), key pecks to either the left or right key across the next 30-s time span produced a superimposed red keylight for 1.5-s, according to a VR 5 schedule. That is, every five key pecks on mean on any one key resulted in the red keylight replacing the green keylight for 1.5-s for that key. As a red keylight had been trained as a negative discriminative stimulus in the previous DT condition, the red keylight therefore functioned as a contingent S- stimulus in the PT condition. If suboptimal choice is able to be punished by the contingent S- stimulus, proportion of time on the right should be less in PT conditions compared to B conditions. That is, choice should be relatively less toward the suboptimal zone in PT conditions, compared with B conditions.

## RESULTS

**Response and Reinforcer rates.** Table 8.1 shows mean overall response rates per minute across the last five sessions of each condition, in each cycle, for individual pigeons. The far right-hand column presents these rates averaged across all conditions. Mean response rate was considerably greater for Pigeon 63, compared with the other five pigeons. However, there was no systematic difference in response rate between A, B and PT conditions, across all three cycles, for all pigeons. Across all DT conditions, the proportion of S+ to S- responses (i.e. the Discrimination Index) ranged from 0.89 (Pigeon 65, Cycle 1) to 1.00 (Pigeon 61, Cycle 3). Therefore, rates of responding in the S+ component were substantially higher than in the S- component. This demonstrates strong discrimination between the two components and shows that the S- discriminative stimulus (a red keylight) was established as a signal for the absence of reinforcers.

Figure 8.2 shows mean obtained overall reinforcer rates from the last five sessions of Conditions A1, A2 and A3, and similarly across replications of Condition B and Condition PT, for individual pigeons. We found no reliable difference in overall reinforcer rates between conditions. However, although Pigeons 61, 62 and 64 obtained a higher overall rate of reinforcers across PT conditions than B conditions. These differences indicate that choice was relatively more optimal when a contingent S- stimulus punished suboptimal choice (Condition PT) than Condition B, for those three pigeons.

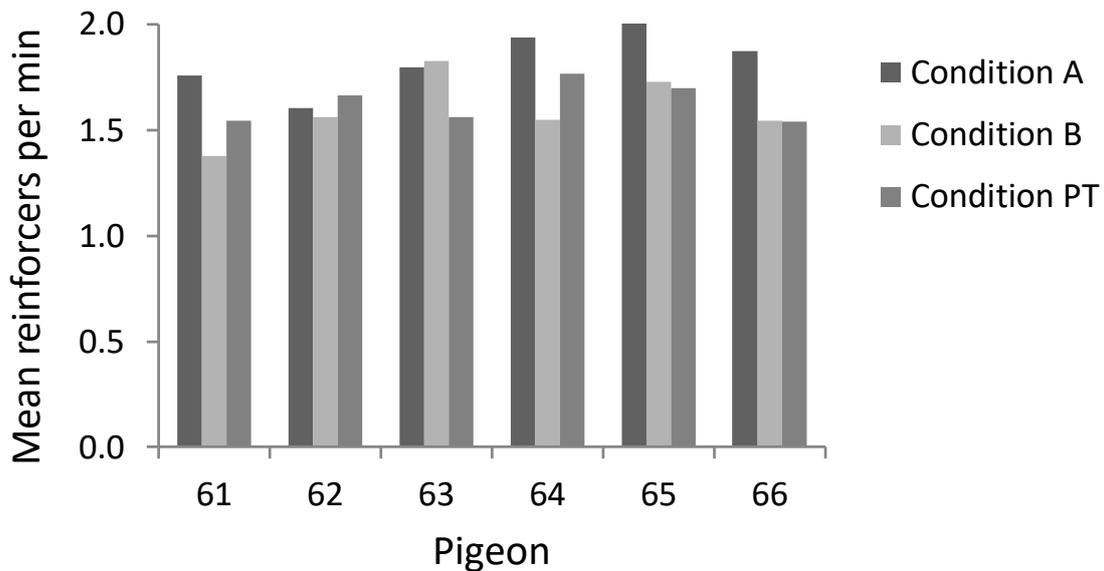
In addition, overall reinforcer rates were higher in Condition A than Condition B for five of six pigeons (Pigeon 63 was an exception); and higher in Condition A than Condition PT for five of six pigeons (Pigeon 62 was an exception). Therefore, a majority of pigeons obtained more overall reinforcers in A conditions in which melioration led to optimal choice, than in B conditions in which melioration led to suboptimal choice. These findings support those of Vaughan (1981) and the extension of Vaughan's study (see Chapter 7).

**Across sessions choice analysis.** Figure 8.3 shows proportion of time on the right key for individual pigeons as the mean of the last five sessions of Conditions A, B and PT, for each of the three cycles. Despite differences in levels of time allocation, choice tended to be toward the optimal zone in A conditions, and toward the suboptimal zone in B conditions, for most pigeons. Pigeon 61 in Cycles 1 and 2 was an exception. In PT conditions, if suboptimal choice is punished by the contingent S- stimulus, proportion of time on the right should be less than that of B conditions. That is, overall choice should shift away from the

Table 8.1

*Mean Overall Response rates per Minute across the Last Five Sessions in each Condition, and each Cycle, for Individual Pigeons. The Far right-hand column presents these Rates Averaged across all Conditions.*

S	CONDITION												MEAN
	A1	B1	DT1 S+/S-	PT1	A2	B2	DT2 S+/S-	PT2	A3	B3	DT3 S+ / S-	PT3	
61	49.41	61.45	55.84/0.03	57.96	62.43	62.75	49.41/0.42	62.33	57.05	61.82	57.30/0.01	65.99	59.40
62	42.11	57.19	38.51/0.33	52.43	42.23	48.57	30.38/0.27	51.71	39.61	51.87	37.31/0.53	50.73	48.22
63	89.37	80.04	65.22/0.03	82.43	82.95	79.04	72.10/1.13	86.70	74.25	76.39	78.26/1.69	76.73	81.40
64	52.01	59.43	46.86/0.05	50.81	60.30	49.01	51.91/0.05	63.07	49.82	44.37	54.55/0.08	48.07	53.60
65	59.28	58.23	59.36/7.62	61.57	67.13	68.45	58.29/4.18	64.65	65.65	76.97	64.54/6.10	68.88	65.24
66	74.85	54.91	59.32/0.08	66.63	68.96	58.85	62.13/1.40	60.21	57.63	59.66	50.87/0.53	58.37	62.71



*Figure 8.2 Overall obtained reinforcer-rate as the mean of the last five sessions of Condition A1, A2 and A3 (black bars); the mean of the last five sessions of Condition B1, B2 and B3 (light grey bars), and the mean of the last five sessions of Condition PT1, PT2, and PT3 (dark grey bars), for individual pigeons.*

suboptimal zone, compared with B conditions. This occurred in all three cycles for Pigeon 64, for two of three cycles for Pigeons 62 and 63, and in one cycle for Pigeons 61 and 65. In comparison, choice did not shift away from the suboptimal zone in PT conditions compared with B conditions across all three cycles for Pigeon 66. Overall, choice shifted

away from the suboptimal zone in PT conditions compared with B conditions, for 11 of 18 comparisons.

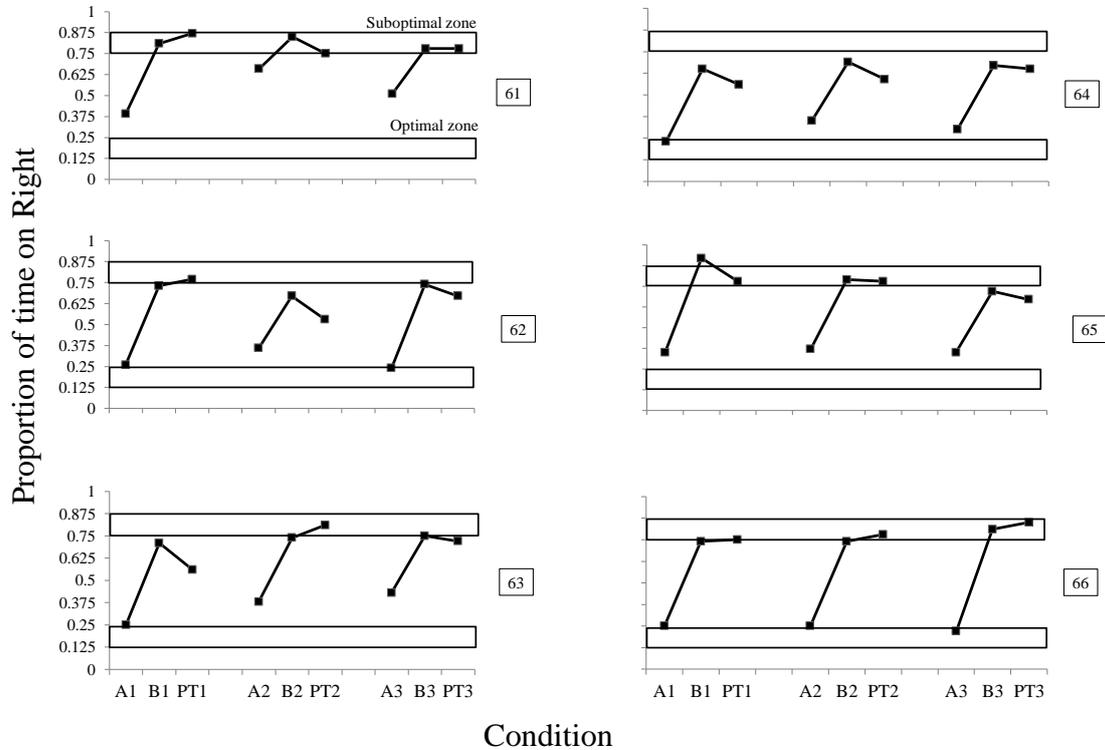


Figure 8.3 Proportion of time on the right key (y-axis) for individual pigeons, averaged across the last five sessions of conditions A1, B1, PT1, A2, B2, PT2, and A3, B3, PT3 (x-axis). Horizontal rectangular boxes indicate optimal and suboptimal zones.

**Within-session choice analysis.** Figure 8.4 shows proportion of time on the right key in each successive preference window as mean of the last five sessions in A, B and PT conditions, for each cycle, for individual pigeons. Window 1 was omitted because no reinforcement contingencies were arranged in a previous window within that session. Despite choice being more variable for some pigeons than others, choice across successive preference windows was generally towards, or remained within, either the optimal or suboptimal zone in A and B conditions, respectively, across all cycles (Pigeon 64 in Cycle 2, and Pigeon 66 in Cycle 3, were exceptions). In PT conditions, if suboptimal choice was punished by the contingent S- stimulus, less time should be allocated to the right key within, and across, successive preference windows, compared with B conditions. This occurred in two of three cycles for Pigeons 63, 64 and 65, and in one cycle for Pigeons 61 and 62. In

comparison, choice across successive preference windows in all PT conditions closely followed that of B conditions for Pigeon 66.

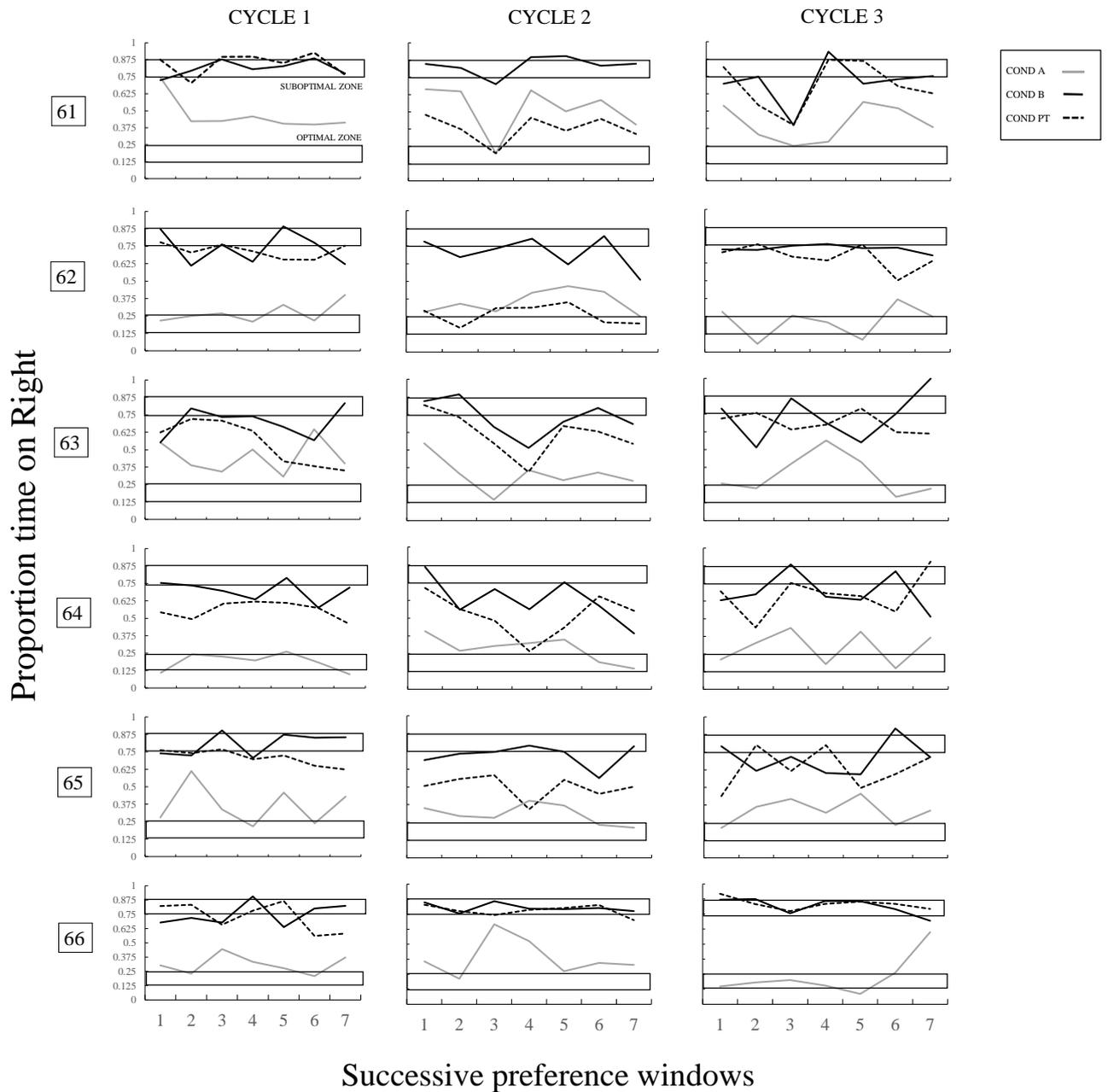


Figure 8.4 Proportion of time on the right key (y-axis) within each successive preference window 1-7 (x-axis), for each of three cycles, for individual pigeons, as the mean of the last five sessions in A conditions (grey line), B conditions (black line) and PT conditions (black dashed line). Horizontal rectangular boxes indicate optimal and suboptimal zones.

## DISCUSSION

In Chapter 7 we showed that within- and across-session choice generally follows arranged local contingencies in support of melioration theory (Herrnstein & Vaughan, 1980; Vaughan, 1981). In Vaughan's (1981) procedure, local choice contingencies led to optimal and suboptimal choice in A and B conditions, respectively. In the present study, we replicated this procedure, but added a punishment test condition in which suboptimal choice in any one 30-s window produced a contingent S- stimulus on a VR 5 schedule for the next 30 s. We then compared choice between the last five sessions of all B and PT conditions, across repeated cycles. If suboptimal choice is punished by the contingent S- stimulus, less time should be allocated to the right key within, and across, successive preference windows in PT conditions compared with B conditions.

Across all pigeons and cycles, choice shifted away from the suboptimal zone in PT conditions compared with B conditions for 11 of 18 comparisons. Therefore, the results of the present study do not provide a reliable effect of the contingent S- stimulus using a binomial sign test ( $p = .12$ ). In addition, although three of six pigeons obtained a higher overall reinforcer rate across all PT conditions compared with across all B conditions (see Figure 8.2), it is not possible to be sure choice was more optimal for these pigeons as a result of the contingent S- stimulus shifting choice away from the suboptimal zone in some PT conditions.

Although choice across successive preference windows shifted away from the suboptimal zone in at least one PT condition for five of six pigeons, for four of those five pigeons choice did not *systematically* shift away from the suboptimal zone in PT conditions compared with B conditions, across repeated cycles. Pigeon 64 was the exception (see Figure 8.3). Therefore, while the results of this study suggest that a contingent S- stimulus *might* have the capacity to shift suboptimal choice, more research is required. To be reliable, the impact of a contingent S- stimulus on suboptimal choice should be consistent across repeated reversals of punished and unpunished conditions, for a majority of subjects.

Analysis of overall obtained reinforcer rates shows that Pigeon 62 achieved more optimal choice across PT conditions than A conditions (see Figure 8.2). This suggests that choice by Pigeon 62 did not always follow the arranged local contingencies. This within-session analysis shows this likely occurred because, in Cycle 2, choice did not reach the optimal zone in any one of the successive preference windows during Condition A for

Pigeon 62, whereas choice reached the optimal zone in three separate preference windows during the PT condition.

Of the pigeons that did not obtain a higher overall rate of reinforcers across PT conditions compared with B conditions, Pigeon 63 also obtained a slightly higher overall rate of reinforcers in B conditions compared with A conditions (see Figure 8.2). It is possible that Pigeon 63 had a right-key bias, or had difficulty in discriminating unequal differences in local reinforcer rates between alternatives (see the Discussion in Chapter 7; and in Vaughan, 1981). This suggestion is supported by the within-session analysis which shows choice was more varied across successive preference windows in both A and B conditions for Pigeon 63, compared with the other five pigeons. If choice did not closely follow the arranged local contingencies in both A and B conditions, this might explain why Pigeon 63 did not obtain a relatively higher rate of reinforcers in PT conditions compared with B conditions.

It is difficult to understand why Pigeon 65 did not obtain a relatively higher reinforcer rate across PT conditions compared with B conditions, given that choice across successive preference windows in those conditions closely resembles that of Pigeon 61. One possibility is that, during Cycle 2, choice during the PT condition was further toward the optimal zone, or more consistently within the optimal zone, for Pigeon 61, compared with Pigeon 65.

It is possible we did not see a consistent impact, or lack of impact, of the contingent S- stimulus on choice across repeated cycles for four of six pigeons for procedural reasons. For example, each PT condition was comprised of only five sessions. The decision to select five sessions for PT conditions was made to allow us to observe the immediate impact of the contingent S- stimulus on choice. However, five sessions is considerably fewer than the number of sessions used in B conditions. To recap, we ran 25, 18, and 25 sessions for Conditions B1, B2 and B3, respectively, with the criterion that relative time on the right remained stable across five consecutive sessions. The difference in session numbers between B and PT conditions means we cannot be sure that the impact of the contingent S- stimulus was not consistent across repeated cycles in PT conditions due to the relatively shorter duration of PT conditions.

Another possibility is that we did not see a consistent impact, or non-impact, of the contingent S- stimulus on choice for four of six pigeons because relative time on the right was averaged across all five sessions of each PT condition, for both across- and within-session choice analyses. It is possible the ability of the contingent S- stimulus to signal an

absence of reinforcers may have been stronger in the first few sessions of some PT conditions, then weakened across remaining sessions. This suggestion is supported by the findings of the study described in Chapter 4. In Condition 5 of that study, a block of only two punishment test sessions followed each block of discrimination training. Results showed the contingent S- stimulus shifted choice away from the punished alternative on all punishment tests for four of five pigeons (see Figure 2.7). However, the arranged reinforcer contingencies between baseline and punishment tests sessions in Condition 5 of the study described in Chapter 4 were the same. In comparison, the arranged reinforcer contingencies in the present study differed between DT and PT conditions, and also differed within sessions in PT conditions. Therefore, reducing the number of sessions in PT conditions to better observe the punishing impact of the contingent S- stimulus might also have the effect of reducing contingency discriminability in PT conditions.

One solution might be to arrange for discrimination training to occur immediately prior to the beginning of every session of A, B and PT conditions, rather than arranging a long standalone DT condition comprised of 20 sessions. For example, in other studies by this author (not included in this thesis), the capacity of a contingent S- stimulus to punish choice was found to increase when a short period of S+/S- discrimination training immediately preceded a choice task, compared to when S+/S- training occurred in a different session or condition.

Overall, the findings of this study are unreliable, but suggest that a contingent S- stimulus might punish suboptimal choice in pigeons. The contingent S- stimulus clearly shifted choice away from the suboptimal zone in PT conditions compared with B conditions across all cycles for Pigeon 64, and in one or more cycles for Pigeons 61, 62 and 65.

These findings provide a starting point for future studies to build upon regarding the use of a contingent S- stimulus to shift suboptimal choice in animals.

## CHAPTER 9

### GENERAL DISCUSSION

This thesis investigated whether a contingent S- stimulus will punish responding maintained by positive reinforcers on simple schedules in pigeons and humans, and shift choice between equally-reinforced alternatives in pigeons. We also investigated whether a contingent S- stimulus will shift suboptimal choice in pigeons, with a view to informing research designed to reduce suboptimal choice in humans. The results of Study 1 show that when a negative discriminative stimulus is made a consequence, the presence of the signal alone is enough to decrease responding maintained by access to positive reinforcers, and to shift choice in pigeons. Despite individual variability in the extent of decrease in S+ response rate, a contingent S- stimulus punished responding for positive reinforcers on most punishment tests for most pigeons. A contingent S- stimulus also shifted choice away from a punished alternative and towards an unpunished alternative, despite an equal reinforcer ratio between concurrent alternatives. Study 1 featured two consecutive 30-min punishment test sessions in which responses or choice produced the contingent S- stimulus. After each block of two punishment sessions, the pigeons were returned to ten sessions of baseline discrimination training. This raises the question of whether regular periods of discrimination training are required for a contingent S- stimulus to retain its aversive properties.

In comparison to the findings of Study 1, a contingent S- stimulus did not reliably decrease S+ response rate in Study 2 with humans. This result may have been impacted by procedural elements identified in the Discussion of Study 2 (see Chapter 5), including brevity of training, the use of reinforcer schedules of high frequency but low magnitude, and the use of an intermittent schedule of punishment. In addition, of the 29 participants who contributed to the human study, only 11 learned to discriminate between S+ and S- in the short training time provided. Clearly, if a single session of training is used to develop S- as a potential punisher in humans, it must be long enough for discrimination learning to be acquired. Further, the findings of Study 1 with pigeons suggest that a contingent S- stimulus will shift choice. Therefore, future research might investigate the potential for a contingent S- stimulus to shift choice in humans.

The findings of Study 3 supported Melioration Theory, and provided a procedural platform for Study 4, in which we investigated the potential of a contingent S- stimulus to shift choice shown to lead to overall suboptimal outcomes. The results of Study 4 suggest

that a contingent S- stimulus might shift choice that is following contingencies shown to lead to suboptimal outcomes in pigeons. However, more research is required to determine whether a contingent S- stimulus will reliably shift suboptimal choice in animals.

If we look back to the literature review in Chapter 2, we see that the efficacy of a punishing stimulus is impacted by variables including whether or not an alternative reinforced behaviour is available, the frequency and intensity of the punishing stimulus, and the schedule on which contingent reinforcers and punishers are delivered. If we consider these variables in light of the studies conducted for this thesis, we find that a measurable alternative reinforced behaviour was not available in punished conditions in which pigeons and humans responded on simple schedules. Although there was a choice condition in the first study with pigeons, that condition presented concurrently available alternatives with each alternative signaled by the same S+ discriminative stimulus (see Table 2.1).

If a different reinforced alternative response had been available in the studies of this thesis, as occurs in the natural world, decreases in S+ response rate, or shifts in S+ choice, may have been greater. In the fourth study, an alternative reinforced behaviour was, in theory, available as the pigeons could choose to *not* follow arranged local choice contingencies, and receive a higher overall rate of reinforcement as a result. However, as this contingency was not signaled, it occurred only to the extent that the pigeons were able to discriminate that a change in choice could produce a more optimal outcome. Studies of choice between discriminative stimuli arranging signaled and unsignaled reinforcers show that animals prefer stimuli arranging signaled over unsignaled schedules of reinforcers (Lewis, Lewin, Muehleisen, & Stoyak, 1974; Moon & Lewis, 1975; also see Jenkins & Boakes, 1973). In addition, signaling the availability of reinforcers for an alternative behaviour can reduce resistance to change of target behaviour, compared to the absence of such a signal (Bland, Bai, Fullerton & Podlesnik, 2016). As preference and resistance to change tend to be positively correlated (Grace, Bedell, & Nevin, 2002; Nevin, 1979; Nevin & Grace, 2000), we can surmise that if alternative optimal choice contingencies *had* been signaled in the fourth study of this thesis, punished suboptimal choice might have decreased to a greater extent. Therefore, one area for further study is the extent to which the availability of an alternative reinforced response would increase the efficacy of a contingent S- stimulus to both punish response rate and suboptimal choice.

The studies reviewed in Chapter 2 also showed that punishment efficacy might depend less on the type of punishing stimulus used, and more on the magnitude, or schedule of delivery, of either reinforcers or punishers. For example, Filby and Appel (1966) showed

that increasing reinforcer frequency can lead to decreases in the efficacy of a punishing stimulus, while Azrin (1959; 1960) showed that increasing the intensity of a punishing stimulus decreased the rate of key-pecking for food in pigeons when food was delivered on a variable-interval (VI) schedule.

The studies that comprise this thesis add an extra layer of complexity to the impact of reinforcer and punisher schedules on punishment efficacy, because the contingent S- stimulus used in each study signals the absence of a *particular* rate of reinforcers (i.e. reinforcers arranged on a VI 15-s or VI 20-s schedule). Given this, we might expect that a contingent S- stimulus signaling the absence of a relatively higher rate of reinforcers would be more aversive, and have greater efficacy as a punisher, than a contingent S- stimulus signaling the absence of a relatively lower rate of reinforcers. Yet, according to Filby and Appel (1966) the presence of a S+ stimulus signaling a relatively higher rate of reinforcers will also *decrease* the efficacy of a punishing stimulus. The conundrum then, is that if the positive discriminative stimulus (S+) used in the studies of this thesis signaled a higher rate of reinforcers (for example, VI 5-s as opposed to VI-15-s), would the punishing efficacy of the contingent S- stimulus be enhanced or attenuated? Further research is required to determine the extent to which the efficacy of a contingent S- stimulus changes as a function of changes in the rate of positive reinforcers for S+ responding.

Schedules of punisher delivery may also impact the efficacy of a punishing stimulus. For example, the studies discussed in Chapter 2 show that if every reinforced response is punished (continuous punishment) behaviour tends to decrease more rapidly, and to a greater degree, than if only a proportion of reinforced responses are punished (i.e. intermittent or partial punishment) (Hendry & Van-Toler, 1964; Lerman, Iwata, Shore, & DeLeon, 1997; Linden, 1976; Zimmerman & Ferster, 1963). However, in the studies presented in this thesis, punishment by the contingent S- stimulus was always arranged on a VR 5 schedule, meaning punishment was always partial. The partial punishment effect and partial reinforcement extinction effect (PPE and PREE, respectively; see Chapter 2) were avoided, as punishment did not move from being partial to being continuous, and because there was no transition from reinforcement to extinction in the presence of the S+ discriminative stimulus.

It is possible that if punishment *had* been continuous in the studies of this thesis, decreases in punished response rate, and shifts in punished choice, would have been greater. However, to make every response produce a 1.5-s contingent S- stimulus would have prevented us from being able to calculate punished S+ response rate, given that S+ response

rate was calculated using only minutes in which S+ was present – an impractical method to apply. A different approach might be to calculate changes in the duration of inter-response times (IRT) as a result of continuous punishment by a contingent S- stimulus. For example, Davidson (1970) trained rats to choose between two concurrently available schedules that required different behaviour – either a set number of responses for a greater amount of food, or a set period of no responses for a relatively lesser amount of food. Although the rats preferred to emit a set number of responses for the greatest overall rate of reinforcers, when the last response was always punished with shock Davidson found inter-response times (IRTs) on that alternative increased, indicating a shift in preference towards the alternate behaviour (also see Dunham, 1972; Sizemore & Maxwell, 1993). Although Davidson’s study did provide an alternative reinforced behaviour, his findings nevertheless suggest it might be possible to assess punishment on simple schedules by measuring changes in IRTs for S+ responding between punished and unpunished conditions.

Research shows that animals sometimes prefer discriminative stimuli that signal punishment contingencies, even when signaled punishment intensity is greater than that of unsignaled punishment intensity (Badia, Coker and Harsh, 1973; Badia, Culbertson & Harsh, 1973; MacDonald, 1973). Although these studies provide a choice between two punished alternatives, rather than between punished and unpunished alternatives, the findings suggest that organisms might be prepared to pay a price to ensure that stimuli signaling punishment contingencies are present rather than absent (but see Doughty, Doughty, O’Donnell, Saunders and Williams, 2007). However, these studies do not shed light on whether signaled punishing stimuli are more effective at reducing behaviour than unsignalled punishing stimuli. In the studies of this thesis, there is no discriminative stimulus signaling that responses in the presence of S+ will produce the contingent S- stimulus. Yet the S- stimulus is itself a discriminative stimulus for an absence of reinforcers. More research is required to determine whether a different discriminative stimulus that signals conditions in which S+ responses produce a S- contingent stimulus would impact punishment efficacy. For example, we might present a yellow keylight once, for several seconds duration, before the onset of the green keylight/s of the S+ component in punishment tests.

In the second study with humans, it is problematic that only 11 of 29 participants showed evidence of S+/S- discrimination learning. Replications of this study need to establish the procedural conditions necessary for discrimination learning to be acquired before arranging a punishment test. Although punishment efficacy has been shown to be

moderated by differences of species and phenotype (e.g. Torres et al., 2017), given the applied implications of the second study, more research is required to understand how environmental variables such as procedure, and reinforcer and punisher magnitude and delivery schedules might produce more consistent effects of punishment by a contingent S-stimulus.

Despite some limitations, the results of the punishment studies that comprise this thesis suggest that a contingent S- stimulus has potential to function as a non-harmful punishing stimulus. The finding that stimuli do not need to be harmful to punish behaviour fits within the context of existing research which shows that behaviour changes according to the extent that stimuli are able to signal likely future events, rather than by the nature of the stimuli (Cowie, Davison & Elliffe, 2011; Cowie & Davison, 2016). This is supported by the findings of the studies within this thesis which show that if a negative discriminative stimulus is made response-contingent, behaviour will decrease, and choice will shift, because the contingent S- stimulus signals an absence of reinforcers.

Overall, the findings of this thesis add to existing research investigating whether a contingent S- stimulus will function as a punisher (e.g. Gaynor & Shull, 2002; Mulvaney, Dinsmoor, Jwaideh & Hughes, 1974). Specifically, we show that a contingent S- stimulus has the potential to punish the same response that produces positive reinforcers, and to shift choice, without recourse to other consequences. Investigation of the potential of a contingent S- stimulus to change behaviour is in the early stages, and more research is required to establish if changes to procedural parameters will improve the reliability of a contingent S-stimulus as a punisher. However, continuing to investigate this phenomenon is beneficial, given that many applied behavioral treatments attempt to reduce problem behaviour that continues to be maintained by access to positive reinforcers (i.e. auto reinforcement).

Given our knowledge of what makes a punishing stimulus relatively more or less effective (see Chapter 2), we can now begin to investigate the impact of changes to environmental variables on the efficacy of a contingent S- stimulus as a punisher. If the punishing effects of a contingent S- discriminative stimulus increase according to changes in procedural variables, as they should, the applied implications of the current findings will be strengthened.

## CHAPTER 10.

### CONCLUSION

Harmful behaviour is an ongoing problem for behavioural practitioners, patients and caregivers. Although ABA-based interventions such as DRA have a high rate of efficacy in reducing harmful behaviour, post-treatment relapse of behaviour can increase beyond pre-treatment levels. Combining DRA-type treatments with punishment can be more efficacious in eliminating or reducing harmful behaviour (see Chapter 2), however the use of harmful punishing stimuli is ethically problematic. Clearly, an opportunity exists for a new approach to the development of punishing stimuli able to decrease harmful behaviour without producing unwanted side-effects.

contingent S- stimulus can punish behaviour, it is possible a contingent S- stimulus might be used in applied research that investigates interventions to reduce or eliminate harmful behaviour. Research that translates basic findings regarding new approaches to behaviour change into applied settings is timely and worthwhile, given that pharmacological interventions designed to reduce harmful behaviour are over-prescribed (Tsiouris, 2010), do not directly target behaviour, and come with a risk of side-effects (McKinney & Renk, 2011; Matson, Sipes, Fodstad, & Fitzgerald, 2011; Tyrer et al., 2008).

How might a contingent S- stimulus be applied as a punisher in a natural world treatment setting? One approach might be to combine S+/S- stimulus discrimination training with a standard DRA training procedure, before making the S- stimulus contingent on harmful behaviour. For example, if S+ signals a relatively lower rate of reinforcers for harmful behaviour, S++ a relatively higher rate of reinforcers for an alternative non-harmful behaviour, and S- the absence of reinforcers, we might train using a two-component concurrent multiple schedule as follows:

Component 1: S+/S++ (standard DRA)

Component 2: S-/S++.

The type of reinforcer made available during S+ and S++, and unavailable during S-, can be predetermined through preference tests. However, the reinforcer selected must be of a type that is able to be fully withheld in the presence of S-. For example, while access to an iPad is easy to withhold, withholding attention during dangerous or harmful behaviour is more

challenging. The ability of S- to signal an absence of reinforcers will be compromised if reinforcers are occasionally made available in the presence of S- during the training period.

When choice is near exclusive to S++ in Component 2, we might then make behaviour in the presence of S+ briefly produce the contingent S- stimulus in Component 1 on a FR 1 schedule. How this occurs will depend upon the participant or subject, the setting and available resources. However, the low rate reinforcers for behaviour in the presence of S+ should not be withheld in Component 1 during punishment so that this procedure, or variations of it, can accommodate the potential for auto-reinforcement of harmful behaviour, whilst at the same time punishing that behaviour. A relatively higher rate of reinforcers for alternative behaviour (S++) should also be provided in Component 2 to avoid an increase in harmful behaviour caused by extinction when training S-. Further, this procedure might avoid the potential for post-DRA treatment relapse if punishment of harmful behaviour by a contingent S- stimulus is continued in the absence of a higher rate of reinforcers for an alternative behaviour.

Going forward, basic research regarding the potential of a contingent S- stimulus to shift choice in human may inform applied research into new approaches to behaviour change. For example, some studies have investigated whether informational cues that pop up on electronic gaming machines to remind gamblers of their losses will decrease gambling behaviour. Recent research in New Zealand shows that these compulsory interruptions of gambling have marginal benefits (du Preez, Landon, Bellringer, Garrett, & Abbott, 2016; Landon, Palmer du Preez, Bellringer, Page, & Abbott, 2016). We might now extend the findings of this thesis to investigate whether a contingent S- stimulus that signals an *experienced* absence of reinforcers in the form of money, or life or relationship quality, will be more effective in reducing incidence of suboptimal gambling behaviour than informational cues. That is, we might develop specific symbols as S- stimuli, embed these within a game, and then make those symbols contingent on excessive or harmful gambling. These, and many other lines of investigation, may be pursued to better understand how a contingent S- stimulus functions as a punishing consequence in applied treatment settings.

For years, research and treatment methods incorporating punishment have been in decline. Yet harmful behaviour continues to negatively impact lives. Using a negative discriminative stimulus as a punisher is one way to avoid ‘throwing the baby out with the bathwater’. If we combine the basic principles of learning with a desire to not cause harm, then punishment has the potential to be a positive learning process in our lives.

## APPENDIX 1

Summary of 82 basic research studies, including authors, year of publication, species, sample size, the behaviour targeted for punishment, the punishing stimulus used, whether the punishing stimulus was used alone (SOLO) or in combination with another stimulus (COMB), manipulated independent variables and dependant measure. KEY: CP = continuous punishment, DISCR. STIM = discriminative stimulus, DUR = duration, FREQ = frequency, INTS = intensity, IP = intermittent punishment, MON-LOSS = money loss, OLF = olfactory punisher, RF = reinforcer, RWD = reward, SHK = shock, SIG = signaled, UNSIG = not signaled

SOURCE	SPECIES	N	PUNISHED BEHAVIOUR	PUNISHER	SOLO/ COMB	VARIABLES	DEPENDANT MEASURE
Abramson & Feinman (1987)	Crab	~10	Eye extension	AIR-PUFF	S	NA	Time to extension
Appel (1963)	Rats	4	Lever-pressing for food	SHK	S	INTS	Response rate
Arbuckle & Lattal (1992)	Pigeons	3	Key-pecking for food	SHK	S	NA	Inter-response time
Auge (1977)	Pigeons	2	Key-pecking for food	S- STIMULUS	S	Time to reinforcers	Response rate
Ayllon & Azrin (1966)	Humans	3	Plunger-pulls for tokens	NOISE	S	NA	Choice
Azrin (1958)	Humans	80	Switch-pressing for task success	NOISE	S	INTS	Response rate
Azrin (1959, 1960)	Pigeons	6	Key-pecking for food	SHK	S	INTS	Response rate
Azrin (1970)	Monkeys	3	Shock-induced biting	SHK	S	INTS	Response rate
Azrin & Holz 1961	Pigeons	2	Key-pecking for food on FI schedule	SHK	S	INTS	Response rate
Azrin, Holz & Hake (1963)	Pigeons	7	Key-pecking for food	SHK	S	INTS/ SCHEDULE	Response rate
Badia, Coker & Harsh (1973)	Rats	4	Lever-pressing for signalled shock	SHK	S	FREQ, DUR, INTS	Preference
Balaban, Rhodes & Neuringer (1990)	Humans	20	Task acquisition	SHK	S	SIG. vs UNSIG	Rate of Acquisition
Banks (1967)	Rats	16	Runway runs for food	SHK	S	FREQ	Response rate
Banks & Vogel-Sprott (1965)	Humans	40	Button-presses for money	SHK	S	DELAY	Response rate
Barker et al. (2010)	Rats	6	Lever-pressing for food	LIGHT	S	NA	Response rate
Bergman & Johnanson (1981)	Monkeys	3	Lever-pressing for cocaine	SHK	S	INTS	Response rate
Baron, Kauffman & Fazzini (1969)	Rats	5	Lever-pressing to avoid shock	SHK	S	FREQ, DELAY	Response rate
Berez (1972)	Humans	133	Cigarette smoking	SHK	S	PUN V. UNPUN	Response rate
Beyra & Spinewine (1973)	Rats	12	Bar-pressing for EBS	SHK	CB - LIGHT	RFT MAGNITUDE	Response rate
Bouton & Schepers (2015)	Rats	32	Lever-pressing for food	SHK	S	CONTEXT	Response rate
Bradshaw, Szabadi, & Bevan (1979)	Humans	3	Button-pressing for money	MON-LOSS	S	FREQ, CHOICE	Response rate
Brethower & Reynolds (1962)	Pigeons	2	Key-pecking for food	SHK	S	INTS	Response rate
Cohen (1968)	Rats	24	Lever-pressing for food	SHK	S	DELAY, INTS	Response rate
Dardano (1968)	Pigeons	3	Key-pecking for food	SHK	S	INTS	Escape rate

SOURCE	SPECIES	N	PUNISHED BEHAVIOUR	PUNISHER	SOLO/ COMB	VARIABLES	DEPENDANT MEASURE
Davidson (1970)	Rats	2	Lever-pressing for food	SHK	S	SHOCK SIG	Response rate
Davison (1970)	Rats	14	Lever-pressing for food	SHK	S	SCHEDULE	Preference
Deluty (1976)	Rats	NA	Lever-pressing for food	SHK	S	FREQ	Response rate
Deur & Parke (1970)	Humans	120	Hitting object for marbles	NOISE	S	CP vs IP	Response rate
Dinsmoor (1952)	Rats	12	Lever-pressing for food	SHK	S	INTS	Preference
Doughty et al. (2007)	Humans	3	Stereotypy	RES	S	DISC. STIM.	Response rate
Dunham (1972)	Rats	4	Drinking	SHK	S	PUN vs NON-PUN	Preference
Ernst & Lee (1975)	Rats	10	Maze running	SHK	S	AGE, TRAINING	Responses
Everly & Perone (2012)	Rats	16	Lever-pressing for food	SHK	S	INTS/IRT	Response rate
Falk & Burnidge (1970)	Rats	8	Drinking	NACL	S	DRUGS	Response rate
Filby & Appel (1966)	Rats	12	Lever-pressing for food	SHK	S	INTS, FREQ	Response rate
Fowler, Hochhauser & Wischner (1981)	Rats	64	Correct discrimination	SHK	S	PUN V. UNPUN	Discrimination
Friedel, Hart & Odum (2017)	Rats	3	Lever-pressing for food	TONES	S	INTS	Response rate
Furakawa et al. (2016)	Humans	145	Preference	RWD LOSS	S	LEVEL OF PUN	Choice
Galbicka & Branch (1983)	Monkeys	2	Lever-pressing for food	SHK	S	SHK DELIVERY	IRT
Geller (1970)	Rats	12	Lever-pressing for food or EBS	SHK	S	"RF TYPE, INTS"	Response rate
Grove & Schuster (1974)	Monkeys	3	Responding for cocaine	SHK	S	INTS, PUN v UNPUN	Punished response rate
Hake, Azrin & Oxford (1967)	Monkeys	7	Bar-pressing for food	SHK	S	INTS	Response rate
Halevy, Feldon & Weiner (1987)	Rats		Runway runs for food	SHK	S	CP/CR vs. IP/IR	Response rate
Hanson et al. (1967)	Monkeys	4	Lever-pressing for food	SHK	CB-RFT/EXT	DRUGS	Response rate
Hendry & Van-Toller (1964)	Rats		Lever-pressing for water	SHK	S	SCHEDULE	Response rate
Herman & Azrin (1964)	Humans	3	Knob-turns for cigarettes	NOISE	S	SCHEDULE	Response rate
Holz et al. (2013)	Rats	63	Lever-pressing for cocaine	HISTAMINE	S	AGE	Response rate
Holz (1968)	Pigeons	2	Preference	SHK	S	SCHEDULE, INTS	Response rate
Holz & Azrin (1961; 1962)	Pigeons	3/2	Key-pecking for food	SHK	CB/RFT	SHOCK S+	Response rate
Hopf et al. (2010)	Rats	NA	Lever-pressing for alcohol	QUININE	S	RF TYPE	Motivation
Laurence, Hinline & Bersch (1994)	Monkeys	4	Lever-pressing to avoid shock	SHK	S	SCHEDULE	Response rate
Leong et al. (2006)	Rats	NA	Lever-pressing for cocaine	LICL	S	RF ACCESS	Response rate
Lerman et al. (1997)	Humans	5	Self-injurious behaviour	RES	S	SCHEDULE	Incident rate
Li, Hsiao, & Li (2013).	Rats	16	Drinking NaCl sucrose solution	LICL	S	DOSAGE	Amount consumed
MacDonald (1973)	Rats	6	Preference	SHK	S	SIG vs. UNSIG	Preference
Martin et al. (1993)	Mice	NA	Lever-pressing for food	SHK	S	DRUGS	Response rate

SOURCE	SPECIES	N	PUNISHED BEHAVIOUR	PUNISHER	SOLO/ COMB	VARIABLES	DEPENDANT MEASURE
Martin et al. (2014)	Humans	20	Key-pressing for monetary gain	MON-LOSS	S	MRI RESULTS	Response rate
Negus (2005)	Monkeys	3	Preference	HISTAMINE	S	RF TYPE	Preference
Myer (1967)	Rats	20	Mice killing	SHK	S	INTS, EXPERIENCE	Kill rate
Orme-Johnson & Yarczower (1974)	Pigeons	6	Key-pecking for food	SHK	S	NCR vs CR	Response rate
Pelloux, Murray & Everitt (2015)	Rats	NA	Cocaine-seeking and taking	SHK	CB-ALT RF	ALT SOURCE RF	Response rate
Pietras & Hackenberg (2005)	Pigeons	4	Key-pecking for food tokens	TOKEN LOSS	S	RESPONSE COST	Response rate
Podlesnik et al. (2010)	Rats	10	Preference	HISTAMINE	S	DOSAGE	Preference
Poling & Thompson (1977)	Rats	2	Lever-pressing for ethanol	SHK	S	INTS	Amount consumed
Rachlin (1966)	Pigeons	4	Key-pecking for food	SHK	S	INTS	Response rate
Reed & Yoshino (2008)	Rats	NA	Lever-pressing for food	NOISE	S	FREQ, DURATION	Response rate
Riccio & Hamm (1972)	Rats	39	Accessing goal box	COLD WATER	S	EXPERIENCE	Speed of access
Roberts & Blasé (1971)	Rats	30	Shock-induced aggression	SHK	S	PUN vs, UNPUN	Rate of aggression
Rodriguez & Logan (1980)	Rats	10	Runway traverses for water	SHK	S	RESPONSE TYPE	Preference
Sawyer et al. (2016)	Humans	82	Key-presses for money	MON-LOSS	S	PTSD vs no PTSD	Reaction time
Scobie & Kaufman (1969)	Humans	5	Key-presses for money	SHK	S	INSTRUCTION	Response rate
Sierra-Machado et al. (2015)	Humans &	22	Button pressing for reward	EYE PUFF	S	RF MAGNITUDE	Preference
Sizemore & Maxwell (1993)	Rats	3	Bar-pressing for food	SHK	S	CONT vs NON-CON	IRT
Smith, Abramson & Tobin (1991)	Bees	NA	Sucrose feeding	SHK	S	PUN vs. UNPUN	Response rate
Steiss et al. (2007)	Dogs	24	Barking	SHK/OLF	S	PUNISHER TYPE	Response rate
Stokman & Glusman (1968)	Cats	4	Escape (hypothalamus-induced)	SHK	S	PUN vs UNPUN	Escape rate
Thompson (1965)	Pigeons	2	Key-pecking for food	DISC. STIM	S	SCHEDULE	Response rate
Todorov (1971)	Pigeons	4	Changeovers between keys	SHK	S	SCHEDULE VALUE, INTS	Number of changeovers
Torres et al. (2017)	Rats	NA	Lever-pressing for meth	SHK	S	PHENOTYPE	Amount consumed
van Oyen et al. (1981)	Rats	20	Lever-pressing for food	SHK	S	SEX	Response rate
Weisman (1975)	Pigeons	2	Key-pecking for food	SHK	S	DISC. STIM.	Response rate
Wesp, Lattal & Poling (1977)	Pigeons	4	Autoshaped key-pecking	SHK	S	INTS	Response rate

## APPENDIX 2

Summary of 62 applied research studies, including authors, year of publication, species, sample size, behaviour targeted for punishment, punishing stimulus, whether the punishing stimulus was used alone (SOLO) or in combination (COMB) with another stimulus, dependent measures, whether alternative methods were used, whether punishment was more effective than alternative methods, percentage of behaviour reduction, and reduction maintenance. KEY: ADD = addictive behaviour, AGG = aggressive behaviour, CBT = cognitive behaviour therapy, DES = property destruction, DRA = differential reinforcement of alternative behaviour, DRL = differential reinforcement of low rates of behaviour, DRO = differential reinforcement of other behaviour, ECT = electro-convulsive therapy, EXT = extinction, FAC. SCR = facial screening, FCT = functional communication training, GROUP = group therapy GUS = gustatory punisher, H = Humans, OC = overcorrection, OLF = olfactory punisher, PRED = predatory behaviour (animals), RES = physical restraint, SHK = shock, SIB = self-injurious behaviour, SOC = socially harmful behaviour, STER = stereotypy, TBI = traumatic brain injury, TO = timeout, VER = verbal reprimand, VIS-SCR = visual screening, WAT-MIST = water mist.

SOURCE	SPECIES	N	BEHAVIOUR	PUNISHER	SOLO OR COMB.	DEPENDENT MEASURE	ALT METHODS	MORE EFFECTIVE	REDUCTION	MAINTAINED (MONTHS)
Altmeyer, Williams, & Sams (1985)	H	1	SIB	GUS	S	Incidents per 12 hours	DRO/TO/RES	NA	79%	20
Altman, Haavik & Cook (1978)	H	2	SIB	OLF	S	Incidents per day	DRO	Y	50%	2
Amtnen & Werner (1999)	H	1	SIB	WAT-MIST	S	AGG incidents per session	N	Y	95%	19
Azrin et al. (1975)	H	11	SIB	OC	COMB - DRA	Incidents per monitored period	N	NA	99%	3
Azrin & Wesolowski (1975)	H	1	SIB	OC	S	Incidents per monitored period	TO/RES	Y	99%	12
Baroff & Tate (1968)	H	1	SIB	SHK	S	Incidents per day	RES	Y	100%	9
Barrett et al. (1981)	H	2	STER	VIS-SCR	S	Incidents per monitored period	DRO	Y	100%	6
Baumeister & Baumeister (1978)	H	2	SIB	OLF	S	Incidents per day	RES/DRA	Y	100%	4
Baumeister & Forehand (1972)	H	8	STER	SHK	COMB - VERBAL	Rocks per min	N	NA	100%	10
Berez (1972, 1973)	H	1	SOC	SHK	COMB - VIS	Smoking/Stuttering	N	NA	60%	1
Birnbrauer (1968)	H	1	AGG/DES	SHK	S	Incidents per session	RES/ECT/TO/DRO	INC	INC	4
Brandma & Stein (1973)	H	1	AGG	SHK	S	Incidents per session	ECT/DRUGS	Y	99%	12
Bucher & King (1971)	H	1	DEST	SHK	S	Touches per 30-sec interval	RES/VER/DRA	N	No	Not reported
Chapman, Smith & Layden (1971)	H	23	ADD	SHK	COMB - DRA, CBT	Cigarettes smoked per day	N	NA	30%	12
Conley & Wolery (1980)	H	2	SIB	OC	S	Incidents per interval	N	NA	100%	9
Corte, Wolf & Locke (1971)	H	4	SIB	SHK	S	Incidents per session	DRO	Y	NA	2
Cunningham & Linscheid (1976)	H	1	SIB	SHK	S	Incidents per monitored period	DRA	Y	100%	3
Dale, Podlesnik & Elliffe (2017)	Dogs	1156	PRED	SHK	S	Avoidance of training stimulus	N	NA	78%	12

SOURCE	SPECIES	N	BEHAVIOUR	PUNISHER	SOLO OR COMB.	DEPENDENT MEASURE	ALT METHODS	MORE EFFECTIVE	REDUCTION	MAINTAINED (MONTHS)
Decatanzaro & Baldwin (1978)	H	2	SIB	OC	COMB - DRO	Incidents per interval	N	NA	95%	Not reported
Dorsey et al. (1980)	H	7	SIB	WAT-MIST	COMB - VER/DRO	Incidents per 20 min session	DRO/VER	Y	60-95%	Not reported
Duker & Seys (2000)	H	8	SIB	SHK	S	Incidents per day	RES	Y	82%	36
Dura (1991)	H	1	AGG	RES	COMB - DRA	Incidents per monitored period	N	NA	100%	54
Eckstein & Hart (1996)	Dogs	4	SIB	SHK	S	Incidents per monitored period	N	NA	100%	12
Fisher et al. (1994)	H	3	SIB/PICA	FAC. SCR	COMB-DRA	Incidents per monitored period	N	NA	90%	Not reported
Foxx & Azrin (1972; 1973)	H	7	SIB/STER	OC	S	Incidents per monitored period	PHYS/DRA/DRO	Y	100%	1
Foxx & Meindl (2007)	H	1	AGG	OC	COMB-DRO/RC	Incidents per day	RES	Y	95%	8
Foxx et al. (1986)	H	1	AGG/DES	SHK	COMB - DRA	Incidents per monitored period	DRUGS/RES/TO	Y	100%	12
Foxx et al. (1975; 1979)	H	6	SIB	OC	S	Incidents per day	RES	Y	50 - 90%	>1-4
Gordon et al. (1986)	H	1	SOC	OC	S	Incidents per day	N	NA	95%	12
Greene & Hoats (1971)	H	2	SIB	TICKLE	S	Incidents per week	N	NA	60%	2
Hagopian et al. (1998)	H	21	SIB/AGG	TO/RES/OC	COMB - FCT	Problem behaviour per minute	FCT/EXT	Y	90%	Not reported
Hanley et al. (2005)	H	2	SIB/AGG	RES/TO	COMB-FCT	Problem behaviour per minute	FCT	Y	99%	Not reported
Harris & Romanczyk (1976)	H	1	SIB	OC	S	Incidents per day	N	NA	99%	9
Harris & Wolchik (1979)	H	4	STER	OC	S	Incidents per 15-min session	DRO/TO	Y	99%	1
Johnson et al. (1982)	H	3	SIB/STER	OC	S/CB-DRA	OC with DRA vs OC alone	N	NA	low levels'	Not monitored
Libet et al. (1973)	H	1	SIB	GS	S	Incidents	N	NA	100%	No
Linscheid et al. (1994)	H	1	SIB	SHK	S	Incidents per 10 min session	DRA/DRL	Y	99%	12
Linscheid & Reichenbach (2002)	H	1	SIB	SHK	COMB-DRUGS	Incidents per day	RES/DRA	Y	99%	60
Luckey et al. (1968)	H	1	SIB	SHK	S	Incidents per day	N	NA	Marked	3
Ludwig et al. (1969)	H	1	AGG	SHK	S	Incidents per day	DRUGS/GROUP/DRA	Y	95%	5
McKenzie et al. (2008)	H	1	SIB	VER REP	S	Incidents per 10-s	N	NA	99%	Maintained via conditioned punishment
McFarlain et al. (1978)	H	1	SIB	SHK	COMB-VER	Incidents per session	N	NA	Significant	Not monitored
Matson, Stephens & Smith (1978)	H	1	SIB	GS	COMB-OC	Incidents per day	TO/RES/DRA	Y	95%	Monthly
Matson & Stephens (1977)	H	1	AGG	OC	S	Incidents per day	DRA	Y	95%	3
Muttar et al. (1975)	H	1	SIB/AGG	SHK	COMB-VER	Incidents per day	DRA/RES	Y	100%	20
Prochaska et al. (1974)	H	1	SIB	SHK (remote)	S	Incidents per day	DRA/RES/DRUGS/TO	Y	100%	7
Quist & Martin (1969)	H	3	STUTTER	VER	S	Incidents per session	N	NA	30-100%	N
Rapoff et al. (1980)	H	1	SIB	RES	COMB-VER	Incidents per day	Non-contingent RES	Y	100%	2
Ricketts, Goza & Matese (1993)	H	1	SIB	SHK	S	Incidents per hour	Not reported	NA	99%	31*
Risley (1968)	H	1	STER/AGG	SHK	COMB-VER	Incidents per day	TO/EXT	Y	90%	Not monitored
Rusch et al. (1976)	H	1	STER	OC	COMB-VER	Incidents per day	Not reported	NA	90%	Not monitored
Sajwaj, Libet & Agras (1974)	H	1	SIB	GS	S	Incidents per day	N	NA	100%	12

SOURCE	SPECIES	N	BEHAVIOUR	PUNISHER	SOLO OR COMB.	DEPENDENT MEASURE	ALT METHODS	MORE EFFECTIVE	REDUCTION	MAINTAINED (MONTHS)
Salvy et al. (2004)	H	1	SIB	SHK	COMB-VER	Incidents per day	N	NA	100%	8
Sargisson, Butler & Elliffe (2011)	Dogs	7	Barking	GS/OLF	S	Barking per 10-s interval	N	NA	Partial	Not monitored
Simmons & Lovaas (1969)	H	9	DES	SHK	COMB-VER	Incidents per day	N	NA	100%	Not monitored
Sisson & Hersen (1993)	H	2	AGG/DES	OC	COMB-DRA	Incidents per day	N	NA	95%	Period not specified
Tanner & Zeiler (1975)	H	1	SIB/Pica	OLF	COMB-DRUGS	Incidents per observation period	RES/SHK/DRUGS	Y	100%	-1
Tate & Baroff (1966)	H	1	SIB/AGG	SHK/EXT	COMB-VER	Incidents per period	EXT/DRUGS/GROUP	N	80%	6
Tiger et al. (2017)	H	1	STER/SIB	EXT	COMB-DISC STIM	Incidents per period	N	NA	80-100%	Not monitored
Toister et al. (1975)	H	1	REG (SIB)	SHK	S	Incidents per min	N	NA	100%	8
Tolman & Mueller (1964)	Monkey	1	SIB	NOISE/TO	S	Incidents per observation period	N	NA	100%	Not monitored
Williams et al. (1993)	H	1	SIB	SHK (High INTS)	S	Incidents per monitored period	SIBIS/DRA/EXT	Y	100%	6

## REFERENCES

- Abramson, C. I., & Feinman, R. D. (1987). Operant punishment of eye elevation in the green crab, *Carcinus maenas*. *Behavioral and Neural Biology*, *48*, 259-277.
- Alsop, B. (2018). Comment on Bland, Cowie, Elliffe, and Podlesnik (2018), Melioration revisited: A systematic replication of Vaughan (1981). *Journal of the Experimental Analysis of Behavior*, *110*, 154-155.
- Altman, K., Haavik, S., & Cook, J. W. (1978). Punishment of self-injurious behavior in natural settings using contingent aromatic ammonia. *Behavior Research and Therapy*, *16*, 85-96.
- Altmeyer, B. K., Williams, D. E., & Sams, V. (1985). Treatment of severe self injurious and aggressive biting. *Journal of Behavior Therapy and Experimental Psychiatry*, *16*, 169-172.
- Antshel, K. M., & Barkley, R. (2008). Psychosocial interventions in attention deficit hyperactivity disorder. *Child and Adolescent Psychiatric Clinics of North America*, *17*, 421-437.
- Appel, J. B. (1963). Punishment and shock intensity. *Science*, *141*, 528-529.
- Arbuckle, J. L., & Lattal, K. A. (1992). Molecular contingencies in schedules of intermittent punishment. *Journal of the Experimental Analysis of Behavior*, *58*, 361-375.
- Arhant, C., Bubna-Littitz, H., Bartels, A., Futschik, A., & Troxler, J. (2010). Behavior of smaller and larger dogs: effects of training methods, inconsistency of owner behavior and level of engagement in activities with the dog. *Applied Animal Behavior Science*, *123*, 131-142.
- Arntzen, E., & Werner, S. B. (1999). Water mist punishment for two classes of problem behavior. *Scandinavian Journal of Behavior Therapy*, *28*, 88-93.
- Auge, R. J. (1977). Stimulus functions within a fixed-interval clock schedule: Reinforcement, punishment, and discriminative stimulus control. *Learning & Behavior*, *5*, 117-123.
- Axelrod, S. A. (1990). Myths that (mis)guide our profession. In A.C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities*, 59-72. Sycamore, IL: Sycamore.
- Axelrod, S. (Ed.). (2013). *Effects of Punishment on Human Behavior*. Academic Press.
- Axelrod, S., & Apsche, J. (Eds.). (1983). *The Effects of Punishment on Human Behavior*.

- New York: Academic.
- Ayllon, T., & Azrin, N. H. (1966). Punishment as a discriminative stimulus and conditioned reinforcer with humans. *Journal of the Experimental Analysis of Behavior*, 9, 411-419.
- Azrin, N. H. (1956). Some effects of two intermittent schedules of immediate and non immediate punishment. *The Journal of Psychology*, 42, 3-21.
- Azrin, N. H. (1958). Some effects of noise on human behavior. *Journal of the Experimental Analysis of Behavior*, 1, 183-200.
- Azrin, N. H. (1959). Punishment and recovery during fixed-ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, 2, 301-305.
- Azrin, N. H. (1960). Effects of punishment intensity during variable-interval reinforcement. *Journal of the Experimental Analysis of Behavior*, 3, 123-142.
- Azrin, N. H. (1970). Punishment of elicited aggression. *Journal of the Experimental Analysis of Behavior*, 14, 7-10.
- Azrin, N. H., & Holz, W. C. (1961). Punishment during fixed-interval reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 343-347.
- Azrin, N. H., & Holz, W. C. (1966). Punishment. *Operant behavior: Areas of research and application*, 380-447.
- Azrin, N. H., Holz, W. C., & Hake, D. F. (1963). Fixed-ratio punishment. *Journal of the Experimental Analysis of Behavior*, 6, 141.
- Azrin, N. H., & Wesolowski, M. D. (1975). Eliminating habitual vomiting in a retarded Adult by positive practice and self-correction. *Journal of Behavior Therapy and Experimental Psychiatry*, 6, 145-148.
- Azrin, N. H., Gottlieb, L., Hughart, L., Wesolowski, M. D., & Rahn, T. (1975). Eliminating self-injurious behavior by educative procedures. *Behavior Research and Therapy*, 13, 101-111.
- Badia, P., Coker, C., & Harsh, J. (1973). Choice of higher density signalled shock over lower density unsignalled shock. *Journal of the Experimental Analysis of Behavior*, 20, 47-55.
- Badia, P., Culbertson, S., & Harsh, J. (1973). Choice of longer or stronger signalled shock over shorter or weaker unsignalled shock. *Journal of the Experimental Analysis of Behavior*, 19, 25-32.
- Balaban, M. T., Rhodes, D. L., & Neuringer, A. (1990). Orienting and defense responses to punishment: Effects on learning. *Biological Psychology*, 30, 203-217.

- Balliet, D., Mulder, L. B., & Van Lange, P. A. (2011). Reward, punishment, and cooperation: a meta-analysis. *Psychological Bulletin*, *137*, 594. Research Collection School of Social Sciences, Singapore Management University.
- Balsam, P. D., & Bondy, A. S. (1983). The negative side effects of reward. *Journal of Applied Behavior Analysis*, *16*, 283.
- Banks, R. K. (1966). Persistence to continuous punishment following intermittent punishment training. *Journal of Experimental Psychology*, *71*, 373-377
- Banks, R. K. (1967). Intermittent punishment effect (IPE) sustained through changed *Learning & Behavior*, *4*, 118-120.
- Banks, R. K., & Vogel-Sprott, M. (1965). Effect of delayed punishment on an immediately rewarded response in humans. *Journal of Experimental Psychology*, *70*, 357.
- Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, *27*, 325-344.
- Barker, D. J., Sanabria, F., Lasswell, A., Thrailkill, E. A., Pawlak, A. P., & Killeen, P. R. (2010). Brief light as a practical aversive stimulus for the albino rat. *Behavioral Brain Research*, *214*, 402-408.
- Baroff, G. S., & Tate, B. G. (1968) The use of aversive stimulation in the treatment of chronic self-injurious behavior. *Journal of the American Academy of Child Psychiatry*, *7*, 454-470.
- Baron, R. A. (1977). The Prevention and Control of Human Aggression. In *Human Aggression* (pp. 225-274). Springer US.
- Baron, R. A., Kaufman, A., & Fazzini, D. (1969). Density and delay of punishment of free operant avoidance. *Journal of the Experimental Analysis of Behavior*, *12*, 1029-1037.
- Barrett, R. P., Matson, J. L., Shapiro, E. S., & Ollendick, T. H. (1981). A comparison of punishment and DRO procedures for treating stereotypic behavior of mentally retarded children. *Applied Research in Mental Retardation*, *2*, 247-256.
- Baum, W.M. (1979). Matching, overmatching and undermatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, *32*, 269-281.
- Baum, W. M. (2012). Rethinking reinforcement: Allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior*, *97*, 101-124.
- Baumeister, A. A., & Baumeister, A. A. (1978). Suppression of repetitive self-injurious behavior by contingent inhalation of aromatic ammonia. *Journal of Autism and Childhood Schizophrenia*, *8*, 71-77.

- Baumeister, A. A., & Forehand, R. (1972). Effects of contingent shock and verbal command on body rocking of retardates. *Journal of Clinical Psychology, 28*, 586-590.
- Baumeister, A. A., & Rollings, J. P. (1976). Self-injurious behavior. *International review of Research in Mental Retardation, 8*, 1-34.
- Beavers, G. A., Iwata, B. A., & Lerman, D. C. (2013). Thirty years of research on the functional analysis of problem behavior. *Journal of Applied Behavior Analysis, 46*, 1–21.
- Bécuwe-Bonnet, V., Bélanger, M. C., Frank, D., Parent, J., & Hélie, P. (2012). Gastrointestinal disorders in dogs with excessive licking of surfaces. *Journal of Veterinary Behavior: Clinical Applications and Research, 7*, 194-204.
- Berecz, J. M. (1972). Modification of smoking behavior through self-administered punishment of imagined behavior: a new approach to aversion therapy. *Journal of Consulting and Clinical Psychology, 38*, 244.
- Berecz, J. M. (1973). The treatment of stuttering through precision punishment and cognitive arousal. *Journal of Speech and Hearing Disorders, 38*, 256-267.
- Bergman, J., & Johanson, C. E. (1981). The effects of electric shock on responding maintained by cocaine in rhesus monkeys. *Pharmacology Biochemistry and Behavior, 14*, 423-426.
- Berkowitz, L. (1993). Pain and aggression: Some findings and implications. *Motivation and Emotion, 17*, 277-293.
- Berkowitz, L., Cochran, S. T., & Embree, M. C. (1981). Physical pain and the goal of aversively stimulated aggression. *Journal of Personality and Social Psychology, 40*, 687.
- Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences, 26*, 507-513.
- Beyra, M., & Spinewine, P. (1973). Inhibition and facilitation of self-stimulation behavior by aversive stimulation. *Physiology & Behavior, 10*, 563-566.
- Birnbrauer, J. S. (1968). Generalization of punishment effects—a case study. *Journal of Applied Behavior Analysis, 1*, 201-211.
- Bland, V. J., Bai, J. Y. H., Fullerton, J. A., & Podlesnik, C. A. (2016). Signaled alternative reinforcement and the persistence of operant behavior. *Journal of the Experimental Analysis of Behavior, 106*, 22-33.
- Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review, 77*, 32.

- Bouton, M. E., & Schepers, S. T. (2015). Renewal after the punishment of free operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 81.
- Boutros, N., Davison, M., & Elliffe, D. (2011). Contingent stimuli signal subsequent reinforcer ratios. *Journal of the Experimental Analysis of Behavior*, *96*, 39-61.
- Bradshaw, C. M., Szabadi, E., & Bevan, P. (1979). The effect of punishment on free operant choice behavior in humans. *Journal of the Experimental Analysis of Behavior*, *31*, 71-81.
- Brandsma, J. M., & Stein, L. I. (1973). The use of punishment as a treatment modality: a case report. *The Journal of Nervous and Mental Disease*, *156*, 30-37.
- Brantner, J. P., & Doherty, M. A. (1983). A review of timeout: A conceptual and methodological analysis. *The Effects of Punishment on Human Behavior*, 87-132.
- Brethower, D. M., & Reynolds, G. S. (1962). A facilitative effect of punishment on unpunished behavior. *Journal of the Experimental analysis of Behavior*, *5*, 191-199.
- Brosnan, J., & Healy, O. (2011). A review of behavioral interventions for the treatment of aggression in individuals with developmental disabilities. *Research in Developmental Disabilities*, *32*, 437-446.
- Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, *11*, 1-8.
- Brown, R. T., & Wagner, A. R. (1964). Resistance to punishment and extinction following training with shock or nonreinforcement. *Journal of Experimental Psychology*, *68*, 503.
- Bucher, B., & King, L. W. (1971). Generalization of punishment effects in the deviant behavior of a psychotic child. *Behavior Therapy*, *2*, 68-77.
- Butterfield, E. C. (1990). The compassion of distinguishing punishing behavioral treatment from aversive treatment. *American Journal on Mental Retardation*, *95*, 137-141.
- Byrne, T., & Poling, A. (2017). Behavioral effects of delayed timeouts from reinforcement. *Journal of the Experimental Analysis of Behavior*, *107*, 208-217.
- Callias, M., Carr, R. J., Corbett, J. & Jenkins, J. (1973). Use of behavior modification techniques in a community service for mentally handicapped children. *Proceedings of the Royal Society of Medicine*, *66*, 1140- 1142.
- Carroll, K. M., & Onken, L. S. (2005). Behavioral therapies for drug abuse. *American Journal of Psychiatry*, *162*, 1452-1460.
- Casey, R. A., Loftus, B., Bolster, C., Richards, G. J., & Blackwell, E. J. (2014). Human

- directed aggression in domestic dogs (*Canis familiaris*): Occurrence in different contexts and risk factors. *Applied Animal Behavior Science*, *152*, 52-63.
- Catania, A. C. (1963). Concurrent performances: A baseline for the study of reinforcers magnitude. *Journal of the Experimental Analysis of Behavior*, *6*, 299-300.
- Chapman, R. F., Smith, J. W., & Layden, T. A. (1971). Elimination of cigarette smoking by punishment and self-management training. *Behavior Research and Therapy*, *9*, 255-264.
- Charach, A., & Fernandez, R. (2013). Enhancing ADHD medication adherence: challenges and opportunities. *Current Psychiatry Reports*, *15*, 371.
- Charlop-Christy, M. H., & Haymes, L. K. (1996). Using obsessions as reinforcers with and without mild reductive procedures to decrease inappropriate behaviors of children with autism. *Journal of Autism and Developmental Disorders*, *26*, 527-546.
- Chow, J. J., Smith, A. P., Wilson, A. G., Zentall, T. R., & Beckmann, J. S. (2017). Suboptimal choice in rats: Incentive salience attribution promotes maladaptive decision-making. *Behavioral Brain Research*, *320*, 244-254.
- Church, R. M. (1963). The varied effects of punishment on behavior. *Psychological Review*, *70*, 369.
- Church, R. M. (1969). Response suppression. *Punishment and aversive behavior*. New York:Appleton-Century-Crofts, 111-156.
- Cohen, P. S. (1968). Punishment: The interactive effects of delay and intensity of shock. *Journal of the Experimental Analysis of Behavior*, *11*, 789-799.
- Cohen-Mansfield, J. (2013). Nonpharmacologic treatment of behavioral disorders in dementia. *Current Treatment Options in Neurology*, *15*, 765-785.
- Conley, O. S., & Wolery, M. R. (1980). Treatment by overcorrection of self-injurious eye gouging in preschool blind children. *Journal of Behavior Therapy and Experimental Psychiatry*, *11*, 121-125.
- Corte, H. E., Wolf, M. M., & Locke, B. J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis*, *4*, 201-213.
- Cowie, S., & Davison, M. (2016). Control by reinforcers across time and space: A review of recent choice research. *Journal of the Experimental Analysis of Behavior*, *105*, 246-269.
- Cowie, S., Davison, M., & Elliffe, D. (2011). Reinforcement: Food signals the time and location of future food. *Journal of the Experimental Analysis of Behavior*, *96*, 63-86.

- Critchfield, T. S., Paletz, E. M., MacAleese, K. R., & Newland, M. C. (2003). Punishment in human choice: Direct or competitive suppression? *Journal of the Experimental Analysis of Behavior, 80*, 1-27.
- Cunningham, C. E., & Linscheid, T. R. (1976). Elimination of chronic infant ruminating by electric shock. *Behavior Therapy, 7*, 231-234.
- Dale, A. R., Podlesnik, C. A., & Elliffe, D. (2017). Evaluation of an aversion-based program designed to reduce predation of native birds by dogs: An analysis of training records for 1156 dogs. *Applied Animal Behavior Science, 191*, 59-66.
- Dale, A. R., Statham, S., Podlesnik, C. A., & Elliffe, D. (2013). The acquisition and maintenance of dogs' aversion responses to kiwi (*Apteryx spp.*) props across time and locations. *Applied Animal Behavior Science, 146*, 107-111.
- Dardano, J. F. (1968). Reversal of preference under progressive-ratio schedules by punishment. *Journal of the Experimental Analysis of Behavior, 11*, 133-146.
- Davidson, R. S. (1970). Conditioned punishment and conditioned negative reinforcement on a multiple schedule. *Psychonomic Science, 20*, 163-165.
- Davis, N. M., Brookes, S., Gray, J. A., & Rawlins, J. N. P. (1981). Chlordiazepoxide and resistance to punishment. *The Quarterly Journal of Experimental Psychology, 33*, 227-239.
- Davison, M. C. (1970). A choice technique to assess the effects of selective punishment on fixed-ratio performance. *Journal of the Experimental Analysis of Behavior, 13*, 57-64.
- Davison, M. C. (1990). Undermatching, melioration and the discrimination of local reinforcer rates. *Behavioral Processes, 21*, 189-195.
- Davison, M., & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Learning & Behavior, 13*, 77-84.
- Davison, M., & McCarthy, D. (1988). *The Matching Law*. Hillsdale, NJ: L.
- Deb, S., Sohanpal, S. K., Soni, R., Lentre, L., & Unwin, G. (2007). The effectiveness of antipsychotic medication in the management of behavior problems in adults with intellectual disabilities. *Journal of Intellectual Disability Research, 51*, 766-777.
- DeCatanzaro, D. A., & Baldwin, G. (1978). Effective treatment of self-injurious behavior through a forced arm exercise. *American Journal of Mental Deficiency, 82*, 433-439.
- Deluty, M. Z. (1976). Choice and the rate of punishment in concurrent schedules. *Journal of the Experimental analysis of Behavior, 25*, 75-80.
- Deroche-Gamonet, V., Belin, D., & Piazza, P. V. (2004). Evidence for addiction-like behavior in the rat. *Science, 305*, 1014-1017.

- Deur, J. L., & Parke, R. D. (1970). Effects of inconsistent punishment on aggression in children. *Developmental Psychology, 2*, 403.
- Devanand, D. P., Mintzer, J., Schultz, S. K., Andrews, H. F., Sultzer, D. L., de la Pena, D., ... & Levin, B. (2012). Relapse risk after discontinuation of risperidone in Alzheimer's disease. *New England Journal of Medicine, 367*, 1497-1507.
- de Villiers, P. A., (1980). Toward a quantitative theory of punishment. *Journal of the Experimental Analysis of Behavior, 33*, 15-25.
- Diden, R., Duker, P. C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals who have mental retardation. *AJMR-American Journal on Mental Retardation, 101*, 387-399.
- Diden, R., Korzilius, H., van Oorsouw, W., & Sturmey, P. (2006). Behavioral treatment of challenging behaviors in individuals with mild mental retardation: Meta-analysis of single subject research. *American Journal on Mental Retardation, 111*, 290-298.
- Dinsmoor, J. A. (1952). A discrimination based on punishment. *Quarterly Journal of Experimental Psychology, 4*, 27-45.
- Dinsmoor, J. A. (1954). Punishment: I. The avoidance hypothesis. *Psychological Review, 61*, 34.
- Dinsmoor, J. A. (1977). Escape, avoidance, punishment: where do we stand? *Journal of the Experimental Analysis of Behavior, 28*, 83-95.
- Dinsmoor, J. A. (1998). Punishment. In W. O'Donohue (Ed.). *Learning and Behavior Therapy*, 188-204. Boston, MA: Allyn & Bacon.
- Doehring, P., Reichow, B., Palka, T., Phillips, C., & Hagopian, L. (2014). Behavioral approaches to managing severe problem behaviors in children with autism spectrum and related developmental disorders. *Child and Adolescent Psychiatric Clinics, 23*, 25-40.
- Dorsey, M. F., Iwata, B. A., Ong, P., & McSween, T. E. (1980). Treatment of self injurious behavior using a water mist: Initial response suppression and generalization. *Journal of Applied Behavior Analysis, 13*, 343-353.
- Doughty, S. S., Anderson, C. M., Doughty, A. H., Williams, D. C., & Saunders, K. J. (2007). Discriminative control of punished stereotyped behavior in humans. *Journal of the Experimental Analysis of Behavior, 87*, 325-336.
- Driscoll, M. P. (2005). Constructivism. *Psychology of learning for instruction, 384-410*.
- Dugatkin, L. A. (2002). Cooperation in animals: an evolutionary overview. *Biology and Philosophy, 17*, 459-476.

- Duker, P. C. (1976). Remotely applied punishment versus avoidance conditioning in the treatment of self-injurious behaviors. *European Journal of Behavioral Analysis and Modification*, 3, 179-185.
- Duker, P. C., & Seys, D. M. (2000). A quasi-experimental study on the effect of electrical aversion treatment on imposed mechanical restraint for severe self injurious behavior. *Research in Developmental Disabilities*, 21, 235-242.
- Dunham, P. J. (1971). Punishment: Method and theory. *Psychological Review*, 78, 58.
- Dunham, P. J. (1972). Some effects of punishment upon unpunished responding. *Journal of the Experimental Analysis of Behavior*, 17, 443-450.
- du Preez, K. P., Landon, J., Bellringer, M., Garrett, N., & Abbott, M. (2016). The effects of pop-up harm minimisation messages on electronic gaming machine gambling behaviour in New Zealand. *Journal of Gambling Studies*, 32, 1115-1126.
- Dura, J. R. (1991). Controlling extremely dangerous aggressive outbursts when functional analysis fails. *Psychological Reports*, 69, 451-459.
- Eckstein, R. A., & Hart, B. L. (1996). Treatment of canine acral lick dermatitis by behavior modification using electronic stimulation. *Journal of the American Animal Hospital Association*, 32, 225-230.
- Egas, M., & Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society of London B: Biological Sciences*, 275, 871-878.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Journal of Accounting Research*, 1-31.
- Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child & Adolescent Psychology*, 38, 439-450.
- Engelmann, D., & Nikiforakis, N. (2015). In the long-run we are all dead: On the benefits of peer punishment in rich environments. *Social Choice and Welfare*, 45, 561-577.
- Ernst, A. J., & Yee, R. W. (1975). A developmental analysis of the relationship between response strength and punishment. *Learning & Behavior*, 3, 107-113.
- Estes, W. K. (1944). An experimental study of punishment. *Psychological Monographs*, 57.
- Everly, J. B., & Perone, M. (2012). Suppressive and facilitative effects of shock intensity and interresponse times followed by shock. *Journal of the Experimental Analysis of Behavior*, 98, 311-340.

- Fabiano, G. A., Pelham, W. E., Coles, E. K., Gnagy, E. M., Chronis-Tuscano, A., & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention deficit/hyperactivity disorder. *Clinical Psychology Review, 29*, 129-140.
- Falk, J. L., & Burnidge, G. K. (1970). Fluid intake and punishment-attenuating drugs. *Physiology & Behavior, 5*, 199-202.
- Farthing, G. W., & Hearst, E. (1968). Generalization gradients of inhibition after different amounts of training. *Journal of the Experimental Analysis of Behavior, 11*, 743-752.
- Favell, J. E., & Lovaas, O. I. (1987). Protection for clients undergoing aversive/restrictive interventions. *Education and Treatment of Children, 311-32*.
- Feigin, V. L., Theadom, A., Barker-Collo, S., Starkey, N. J., McPherson, K., Kahan, M., ... & Jones, K. (2013). Incidence of traumatic brain injury in New Zealand: a population-based study. *The Lancet Neurology, 12*, 53-64.
- Filby, Y., & Appel, J. B. (1966). Variable-interval punishment during variable-interval reinforcement. *Journal of the Experimental Analysis of Behavior, 9*, 521-527.
- Fisher, K. (1987). Agreement reached on aversives. *APA Monitor, 18*, 24.
- Fisher, W. W., Piazza, C. C., & Roane, H. S. (Eds.). (2011). *Handbook of Applied Behavior Analysis*. Guilford Press.
- Fisher, W. W., Bowman, L. G., Thompson, R. H., Contrucci, S. A., Burd, L., & Alon, G. (1998). Reductions in self-injury produced by transcutaneous electrical nerve stimulation. *Journal of Applied Behavior Analysis, 31*, 493-496.
- Fisher, W. W., Piazza, C. C., Bowman, L. G., Kurtz, P. F., Sherer, M. R., & Lachman, S. R. (1994). A preliminary evaluation of empirically derived consequences for the treatment of pica. *Journal of Applied Behavior Analysis, 27*, 447-457.
- Fisher, W., Piazza, C., Cataldo, M., Harrell, R., Jefferson, G., & Conner, R. (1993). Functional communication training with and without extinction and punishment. *Journal of Applied Behavior Analysis, 26*, 23-36.
- Flanagan, B., Goldiamond, I., & Azrin, N. (1958). Operant stuttering: the control of stuttering behavior through response-contingent consequences. *Journal of the Experimental Analysis of Behavior, 1*, 173-177.
- Fleshler, M., & Hoffman, H.S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 5*, 529-530.
- Follick, M. J., & Knutson, J. F. (1978). Punishment of irritable aggression. *Aggressive Behavior, 4*, 1-17.
- Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of*

- the National Academy of Sciences of the United States of America*, 102, 7047-7049.
- Fowler, H., Hochhauser, M., & Wischner, G. J. (1981). Facilitating stimulus effects of reward and punishment in discrimination learning. *Learning & Behavior*, 9, 16-20.
- Fortes, I., Case, J. P., & Zentall, T. R. (2017). Pigeons, unlike humans, do not prefer near hits in a slot-machine-like task. *Behavioral Processes*, 138, 67-72.
- Fox, A. E., & Pietras, C. J. (2013). The effects of response-cost punishment on instructional control during a choice task. *Journal of the Experimental Analysis of Behavior*, 99, 346-361.
- Foxx, R. M., & Azrin, N. H. (1972). Restitution: A method of eliminating aggressive disruptive behavior of retarded and brain damaged patients. *Behavior Research and Therapy*, 10, 15-27.
- Foxx, R. M., & Azrin, N. H. (1973). The elimination of autistic self-stimulatory behavior by overcorrection. *Journal of Applied Behavior Analysis*, 6, 1-14.
- Foxx, R. M., & Martin, E. D. (1975). Treatment of scavenging behavior (coprophagy and pica) by overcorrection. *Behavior Research and Therapy*, 13, 153-162.
- Foxx, R. M., McMorrow, M. J., Bittle, R. G., & Bechtel, D. R. (1986). The successful treatment of a dually-diagnosed deaf man's aggression with a program that included contingent electric shock. *Behavior Therapy*, 17, 170-186.
- Foxx, R. M., & Meindl, J. (2007). The long term successful treatment of the aggressive/destructive behaviors of a preadolescent with autism. *Behavioral Interventions*, 22, 83-97.
- Foxx, R. M., Snyder, M. S., & Schroeder, F. (1979). A food satiation and oral hygiene punishment program to suppress chronic rumination by retarded persons. *Journal of Autism and Developmental Disorders*, 9, 399-412.
- Frick, P. J. (2012). *Conduct Disorders and Severe Antisocial Behavior*. Springer Science & Business Media.
- Friedel, J. E., DeHart, W. B., & Odum, A. L. (2017). The effects of 100 dB 1-kHz and 22 kHz tones as punishers on lever pressing in rats. *Journal of the Experimental Analysis of Behavior*, 107, 354-368.
- Furukawa, E., Alsop, B., Sowerby, P., Jensen, S., & Tripp, G. (2017). Evidence for increased behavioral control by punishment in children with attention-deficit hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, 58, 248-257.
- Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *Science*, 322, 1510-1510.

- Galbicka, G., & Branch, M. N. (1981). Selective punishment of interresponse times. *Journal of the Experimental Analysis of Behavior*, 35, 311-322.
- Galbicka, G., & Platt, J. R. (1984). Inter-response time punishment: a basis for shock maintained behavior. *Journal of the Experimental Analysis of Behavior*, 41, 291-308.
- Galbraith, D. A., Byrik, R. J. and Rutledge, J. T. (1970). An aversive conditioning approach to the inhibition of chronic vomiting. *Canadian Psychiatric Association Journal*, 15, 311-313.
- Gardner, W. I. (1969). Use of punishment procedures with the severely retarded: A review. *American Journal of Mental Deficiency*, 74, 86-103.
- Gaynor, S. T., & Shull, R. L. (2002). The generality of selective observing. *Journal of the Experimental Analysis of Behavior*, 77, 171-187.
- Geller, I. (1970). Effect of punishment on lever pressing maintained by food reward or brain stimulation. *Physiology & Behavior*, 5, 203-206.
- Gerhardt, P., Holmes, D.L., Alessandri, M., & Goodman, M. (1991). Social policy on the use of aversive interventions: Empirical, ethical, and legal considerations. *Journal of Autism and Developmental Disorders* 21, 265-277.
- Gershoff, E. T. (2002). Corporal punishment by parents and associated child behaviors and experiences: a meta-analytic and theoretical review. *Psychological Bulletin*, 128(4), 539.
- Gipson, C. D., Alessandri, J. J., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning & Behavior*, 37, 289-298.
- Gordon, R., Handleman, J. S., & Harris, S. L. (1986). The effects of contingent versus non contingent running on the out-of-seat behavior of an autistic boy. *Child & Family Behavior Therapy*, 8, 37-44.
- Grace, R. C., Bedell, M. A., & Nevin, J. A. (2002). Preference and resistance to change with constant-and variable-duration terminal links: Independence of reinforcement rate and magnitude. *Journal of the Experimental Analysis of Behavior*, 77(3), 233-255.
- Grant B. F., Dawson D. A. (2006). Introduction to the national epidemiologic survey on alcohol and related conditions. *Alcohol Research Health* 29, 74–78.
- Green, L., & Rachlin, H. (1996). Commitment using punishment. *Journal of the Experimental Analysis of Behavior*, 65, 593-601.
- Greene, R. J., & Hoats, D. L. (1971). Aversive tickling: A simple conditioning technique. *Behavior Therapy*, 2, 389-393.
- Greenway, A. P. (1975). Amount of training and reversal learning with singly presented

- stimuli in the rat. *Animal Learning & Behavior*, 3, 235-238.
- Grey, I. M., & Hastings, R. P. (2005). Evidence-based practices in intellectual disability and behavior disorders. *Current Opinion in Psychiatry*, 18, 469-475.
- Griffith, R. G. (1983). The Administrative Issues: An Ethical and Legal Perspective. In: Axelrod, Saul; Apsche, Jack, eds. *The Effects of Punishment on Human Behavior*. Orlando, Florida, Academic Press; 317-338.
- Grohmann, K., Dickomeit, M. J., Schmidt, M. J., & Kramer, M. (2013). Severe brain damage after punitive training technique with a choke chain collar in a German shepherd dog. *Journal of Veterinary Behavior: Clinical Applications and Research*, 8, 180-184.
- Grove, R. N., & Schuster, C. R. (1974). Suppression of cocaine self-administration by extinction and punishment. *Pharmacology Biochemistry and Behavior*, 2, 199-208.
- Guzmán, R. A., Rodríguez-Sickert, C., & Rowthorn, R. (2007). When in Rome, do as the Romans do: the coevolution of altruistic punishment, conformist learning, and cooperation. *Evolution and Human Behavior*, 28, 112-117.
- Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquistio, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis*, 31, 211-235.
- Hake, D. F., & Azrin, N. H. (1965). Conditioned punishment. *Journal of the Experimental Analysis of Behavior*, 8, 279-293.
- Hake, D. F., Azrin, N. H., & Oxford, R. (1967). The effects of punishment intensity on squirrel monkeys. *Journal of the Experimental Analysis of Behavior*, 10, 95-107.
- Halevy, G., Feldon, J., & Weiner, I. (1987). Resistance to extinction and punishment following training with shock and non-reinforcement: Failure to obtain cross tolerance. *The Quarterly Journal of Experimental Psychology*, 39, 147-160.
- Hamilton, D. I. (1982). The side effects of electric shock employed with mentally retarded persons. *Australia and New Zealand Journal of Developmental Disabilities*, 8, 133-140.
- Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function based interventions. *Journal of Applied Behavior Analysis*, 38, 51-65.
- Hanson, H. M., Witoslawski, J. J., & Campbell, E. H. (1967). Drug effects in squirrel monkeys trained on a multiple schedule with a punishment contingency. *Journal of the*

- Experimental Analysis of Behavior*, 10, 565-569.
- Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, 56, 51.
- Harris, S. L., & Ersner-Hershfield, R. (1978). Behavioral suppression of seriously disruptive behavior in psychotic and retarded patients: a review of punishment and its alternatives. *Psychological Bulletin*, 85, 1352.
- Harris, S. L., & Romanczyk, R. G. (1976). Treating self-injurious behavior of a retarded child by overcorrection. *Behavior Therapy*, 7, 235-239.
- Harris, S. L., & Wolchik, S. A. (1979). Suppression of self-stimulation: Three alternative strategies. *Journal of Applied Behavior Analysis*, 12, 185-198.
- Harvey, S. T., Boer, D., Meyer, L. H., & Evans, I. M. (2009). Updating a meta-analysis of intervention research with challenging behavior: Treatment validity and standards of practice. *Journal of Intellectual and Developmental Disability*, 34, 67-80.
- Hastings, R. P., Boulton, H. V., Monzani, L. C., & Tombs, A. K. (2004). Behavioral function effects on intervention acceptability and effectiveness for self-injurious behavior. *Research in Developmental Disabilities*, 25, 139-154.
- Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action*. Psychonomic Society.
- Heather, N. (2017). Q: Is Addiction a Brain Disease or a Moral Failing? A: Neither. *Neuroethics*, 10, 115-124.
- Helbing, D., Szolnoki, A., Perc, M., & Szabó, G. (2010). Punish, but not too hard: how costly punishment spreads in the spatial public goods game. *New Journal of Physics*, 12, 083005.
- Hendry, D. P., & Van-Toller, C. (1964). Fixed-ratio punishment with continuous reinforcement. *Journal of the Experimental Analysis of Behavior*, 7, 293-300.
- Herman, R. L., & Azrin, N. H. (1964). Punishment by noise in an alternative response situation. *Journal of the Experimental Analysis of Behavior*, 7, 185-188.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267-272.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243-266.
- Herrnstein, R. J., & Prelec, D. (1991). Melioration: A theory of distributed choice. *The Journal of Economic Perspectives*, 137-156.
- Herrnstein, R. J., & Vaughan, W. (1980). Melioration and behavioral allocation. *Limits to*

- action: The allocation of individual behavior*, 143-176.
- Herrnstein, R. J., Prelec., D., & Vaughan, W. (1986). *An intra-personal prisoners' dilemma*. Paper presented at the IX Symposium on Quantitative Analysis of Behavior: Behavioral Economics, Harvard University.
- Herron, M. E., Shofer, F. S., & Reisner, I. R. (2009). Survey of the use and outcome of confrontational and non-confrontational training methods in client-owned dogs showing undesired behaviors. *Applied Animal Behavior Science*, *117*, 47-54.
- Heyman, G.M. (1996). Resolving the contradictions of addiction. *Behavioral and Brain Sciences*, *19*, 193-203.
- Heyman, G. M. (2013). Addiction and Choice: Theory and New Data. *Frontiers in Psychiatry*, *4*, 31.
- Heyman, G. M., & Dunn, B. (2002). Decision biases and persistent illicit drug use: an experimental study of distributed choice and addiction. *Drug and Alcohol Dependence*, *67*, 193-203.
- Heyman, G. M., & Tanz, L. (1995). How to teach a pigeon to maximize overall reinforcement rate. *Journal of the Experimental Analysis of Behavior*, *64*, 277-297.
- Heyvaert, M., Saenen, L., Campbell, J. M., Maes, B., & Onghena, P. (2014). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: An updated quantitative synthesis of single-subject research. *Research in Developmental Disabilities*, *35*, 2463-2476.
- Hiby, E. F., Rooney, N. J., & Bradshaw, J. W. S. (2004). Dog training methods: their use, effectiveness and interaction with behavior and welfare. *Animal Welfare* *13*, 63-70.
- Hineline, Philip N., & Rosales-Ruiz, J. (2013). Behavior in relation to aversive events: Punishment and negative reinforcement. In G.J. Madden (Ed.), *APA Handbook of Behavior Analysis: Vol. 1. Methods and Principles*, American Psychological Association.
- Hinnenkamp, J.E., Shahan, T.A., & Madden, G.J. (2017). How suboptimal is suboptimal choice? *Journal of the Experimental Analysis of Behavior*, *107*, 136-150.
- Hodgson, K., Hutchinson, A. D., & Denson, L. (2014). Nonpharmacological treatments for ADHD: a meta-analytic review. *Journal of Attention Disorders*, *18*, 275-282.
- Holland, P. C., & Morell, J. R. (1996). The effects of intertrial and feature–target intervals on operant serial feature negative discrimination learning. *Learning and Motivation*, *27*, 21-42.
- Holland, P. C., Asem, J. S., Galvin, C. P., Keeney, C. H., Hsu, M., Miller, A., & Zhou, V.

- (2014). Blocking in autoshaped lever-pressing procedures with rats. *Learning & Behavior*, *42*, 1-21.
- Holtz, N. A., Anker, J. J., Regier, P. S., Claxton, A., & Carroll, M. E. (2013). Cocaine self administration punished by iv histamine in rat models of high and low drug abuse vulnerability: effects of saccharin preference, impulsivity, and sex. *Physiology & Behavior*, *122*, 32-38.
- Holtz, N. A., & Carroll, M. E. (2015). Cocaine self-administration punished by intravenous histamine in adolescent and adult rats. *Behavioral Pharmacology*, *26*, 393.
- Holz, W. C. (1968). Punishment and rate of positive reinforcement. *Journal of the Experimental Analysis of Behavior*, *11*, 285-292.
- Holz, W. C., & Azrin, N. H. (1961). Discriminative properties of punishment. *Journal of the Experimental Analysis of Behavior*, *4*, 225-232.
- Holz, W. C., & Azrin, N. H. (1962). Interactions between the discriminative and aversive properties of punishment. *Journal of the Experimental Analysis of Behavior*, *5*, 229-234.
- Honig, W. K. (1966). The role of discrimination training in the generalization of punishment. *Journal of the Experimental Analysis of Behavior*, *9*, 377-384.
- Honig, W. K., & Slivka, R. M. (1964). Stimulus generalization of the effects of punishment. *Journal of the Experimental Analysis of Behavior*, *7*, 21-25.
- Honig, W. K., Boneau, A., Burstein, R., & Pennypacker, S. (1963). Positive and negative generalization gradients obtained after equivalent training conditions. *Journal of Comparative Physiology*, *56*, 111-116.
- Hopf, F. W., Chang, S. J., Sparta, D. R., Bowers, M. S., & Bonci, A. (2010). Motivation for alcohol becomes resistant to quinine adulteration after 3 to 4 months of intermittent alcohol self-administration. *Alcoholism: Clinical and Experimental Research*, *34*, 1565-1573.
- Horner, R. H., Dunlap, G., Koegel, R. L., Carr, E. G., Sailor, W., Anderson, J., ... & O'Neill, R. E. (2005). Toward a technology of "nonaversive" behavioral support. *Research and Practice for Persons with Severe Disabilities*, *30*, 3-10.
- Humphreys, L. G. (1939). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *Journal of Experimental Psychology General*, *25*, 141-158.
- Hustad, J. T., Short, E. E., Borsari, B., Barnett, N. P., Tevyaw, T. O. L., & Kahler, C. W.

- (2011). College alcohol citations result in modest reductions in student drinking. *Journal of Substance Abuse Treatment, 40*, 281-286.
- Hymowitz, N. (1979). Suppression of responding during signaled and unsignaled shock. *Psychological Bulletin, 86*, 175.
- Ilango, A., Wetzell, W., Scheich, H., & Ohl, F. W. (2010). The combination of appetitive and aversive reinforcers and the nature of their interaction during auditory learning. *Neuroscience, 166*, 752-762.
- Janssen, M. A., & Bushman, C. (2008). Evolution of cooperation and altruistic punishment when retaliation is possible. *Journal of Theoretical Biology, 254*, 541-545.
- Jenkins, H.M. & Boakes, R.A. (1973). Observing stimulus sources that signal food or no food. *Journal of the Experimental Analysis of Behavior, 20*, 197–207.
- Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 365*, 2635-2650.
- Johnson, W. L., Baumeister, A. A., Penland, M. J., & Inwald, C. (1982). Experimental analysis of self-injurious, stereotypic, and collateral behavior of retarded persons: Effects of overcorrection and reinforcement of alternative responding. *Analysis and Intervention in Developmental Disabilities, 2*, 41-66.
- Johnston, J. M. (1991). What can behavior analysis learn from the aversives controversy? *The Behavior Analyst, 14*, 187-196.
- Jonckheer-Sheehy, V. S., Delesalle, C. J., van den Belt, A. J. M., & Van den Boom, R. (2012). Bad behavior or a physical problem? Rearing in a Dutch Warmblood mare. *Journal of Veterinary Behavior: Clinical Applications and Research, 7*, 380-385.
- Jones, B., & Davison, M. (1997). The control of choice by its consequences. *Journal of the Experimental Analysis of Behavior, 68*, 329–348.
- Jones, C. M., Campopiano, M., Baldwin, G., & McCance-Katz, E. (2015). National and state treatment need and capacity for opioid agonist medication-assisted treatment. *American Journal of Public Health.*
- Jonker, N. C., Ostafin, B. D., Glashouwer, K. A., van Hemel-Ruiter, M. E., & de Jong, P. J. (2014). Reward and punishment sensitivity and alcohol use: the moderating role of executive control. *Addictive Behaviors, 39*, 945-948.
- Kahng, S., Iwata, B. A., & Lewin, A. B. (2002). Behavioral treatment of self-injury, 1964 to 2000. *American Journal on Mental Retardation, 107*, 212-221.
- Kampman, K., & Jarvis, M. (2015). American Society of Addiction Medicine (ASAM) National Practice Guideline for the use of medications in the treatment of addiction

- involving opioid use. *Journal of Addiction Medicine*, 9, 358.
- Kaufman, A., & Baron, A. (1968). Suppression of behavior by timeout punishment when suppression results in loss of positive reinforcement. *Journal of the Experimental Analysis of Behavior*, 11, 595-607.
- Kazdin, A. E. (1980a). Acceptability of time out from reinforcement procedures for disruptive child behavior. *Behavior Therapy*, 11, 329-344.
- Kazdin, A. E. (1980b). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis*, 13, 259-273.
- Kelleher, R. T., & Morse, W. H. (1968). Schedules using noxious stimuli: Responding maintained with response-produced electric shocks. *Journal of the Experimental Analysis of Behavior*, 11, 819-838.
- Kelly, J. A., & Drabman, R. S. (1977). Overcorrection: An effective procedure that failed. *Journal of Clinical Child & Adolescent Psychology*, 6, 38-40.
- Kemp, F. (1996). The ideology of aversive treatment as applied to clients and colleagues. *Child and Family Behavior Therapy*, 18, 9-27.
- Kendall, S.B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, 21, 463-473.
- Kendall, S.B. (1985). A further study of choice and percentage reinforcement. *Behavioral Processes*, 10(4), 399-413.
- Killeen, P. R., & Jacobs, K. W. (2017). Coal is not black, snow is not white, food is not a reinforcer: the roles of affordances and dispositions in the analysis of behavior. *The Behavior Analyst*, 40, 17-38.
- Landon, J., Palmer du Preez, K., Bellringer, M., Page, A., & Abbott, M. (2016). Pop-up messages on electronic gaming machines in New Zealand: Experiences and views of gamblers and venue staff. *International Gambling Studies*, 16, 49-66.
- Landsberg, G. (1990). A veterinarian's guide to the correction of dominance aggression. *The Canadian Veterinary Journal*, 31, 121.
- Laude, J. R., Beckmann, J. S., Daniels, C. W., & Zentall, T. R. (2014). Impulsivity affects suboptimal gambling-like choice by pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40, 2.
- Laude, J. R., Pattison, K. F., & Zentall, T. R. (2012). Hungry pigeons make suboptimal choices, less hungry pigeons do not. *Psychonomic Bulletin & Review*, 19, 884-891.
- Laurence, M. T., Himeline, P. N., & Bersh, P. J. (1994). The puzzle of responding maintained by response-contingent shock. *Journal of the Experimental Analysis of*

- Behavior*, 61, 135-153.
- LaVigna, G. W., & Donnellan, A. M. (1986). *Alternatives to punishment: Solving behavior problems with non-aversive strategies*. Ardent Media.
- Lenoir, M., Serre, F., Cantin, L., & Ahmed, S. H. (2007). Intense sweetness surpasses cocaine reward. *PloS one*, 2, 698.
- Leong, K. C., Berini, C. R., Ghee, S. M., & Reichel, C. M. (2016). Extended cocaine seeking produces a shift from goal-directed to habitual responding in rats. *Physiology & Behavior*, 164, 330-335.
- Lerman, D. C., & Vorndran, C. M. (2002). On the status of knowledge for using punishment: Implications for treating behavior disorders. *Journal of Applied Behavior Analysis*, 35, 431-464.
- Lerman, D. C., Iwata, B. A., Shore, B. A., & DeLeon, I. G. (1997). Effects of intermittent punishment on self - injurious behavior: an evaluation of schedule thinning. *Journal of Applied Behavior Analysis*, 30, 187-201.
- Lesscher, H., Van Kerkhof, L. W., & Vanderschuren, L. J. (2010). Inflexible and indifferent alcohol drinking in male mice. *Alcoholism: Clinical and Experimental Research*, 34, 1219-1225.
- Lewis, P., Lewin, L., Muehleisen, P., & Stoyak, M. (1974). Preference for signalled reinforcement. *Journal of the Experimental Analysis of Behavior*, 22, 143-150.
- Li, K. C., Hsiao, S., & Li, J. S. (2013). Conditioned taste aversion as instrumental punishment. *Journal of Experimental Psychology: Animal Behavior Processes*, 39, 294.
- Libet, J. M., Sajwaj, T., & Agras, W. S. (1973). Elimination of persistent vomiting in an infant by response-contingent punishment using lemon juice. In *Proceedings of the Annual Convention of the American Psychological Association*. American Psychological Association.
- Lichstein, K. L., & Schreibman, L. (1976). Employing electric shock with autistic children. *Journal of Autism and Developmental Disorders*, 6, 163-173.
- Linden, D. R. (1976). The effect of intensity and frequency of intermittent punishment in acquisition on resistance to extinction of an approach response in the rat. *Learning & Behavior*, 4, 451-456.
- Linscheid, T. R., & Reichenbach, H. (2002). Multiple factors in the long-term effectiveness of contingent electric shock treatment for self-injurious behavior: A case

- example. *Research in Developmental Disabilities*, 23, 161-177.
- Linscheid, T. R., Pejeau, C., Cohen, S., & Footo-Lenz, M. (1994). Positive side effects in the treatment of SIB using the self-injurious behavior inhibiting system (SIBIS): Implications for operant and biochemical explanations of SIB. *Research in Developmental Disabilities*, 15, 81-90.
- Lohrmann-O'Rourke, S., & Zirkel, P. A. (1998). The case law on aversive interventions for students with disabilities. *Exceptional Children*, 65, 101.
- Luckey, R. E., Watson, C. M. and Musick, J. K. (1968) Aversive conditioning as a means of inhibiting vomiting and rumination. *American Journal of Mental Deficiency*, 73, 139-142.
- Ludwig, A. M., Marx, A. J., Hill, P. A., & Browning, R. M. (1969). The control of violent behavior through faradic shock: a case study. *The Journal of Nervous and Mental Disease*, 148, 624-637.
- Lydon, S., Healy, O., O'Reilly, M., & McCoy, A. (2013). A systematic review and evaluation of response redirection as a treatment for challenging behavior in individuals with developmental disabilities. *Research in Developmental Disabilities*, 34, 3148-3158.
- MacDonald, L. (1973). The relative aversiveness of signaled versus unsignalled shock punishment. *Journal of the Experimental Analysis of Behavior*, 20, 37-46.
- Mace, F. C., McComas, J. J., Mauro, B. C., Progar, P. R., Taylor, B., Ervin, R., & Zangrillo, A. N. (2010). Differential reinforcement of alternative behavior increases resistance to extinction: Clinical demonstration, animal modeling, and clinical test of one solution. *Journal of the Experimental Analysis of Behavior*, 93, 349-367.
- Mackintosh, N. J. (1965). Selective attention in animal discrimination learning. *Psychological Bulletin*, 64, 124.
- Mackintosh, N. J. (1974). *The Psychology of Animal Learning*. London: Academic Press.
- Margolius, G. (1955). Stimulus generalization of an instrumental response as a function of the number of reinforced trials. *Journal of Experimental Psychology*, 49, 105.
- Marschark, E. D., & Baenninger, R. (2002). Modification of instinctive herding dog behavior using reinforcement and punishment. *Anthrozoös*, 15, 51-68.
- Martin, L. E., Cox, L. S., Brooks, W. M., & Savage, C. R. (2014). Winning and losing: differences in reward and punishment sensitivity between smokers and nonsmokers. *Brain and Behavior*, 4, 915-924.
- Martin, J. R., Moreau, J. L., Jenck, F., & Cumin, R. (1993). Acute and chronic

- administration of buspirone fails to yield anxiolytic-like effects in a mouse operant punishment paradigm. *Pharmacology Biochemistry and Behavior*, *46*, 905-910.
- Martínez, M., Alba, R., Rodríguez, W., & Orduña, V. (2017). Incentive salience attribution is not the sole determinant of suboptimal choice in rats: conditioned inhibition matters. *Behavioral Processes*, *142*, 99-105.
- Mason, G., & Rushen, J. (Eds.). (2008). *Stereotypic Animal Behavior: Fundamentals and Applications to Welfare*. Cabi.
- Matson, J. L., & DiLorenzo, T. M. (1984). *Punishment and its alternatives: A new perspective for behavior modification*, 13, Springer Publishing Company.
- Matson, J. L., & Stephens, R. M. (1977). Overcorrection of aggressive behavior in a chronic psychiatric patient. *Behavior Modification*, *1*, 559-564.
- Matson, J. L., Sipes, M., Fodstad, J. C., & Fitzgerald, M. E. (2011). Issues in the management of challenging behaviors of adults with autism spectrum disorder. *CNS Drugs*, *25*, 597-606.
- Matson, J. L., Stephens, R. M., & Smith, C. (1978). Treatment of self-injurious behavior with overcorrection. *Journal of Intellectual Disability Research*, *22*, 175-178.
- May, J. G., McAllister, J., Risley, T., Twardosz, S., & Cox, C. H. (1974). Florida guidelines for the use of behavioral procedures in state programs for the retarded. *Erhältlich von Jack May, Department of Psychology, Florida State University, Tallahassee, Florida, 32306*.
- Mazur, J. E. (2015). *Learning and Behavior*. Psychology Press.
- McFarlain, R. A., Andy, O. J., Scott, R. W., & Wheatley, M. L. (1975). Suppression of headbanging on the ward. *Psychological Reports*, *36*, 315-321.
- McGreevy, P. D., & McLean, A. N. (2009). Punishment in horse-training and the concept of ethical equitation. *Journal of Veterinary Behavior: Clinical Applications and Research*, *4*, 193-197.
- McHugh, R. K., Whitton, S. W., Peckham, A. D., Welge, J. A., & Otto, M. W. (2013). Patient preference for psychological vs. pharmacological treatment of psychiatric disorders: a meta analytic review. *The Journal of Clinical Psychiatry*, *74*, 595.
- McKenzie, S. D., Smith, R. G., Simmons, J. N., & Soderlund, M. J. (2008). Using a stimulus correlated with reprimands to suppress automatically maintained eye poking. *Journal of Applied Behavior Analysis*, *41*, 255-259.
- McKinney, C., & Renk, K. (2011). Atypical antipsychotic medications in the management of disruptive behaviors in children: safety guidelines and recommendations. *Clinical*

- Psychology Review*, 31, 465-471.
- Mestre-Bach, G. Granero, R., Steward, T., ...et al. (2016). Reward and punishment sensitivity in women with gambling disorder or compulsive buying: Implications in treatment outcome *Journal of Behavioral Addictions*, 5, 658-665.
- Miller, N. E. (1960). Learning resistance to pain and fear: Effects of overlearning exposure and rewarded exposure in context. *Journal of Experimental Psychology*, 60, 137-145.
- Miller, N. E., & Dollard, J. (1941) Social Learning. "Imitation." *Yale University, New Haven*.
- Molet, M., Miller, H.C., Laude, J.R., Kirk, C., Manning, B., & Zentall, T.R. (2012). Decision making by humans in a behavioral task: Do humans, like pigeons, show suboptimal choice? *Learning and Behavior*, 40, 439-447.
- Moon, L., & Lewis, P. (1975). Effect of signal duration on preference for signalled reinforcement. *Psychological Reports*, 37, 263-269.
- Morris, D. H., Treloar, H., Tsai, C.,...et al. (2016). Acute subjective response to alcohol as a function of reward and punishment sensitivity, *Addictive Behaviors*, 60, 90-96.
- Mulvaney, D. E., Dinsmoor, J. A., Jwaideh, A. R., & Hughes, L. H. (1974). Punishment of observing by the negative discriminative stimulus. *Journal of the Experimental Analysis of Behavior*, 21, 37-44.
- Munavi, K. M. K. (1970). *Effects of Contigent Punishment on Response Probabilities of Unpunished Operants* (Doctoral dissertation, Dalhousie University).
- Murphy, G., Beadle-Brown, J., Wing, L., Gould, J., Shah, A., & Homes, N. (2005). Chronicity of challenging behaviors in people with severe intellectual disabilities and/or autism: A total population sample. *Journal of Autism and Developmental Disorders*, 35, 405-418.
- Muttar, A. K., Peck, D., Whitlow, O. & Fraser, W. (1975). Reversal of a severe case of self mutilation. *Journal of Mental Deficiency Research*, 19, 3-9.
- Muttar, A. K., Peck, D., Whitlow, D., & Fraser, W. (1975). Reversal of a severe case of self mutilation. *Journal of Intellectual Disability Research*, 19, 3-9.
- Myer, J. S. (1967). Prior killing experience and the suppressive effects of punishment on the killing of mice by rats. *Animal Behavior*, 15, 59-61.
- Myers, C. E., Rego, J., Haber, P., Morley, K., Beck, K. D., Hogarth, L., & Moustafa, A. A. (2017). Learning and generalization from reward and punishment in opioid addiction. *Behavioral Brain Research*, 317, 122-131.
- Myers, C. E., Sheynin, J., Balsdon, T., Luzardo, A., Beck, K. D., Hogarth, L., ... &

- Moustafa, A. (2016). Probabilistic reward-and punishment-based learning in opioid addiction: experimental and computational data. *Behavioral Brain Research*, *296*, 240-248.
- Myers, D.L., & Myers, L.E. (1977). Undermatching: a reappraisal of performance on concurrent variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *27*, 203-214.
- Nakatani, Y., Matsumoto, Y., Mori, Y., Hirashima, D., Nishino, H., Arikawa, K., & Mizunami, M. (2009). Why the carrot is more effective than the stick: Different dynamics of punishment memory and reward memory and its possible biological basis. *Neurobiology of Learning and Memory*, *92*, 370-380.
- Negus, S. S. (2005). Effects of punishment on choice between cocaine and food in rhesus monkeys. *Psychopharmacology*, *181*, 244-252.
- Netherland, J. (Ed.). (2012). *Critical Perspectives on Addiction* (Vol. 14). Emerald Group Publishing.
- Nevin, J.A. (1979). Reinforcement schedules and response strength. In M.D. Zeiler & P. Harzem (Eds.), *Advances in Analysis of Behavior: Vol. 1. Reinforcement and the organisation of behavior*, p. 117-158. Chichester, England: Wiley.
- Nevin, J.A., & Grace, R.C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, *23*, 73-120.
- Newsom, C., Favell, J. E., & Rincover, A. (1983). The side effects of punishment. *The Effects of Punishment on Human Behavior*, 285-316.
- O'Brien, F. (1989). Punishment for people with developmental disabilities. In E. Cipani (Ed.), *The treatment of severe behavior disorders: Behavior analysis approaches*, 37-58. Washington, DC: American Association of Mental Retardation.
- Ogata, N., & Dodman, N. H. (2011). The use of clonidine in the treatment of fear-based behavior problems in dogs: an open trial. *Journal of Veterinary Behavior: Clinical Applications and Research*, *6*, 130-137.
- Orme-Johnson, D. W., & Yarczower, M. (1974). Conditioned suppression, punishment, and aversion. *Journal of the Experimental Analysis of Behavior*, *21*, 57-74.
- Orsini, C. A., Willis, M. L., Gilbert, R. J., Bizon, J. L., & Setlow, B. (2016). Sex differences in a rat model of risky decision making. *Behavioral Neuroscience*, *130*, 50.
- Pattison, K. F., Laude, J. R., & Zentall, T. R. (2013). Environmental enrichment affects suboptimal, risky, gambling-like choice by pigeons. *Animal Cognition*, *16*, 429-434.

- Pellón, R., Mas, B., & Blackman, D. E. (1992). Effects of d-amphetamine and of diazepam on non-punished and punished schedule-induced drinking in rats. *Behavioral Pharmacology*, *3*, 75-81.
- Pelloux, Y., Everitt, B. J., & Dickinson, A. (2007). Compulsive drug seeking by rats under punishment: effects of drug taking history. *Psychopharmacology*, *194*, 127-137.
- Pelloux, Y., Murray, J. E., & Everitt, B. J. (2015). Differential vulnerability to the punishment of cocaine related behaviors: effects of locus of punishment, cocaine taking history and alternative reinforcer availability. *Psychopharmacology*, *232*, 125-134.
- Perone, M. (2003). Negative effects of positive reinforcement. *The Behavior Analyst*, *26*, 1-14.
- Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*, 60-69.
- Petscher, E. S., Rey, C., & Bailey, J. S. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities*, *30*, 409-425.
- Picker, M., Poling, A., & Parker, A. (1979). A review of children's self-injurious behavior. *The Psychological Record*, *29*, 435-452.
- Pietras, C. J., Brandt, A. E., & Searcy, G. D. (2010). Human responding on random interval schedules of response-cost punishment: The role of reduced reinforcement density. *Journal of the Experimental Analysis of Behavior*, *93*, 5-26.
- Pietras, C. J., & Hackenberg, T. D. (2005). Response-cost punishment via token loss with pigeons. *Behavioral Processes*, *69*, 343-356.
- Podlesnik, C. A., & Jimenez-Gomez, C. (2013). Punishing and cardiovascular effects of intravenous histamine in rats: Pharmacological selectivity. *Journal of the Experimental Analysis of Behavior*, *100*, 333-354.
- Podlesnik, C. A., & Kelley, M. E. (2015). Translational research on the relapse of operant behavior. *Revista Mexicana de Análisis de la Conducta*, *41*, 226-251.
- Podlesnik, C. A., & Kelley, M. E. (2017). Beyond Intervention: Shaping Policy for Addressing Persistence and Relapse of Severe Problem Behavior. *Policy Insights from the Behavioral and Brain Sciences*, *4*, 17-24.
- Podlesnik, C. A., Jimenez-Gomez, C., & Woods, J. H. (2010). A choice procedure to

- assess the aversive effects of drugs in rodents. *Journal of the Experimental Analysis of Behavior*, 93, 203-223.
- Podlesnik, C. A., Jimenez-Gomez, C., Ward, R. D., & Shahan, T. A. (2009). Resistance to change and frequency of response-dependent stimuli uncorrelated with reinforcement. *Journal of the Experimental Analysis of Behavior*, 92, 199-214.
- Podlesnik, C. A., Kelley, M. E., Jimenez-Gomez, C., & Bouton, M. E. (2017). Renewed behavior produced by context change and its implications for treatment maintenance: A review. *Journal of Applied Behavior Analysis*, 50, 675-697.
- Poling, A., & Thompson, T. (1977). Attenuation of ethanol intake by contingent punishment of food-maintained responding. *Pharmacology Biochemistry and Behavior*, 7, 393-399.
- Polsky, R. H. (1994). Electronic shock collars: are they worth the risks? *Journal of the American Animal Hospital Association*, 30, 463-468.
- Powell, J., & Azrin, N. (1968). The effects of shock as a punisher for cigarette smoking. *Journal of Applied Behavior Analysis*, 1, 63-71.
- Prendergast, M., Podus, D., Finney, J., Greenwell, L., & Roll, J. (2006). Contingency management for treatment of substance use disorders: A meta-analysis. *Addiction*, 101, 1546-1560.
- Prochaska, J., Smith, N., Marzilli, R., Colby, J., & Donovan, W. (1974). Remote-control aversive stimulation in the treatment of head-banging in a retarded child. *Journal of Behavior Therapy and Experimental Psychiatry*, 5, 285-289.
- Quist, R. W., & Martin, R. R. (1967). The effect of response contingent verbal punishment on stuttering. *Journal of Speech and Hearing Research*, 10, 795.
- Rachlin, H. (1966). Recovery of responses during mild punishment. *Journal of the Experimental Analysis of Behavior*, 9, 251-263.
- Rachlin, H. (1997). Four teleological theories of addiction. *Psychonomic Bulletin & Review*, 4, 462-473.
- Rapoff, M. A., Altman, K., & Christophersen, E. R. (1980). Elimination of a retarded blind child's self-hitting by response-contingent brief restraint. *Education and Treatment of Children*, 231-237.
- Read, D., & Roelofsma, P. (1999). Hard choices and weak wills: the theory of intrapersonal dilemmas. *Philosophical Psychology*, 12, 341-356.
- Reed, P., & Doughty, A. H. (2005). Within-subject testing of the signaled-reinforcement effect on operant responding as measured by response rate and resistance to change.

- Journal of the Experimental Analysis of Behavior*, 83, 31-45.
- Reed, P., & Yoshino, T. (2008). Effect of contingent auditory stimuli on concurrent schedule performance: An alternative punisher to electric shock. *Behavioral Processes*, 78, 421-428.
- Reichow, B. (2012). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42, 512-520.
- Repp, A. C., & Singh, N. N. (1990). *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities*. Sycamore Publishing Company.
- Ricketts, R. W., Goza, A. B., Matese, M. (1993). A 4-year follow-up of treatment of self injury. *Journal of Behavior Therapy & Experimental Psychiatry*, 24, 57-62.
- Riccio, D. C., & Hamm, R. J. (1972). Increased resistance to punishment produced by adaptation to a cold stressor. *Physiology & Behavior*, 9, 209-212.
- Risley, T. R. (1968). The effects and side effects of punishing the autistic behaviors of a deviant child. *Journal of Applied Behavior Analysis*, 1, 21-34.
- Roberts, C. L., & I, K. (1971). Elicitation and punishment of intraspecies aggression by the same stimulus. *Journal of the Experimental Analysis of Behavior*, 15, 193-196.
- Rockenbach, B., & Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444, 718-723.
- Rodriguez, W. A., & Logan, F. A. (1980). Preference for punishment of the instrumental or the consummatory response. *Learning & Behavior*, 8, 116-119.
- Roll, A., & Unshelm, J. (1997). Aggressive conflicts amongst dogs and factors affecting them. *Applied Animal Behavior Science*, 52, 229-242.
- Rooney, N. J., & Cowan, S. (2011). Training methods and owner–dog interactions: Links with dog behavior and learning ability. *Applied Animal Behavior Science*, 132, 169- 177.
- Roper, K. L., & Baldwin, E. R. (2004). The two-alternative observing response procedure in rats: Preference for nondiscriminative stimuli and the effect of delay. *Learning and Motivation*, 35, 275-302.
- Rusch, F., Close, D., Hops, H., & Agosta, J. (1976). Overcorrection: Generalization and maintenance. *Journal of Applied Behavior Analysis*, 9, 498-498.
- Sajwaj, T., Libet, J., & Agras, S. (1974). Lemon-juice therapy: the control of life threatening rumination in a six-month old infant. *Journal of Applied Behavior Analysis*, 7, 557-563.

- Salvy, S. J., Mulick, J. A., Butter, E., Bartlett, R. K., & Linscheid, T. R. (2004). Contingent electric shock (SIBIS) and a conditioned punisher eliminate severe head banging in a preschool child. *Behavioral Interventions, 19*, 59-72.
- Sargisson, R. J., Butler, R., & Elliffe, D. (2011). An evaluation of the Aboistop citronella spray collar as a treatment for barking of domestic dogs. *ISRN Veterinary Science*, Article ID 759379, <http://dx.doi.org/10.5402/2011/759379>
- Sawyer, A. T., Liverant, G. I., Jun, J. J., Lee, D. J., Cohen, A. L., Dutra, S. J., ... & Sloan, D. M. (2016). Punishment learning in US veterans with posttraumatic stress disorder. *Journal of Traumatic Stress, 29*, 374-378.
- Schalke, E., Stichnoth, J., Ott, S., & Jones-Baade, R. (2007). Clinical signs caused by the use of electric training collars on dogs in everyday life situations. *Applied Animal Behavior Science, 105*, 369-380.
- Schilder, M. B., & van der Borg, J. A. (2004). Training dogs with help of the shock collar: short and long term behavioral effects. *Applied Animal Behavior Science, 85*, 319-334.
- Schroeder, S. R., Mulick, J. A., & Schroeder, C. S. (1979). Management of severe behavior problems of the retarded. *Handbook of Mental Deficiency*, 341-366.
- Schuster, R., & Rachlin, H. (1968). Indifference between punishment and free shock: Evidence for the negative law of effect. *Journal of the Experimental Analysis of Behavior, 11*, 777-786.
- Scobie, S. R., & Kaufman, A. (1969). Intermittent punishment of human responding maintained by intermittent reinforcement. *Journal of the Experimental Analysis of Behavior, 12*, 137-147.
- Selekman, W. (1973). Behavioral contrast and inhibitory stimulus control as related to extended training. *Journal of the Experimental Analysis of Behavior, 20*, 245-252.
- Seligman, M. E. (1968). Chronic fear produced by unpredictable electric shock. *Journal of Comparative and Physiological Psychology, 66*, 402.
- Sewell, W. R., Nickel, R., Johnson, J. T., & Williamson, D. (1979). Effect of length of training on stimulus control using a low-density reinforcement schedule. *The Psychological Record, 29*, 401-408.
- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior, 93*, 269-289.
- Shahan, T. A. (2017). Moving Beyond Reinforcement and Response Strength. *The Behavior Analyst, 40*, 107-121.
- Shahan, T. A., & Cunningham, P. (2015). Conditioned reinforcement and information

- theory reconsidered. *Journal of the Experimental Analysis of Behavior*, *103*, 405-418.
- Shahan, T. A., & Podlesnik, C. A. (2008). Quantitative analysis of observing and attending. *Behavioral Processes*, *78*, 145-157.
- Sharma, A., & Shaw, S. R. (2012). Efficacy of risperidone in managing maladaptive behaviors for children with autistic spectrum disorder: a meta-analysis. *Journal of Pediatric Health Care*, *26*, 291-299.
- Shimp, K. G., Mitchell, M. R., Beas, B. S., Bizon, J. L., & Setlow, B. (2015). Affective and cognitive mechanisms of risky decision making. *Neurobiology of Learning and Memory*, *117*, 60-70.
- Sidman, M. (1989). *Coercion and its fallout*. Boston: Authors Cooperative.
- Sierra-Mercado, D., Deckersbach, T., Arulpragasam, A. R., Chou, T., Rodman, A. M., Duffy, A., ... & Eskandar, E. N. (2015). Decision making in avoidance–reward conflict: a paradigm for non-human primates and humans. *Brain Structure and Function*, *220*, 2509-2517.
- Silva, F. J., Silva, K., & Pear, J. J. (1992). Sign-versus goal-tracking: effects of conditioned stimulus-to-unconditioned-stimulus distance. *Journal of the Experimental Analysis of Behavior*, *57*, 17-31.
- Silverman, P. J. (1971). Chained and tandem fixed-interval schedules of punishment. *Journal of the Experimental Analysis of Behavior*, *16*, 1-13.
- Simmons, J. Q., & Lovaas, O. I. (1969). Use of pain and punishment as treatment techniques with childhood schizophrenics. *American Journal of Psychotherapy*, *23*, 23-36.
- Singer, G. H., Gert, B., & Koegel, R. L. (1999). A moral framework for analyzing the controversy over aversive behavioral interventions for people with severe mental retardation. *Journal of Positive Behavior Interventions*, *1*, 88-100.
- Sisson, L. A., & Hersen, M. (1993). Improving the performance of youth with dual sensory impairment: Analyses and social validation of procedures to reduce maladaptive responding in vocational and leisure settings. *Behavior Therapy*, *24*, 553-571.
- Sizemore, O. J., & Maxwell, F. R. (1983). Differential punishment of other behavior: Effects of pause-dependent shock on variable-interval performance. *Animal Learning & Behavior*, *11*, 266-274.
- Skiba, R. J., & Deno, S. L. (1991). Terminology and behavior reduction: The case against “punishment.”. *Exceptional Children*.

- Skinner, B.F. (1953). *Science and Human Behavior*. New York: Macmillan.
- Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 212.
- Smith, B. H., Abramson, C. I., & Tobin, T. R. (1991). Conditional withholding of proboscis extension in honeybees (*Apis mellifera*) during discriminative punishment. *Journal of Comparative Psychology*, *105*, 345.
- Smith, J. M., & Harper, D. (2003). *Animal Signals*. Oxford University Press.
- Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcers and automatic punishment in infant vocal behavior. *The Analysis of Verbal Behavior*, *13*, 39.
- Spencer, D., Marshall, J., Post, B., Kulakodlu, M., Newschaffer, C., Dennen, T., ... & Jain, A. (2013). Psychotropic medication use and polypharmacy in children with autism spectrum disorders. *Pediatrics*, *132*, 833-840.
- Spetch, M.L., Belke, T.W., Barnet, R.C., Dunn, R., & Pierce, W.D. (1990). Suboptimal choice in a percentage-reinforcement procedure: Effects of signal condition and terminal link length. *Journal of the Experimental Analysis of Behavior*, *53*, 219-234.
- Spradlin, J. E. (2002). Punishment a positive process? *Journal of Applied Behavior Analysis*, *35*, 475.
- Staddon, J.E., Motheral, S. (1978). On matching and maximizing in operant choice experiments. *Psychological Review*, *85*, 436-444.
- Steiss, J. E., Schaffer, C., Ahmad, H. A., & Voith, V. L. (2007). Evaluation of plasma cortisol levels and behavior in dogs wearing bark control collars. *Applied Animal Behavior Science*, *106*, 96-106.
- Stokman, C. L., & Glusman, M. (1969). Suppression of hypothalamically produced flight responses by punishment. *Physiology & Behavior*, *4*, 523-525.
- Subotnik, K. L., Nuechterlein, K. H., Ventura, J., Gitlin, M. J., Marder, S., Mintz, J., ... & Singh, I. R. (2011). Risperidone nonadherence and return of positive symptoms in the early course of schizophrenia. *American Journal of Psychiatry*, *168*, 286-292.
- Symons F.J. & Thompson T. (1997) Self-injurious behavior and body site preference. *Journal of Intellectual Disability Research*, *41*, 456-468.
- Tanner, B. A., & Zeiler, M. (1975). Punishment of self-injurious behavior using aromatic ammonia as the aversive stimulus. *Journal of Applied Behavior Analysis*, *8*, 53-57.
- Tate, B. G., & Baroff, G. S. (1966). Aversive control of self-injurious behavior in a

- psychotic boy. *Behavior Research and Therapy*, 4, 281-287.
- Tatham, T. A., Gyorda, A. M., & Barrett, J. E. (1993). Generalization of behavioral history across responses in the reversal of the effects of cocaine and d-amphetamine on the punished behavior of squirrel monkeys. *Behavioral Pharmacology*, 4, 61-68.
- Taylor, L., Oliver, C., & Murphy, G. (2011). The chronicity of self-injurious behavior: A long-term follow-up of a total population study. *Journal of Applied Research in Intellectual Disabilities*, 24, 105-117.
- Taylor, R. & Davison, M. (1983). Sensitivity to reinforcement in concurrent arithmetic and exponential schedules. *Journal of the Experimental Analysis of Behavior*, 39, 191-198.
- Terrace, H. S. (1971). Escape from S-. *Learning and Motivation*, 2, 148-163.
- Terrace, H. S. (1973). Conditioned inhibition in successive discrimination learning. *Transactions of the New York Academy of Sciences*, 35, 39-50.
- Thompson, D. M. (1965). Punishment by SD associated with fixed-ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, 8, 189-194.
- Thorndike, E. L. (1913). *The Psychology of Learning* (Vol. 2). Teachers College, Columbia University.
- Thorndike, E. L. (1932). Reward and punishment in animal learning. *Comparative Psychology Monographs*.
- Tiger, J. H., Wierzba, B. C., Fisher, W. W., & Benitez, B. B. (2017). Developing and demonstrating inhibitory stimulus control over repetitive behavior. *Behavioral Interventions*, 32, 160-174.
- Todorov, J. C. (1971). Concurrent performances: Effect of punishment contingent on the switching response. *Journal of the Experimental Analysis of Behavior*, 16, 51-62.
- Toister, R. P., Condron, C. J., Worley, L., & Arthur, D. (1975). Faradic therapy of chronic vomiting in infancy: A case study. *Journal of Behavior Therapy and Experimental Psychiatry*, 6, 55-59.
- Tolman, C. W., & Mueller, M. R. (1964). Laboratory control of toe-sucking in a young rhesus monkey by two kinds of punishment *Journal of the Experimental Analysis of Behavior*, 7, 323-325.
- Tonneau, F. (2005). Windows. *Behavioral Processes*, 69, 237-247.
- Torres, O. V., Jayanthi, S., Ladenheim, B., McCoy, M. T., Krasnova, I. N., & Cadet, J. L. (2017). Compulsive methamphetamine taking under punishment is associated with greater cue-induced drug seeking in rats. *Behavioral Brain Research*, 326, 265-271.

- Trujano, R. E., & Orduña, V. (2015). Rats are optimal in a choice task in which pigeons are not. *Behavioral Processes, 119*, 22-27.
- Tsiouris, J. A. (2010). Pharmacotherapy for aggressive behaviors in persons with intellectual disabilities: treatment or mistreatment? *Journal of Intellectual Disability Research, 54*, 16.
- Tunney, R. J., & Shanks, D. R. (2002). A re-examination of melioration and rational choice. *Journal of Behavioral Decision Making, 15*, 291-311.
- Tyrer, P., Oliver-Africano, P. C., Ahmed, Z., Bouras, N., Cooray, S., Deb, S., ... & Kramo, K. (2008). Risperidone, haloperidol, and placebo in the treatment of aggressive challenging behavior in patients with intellectual disability: a randomised controlled trial. *The Lancet, 371*, 57-63.
- van den Berg, P., Molleman, L., & Weissing, F. J. (2012). The social costs of punishment. *Behavioral and Brain Sciences, 35*, 42-43.
- Vanderschuren, L. J., Minnaard, A. M., Smeets, J. A., & Lesscher, H. M. (2017). Punishment models of addictive behavior. *Current Opinion in Behavioral Sciences, 13*, 77-84.
- van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis, 21*, 381-384.
- van Oyen, H. G., van der Zwan, S. M., van de Poll, N. E., & Walg, H. (1981). Punishment of food rewarded lever holding in male and female rats. *Physiology & Behavior, 26*, 1037-1040.
- Vaughan, W. (1981). Melioration, matching, and maximization. *Journal of the Experimental Analysis of Behavior, 36*, 141-149.
- Vismara, L. A., & Rogers, S. J. (2010). Behavioral treatments in autism spectrum disorder: what do we know? *Annual Review of Clinical Psychology, 6*, 447-468.
- Volkert, V. M., Lerman, D. C., Call, N. A., & Trosclair-Lasserre, N. (2009). An evaluation of resurgence during treatment with functional communication training. *Journal of Applied Behavior Analysis, 42*, 145-160.
- Vollmer, T. R., & Iwata, B. A. (1993). Implications of a functional analysis technology for the use of restrictive behavioral interventions. *Child & Adolescent Mental Health Care, 3*, 95-113.
- Vollmer, T.R., Hagopian, L.P., & Bailey, J.S. (2011). The association for behavior analysis international position statement on restraint and seclusion, *Behavior*

- Analysis*, 34, 103-110.
- Vyazovska, O. V., Navarro, V. M., & Wasserman, E. A. (2016). Stagemwise multidimensional visual discrimination by pigeons. *Journal of the Experimental Analysis of Behavior*, 106, 58-74.
- Wacker, D. P., Steege, M. W., Northup, J. O. H. N., Sasso, G., Berg, W., Reimers, T., et al., (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Behavior Analysis*, 23, 417-429.
- Wagner, A. R. (1966). Frustration and punishment. *Current research in motivation*. New York: Holt, Rinehart & Winston, 229-239.
- Wang, C. Y., Xiang, Y. T., Cai, Z. J., Weng, Y. Z., Bo, Q. J., Zhao, J. P., ... & Chen, D. F. (2010). Risperidone maintenance treatment in schizophrenia: a randomized, controlled trial. *American Journal of Psychiatry*, 167, 676-685.
- Wearden, J.H. & Burgess, I.S. (1982). Matching since Baum (1979). *Journal of the Experimental Analysis of Behavior*, 38, 339-348.
- Weisman, R. G. (1975). Stimulus control by response-dependent shock in discriminated punishment. *Bulletin of the Psychonomic Society*, 5, 427-428.
- Wells, D. L. (2001). The effectiveness of a citronella spray collar in reducing certain forms of barking in dogs. *Applied Animal Behavior Science*, 73, 299-309.
- Wesp, R. K., Lattal, K. A., & Poling, A. D. (1977). Punishment of autoshaped key-peck responses of pigeons. *Journal of the Experimental Analysis of Behavior*, 27, 407-418.
- Whaley, D. L. & Tough, J. (1970). Treatment of a self-injuring mongoloid with shock induced suppression and avoidance. In R. Ulrich, T. Stachnik and J. Mabry (Eds.), *Control of Human Behavior*, Glenview, 111.: Scott, Foresman.
- Williams, D. E., Kirkpatrick-Sanchez, S., & Iwata, B. A. (1993). A comparison of shock intensity in the treatment of longstanding and severe self-injurious behavior. *Research in Developmental Disabilities*, 14, 207-219.
- Wood, R. (1984). Behavior disorders following severe brain injury: their presentation and psychological management. In Closed Head Injury, Brooks N., ed., Oxford University Press, Oxford, UK, pp. 195-219.
- Woolverton, W. L., Freeman, K. B., Myerson, J., & Green, L. (2012). Suppression of cocaine self-administration in monkeys: effects of delayed punishment. *Psychopharmacology*, 220, 509-517.

- Worsham, R. W., Israel, M. L., von Heyn, R. E., & Connolly, D. A. (1992). Treatment of life threatening vomiting and rumination with contingent electric shock. *Unpublished manuscript*.
- Wright, A. A. (1997). Concept learning and learning strategies. *Psychological Science*, 8, 119-123.
- Wright, L. (1973) Aversive conditioning of self-induced seizures. *Behavior Therapy*, 4, 712-713.
- Yamamoto, M., Kikusui, T., & Ohta, M. (2009). Influence of delayed timing of owners' actions on the behaviors of their dogs, *Canis familiaris*. *Journal of Veterinary Behavior: Clinical Applications and Research*, 4, 11-18.
- Yulevich, L., & Axelrod, S. A. U. L. (1983). Punishment: A concept that is no longer necessary. *Progress in Behavior Modification*, 14, 355-382.
- Zentall, T. R. (2014). Suboptimal choice by pigeons: An analog of human gambling behavior. *Behavioral Processes*, 103, 156-164.
- Zentall, T.R. (2016a). When humans and other animals behave irrationally. *Comparative Cognition & Behavior Reviews*, 11, 25-48.
- Zentall, T.R. (2016b). Resolving the paradox of suboptimal choice. *Journal of Experimental Psychology-Animal Learning & Cognition*, 42, 1-14.
- Zentall, T.R., & Laude, J.R. (2013). Do pigeons gamble? I wouldn't bet against it. *Current Directions in Psychological Science*, 22, 271-277.
- Zentall, T. R., & Stagner, J. (2011). Maladaptive choice behavior by pigeons: an animal analogue and possible mechanism for gambling (suboptimal human decision-making behavior). *Proceedings of the Royal Society of London B: Biological Sciences*, 278, 1203-1208.
- Zentall, T. R., Laude, J. R., Stagner, J. P., & Smith, A. P. (2015). Suboptimal choice by pigeons: Evidence that the value of the conditioned reinforcer rather than its frequency determines choice. *The Psychological Record*, 65, 223-229.
- Zimmerman, J., & Ferster, C. B. (1963). Intermittent punishment of SΔ responding in matching to sample. *Journal of the Experimental Analysis of Behavior*, 6, 349-356.