Punishment and Choice

Rafaela M. Fontes
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Psychology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
PUNISHMENT AND CHOICE

by

Rafaela M. Fontes

A dissertation submitted in partial fulfillment

of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Psychology

APPROVED:

Timothy A. Shahan
Major Professor

Sarfaraz Serang
Committee Member

Gregory J. Madden
Committee Member

Katherine R. Brown
Committee Member

Timothy A. Slocum
Committee Member

D. Richard Cutler, Ph.D.
Interim Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2022
ABSTRACT

Punishment and Choice

by

Rafaela M. Fontes

Utah State University, 2022

Major Professor: Dr. Timothy A. Shahan
Department: Psychology

ABSTRACT

Punishment is defined as a consequence that decreases the frequency of behavior that produces it and is an important behavior regulation mechanism for learning to stop engaging in maladaptive behavior. Punishment has implications for behavioral disorders and treatments and plays an important role in both programmed and natural contingencies. Despite the clear relevance of punishment for behavioral regulation, little is known about how punishment works. Furthermore, punishment research has been in evident decline, leaving important empirical and theoretical gaps in the literature. Therefore, the overall goal of the present set of studies is to fill in some of these gaps. The idea that punishment results in undesirable collateral effects is one of the major arguments against the use of punishment and has likely contributed to a reduction in research on the topic. Therefore, the first study (Chapter 2) re-examines some of these putative side effects and highlights the lack of strong empirical support for the notion that these potential side effects are necessarily ubiquitous, long-lasting, or specific to punishment. Furthermore, there is an evident need for a complete and robust quantitative
model of punishment. Therefore, the following two studies (Chapters 3 and 4) focus on the effects of different punishment parameters on choice. More specifically, Chapter 3 investigates the effects of punishment intensity on response allocation and Chapter 4 investigates the effects of punishment rate on response allocation. Together, these studies aim to improve our understanding of punishment effects and underlying mechanisms and contribute to a more comprehensive theory of punishment that integrates punishment into more general theories of behavior.
Punishment is defined as a consequence that decreases the frequency of behavior that produces it and is an important behavior regulation mechanism for learning to stop engaging in maladaptive behavior. Punishment has implications for behavioral disorders and treatments and plays an important role in both programmed and natural contingencies. Despite the clear relevance of punishment for behavioral regulation, little is known about how punishment works. Furthermore, punishment research has been in evident decline, leaving important empirical and theoretical gaps in the literature. Therefore, the overall goal of the present set of studies is to fill in some of these gaps. The idea that punishment results in undesirable collateral effects is one of the major arguments against the use of punishment and has likely contributed to a reduction in research on the topic. Therefore, the first study (Chapter 2) re-examines some of these putative side effects and highlights the lack of strong empirical support for the notion that these potential side effects are necessarily ubiquitous, long-lasting, or specific to punishment. Furthermore, there is an evident need for a complete and robust quantitative model of punishment. Therefore, the following two studies (Chapters 3 and 4) focus on the effects of different punishment parameters on choice. More specifically, Chapter 3 investigates the effects of punishment intensity on response allocation and Chapter 4 investigates the effects of punishment rate on response allocation. Together, these studies aim to improve our understanding of punishment effects and underlying mechanisms and
contribute to a more comprehensive theory of punishment that integrates punishment into more general theories of behavior.
ACKNOWLEDGMENTS

I would like to thank my advisor, Tim Shahan, for having me back in his lab and helping me navigate a second Ph.D. I only accepted the challenge of going back to grad school because I knew I still had much to learn from you, and I was not wrong. Thank you for pushing me and forcing me out of my comfort zone. Thanks for teaching me an entirely new perspective on how to do science. I walk out of this program certain that I am a much better scientist today than I was three and a half years ago. I am also thankful for my former advisor, João Claudio Todorov, for teaching me the love for basic research and for encouraging me to always walk the extra mile. I am sure he would be happy to see me pursuing two of the topics he dedicated so many years of his career to: choice and punishment. Thanks to my dissertation committee, Greg Madden, Katie Brown, Sarfaraz Serang, and Tim Slocum for their support and guidance in this dissertation. It has been an honor to have you on my committee. I am especially thankful for my old (Kaitlyn Browning, Rusty Nall, Paul Cunningham, and Tony Nist) and new (Gabrielle Sutton and Jack Van Allsburg) lab mates. I have learned so much from all of you. Thanks for all the help and support over the years. I could not have done it without you all. I would like to thank my family, especially my parents Rosa and Márcio, for their love and support. I am grateful for having you all so close despite being so far away. An especial thanks to my husband, Rodrigo, for embracing this journey with me. Your love has helped me overcome it all. You gave me the peace and courage to keep going, and I cannot thank you enough for that. Lastly, thanks to all my dear rats. Their cuteness made the lab life a lot easier. Para todos que me ajudaram a chegar até aqui, muito obrigada!

Rafaela Fontes
# CONTENTS

Page  

ABSTRACT .......................................................................................................................... iii  

PUBLIC ABSTRACT ........................................................................................................... v  

ACKNOWLEDGMENTS ...................................................................................................... vii  

LIST OF TABLES ................................................................................................................ x  

LIST OF FIGURES ............................................................................................................. xi  

CHAPTER  

1. INTRODUCTION ........................................................................................................... 1  

References ........................................................................................................................ 6  

2. PUNISHMENT AND ITS PUTATIVE FALLOUT: A REAPRAISAL ......................... 14  

Abstract .......................................................................................................................... 14  

Introduction ...................................................................................................................... 15  

What is punishment and how does it work? ................................................................. 16  

Putative Shortcomings and Side Effects of Punishment ............................................. 21  

Response Suppression is Temporary ........................................................................... 21  

Conditioned Punishment ............................................................................................... 26  

Increase in Escape and Avoidance Behavior ............................................................... 31  

Punishment-Induced Aggression ................................................................................... 35  

Countercontrol .................................................................................................................. 38  

Conclusions and Future Directions ............................................................................. 41  

Acknowledgments ........................................................................................................... 46  

References ....................................................................................................................... 47  

3. PUNISHMENT INTENSITY AND RESPONSE ALLOCATION ......................... 72  

Abstract .......................................................................................................................... 72  

Introduction ...................................................................................................................... 73  

Methods .............................................................................................................................. 79  

Subjects .............................................................................................................................. 79  

Apparatus .......................................................................................................................... 79  

Procedures ......................................................................................................................... 80  

Data Analysis .................................................................................................................... 81  

Results ................................................................................................................................. 84  

Discussion .......................................................................................................................... 90
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHAPTER 3</strong></td>
<td></td>
</tr>
<tr>
<td>1 Parameter Estimates for Multilevel Models (MLM) for Separate Outcome of Response Rate and Response Ratio</td>
<td>89</td>
</tr>
<tr>
<td><strong>CHAPTER 4</strong></td>
<td></td>
</tr>
<tr>
<td>1 Shock Intensities for Experiment 1</td>
<td>121</td>
</tr>
<tr>
<td>2 Shock Intensities for Experiment 2</td>
<td>135</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>CHAPTER 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Absolute Response Rates ..................................................</td>
</tr>
<tr>
<td>2</td>
<td>Response Ratio ...............................................................</td>
</tr>
<tr>
<td>3</td>
<td>Response Ratio as a Function of Reinforcer Ratio ..................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure</th>
<th>CHAPTER 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Simulations of Eq. 7 When Equal Punishment Rates Are Applied to Both Options ..........................................................</td>
</tr>
<tr>
<td>2</td>
<td>Simulations of Eq. 7 When Equal Punishment Rates Are Applied to Both Options ..........................................................</td>
</tr>
<tr>
<td>3</td>
<td>Exp. 1 – Generalized Matching Law (Eq. 2) Fits .......................</td>
</tr>
<tr>
<td>4</td>
<td>Exp. 1 – GML (Eq. 2) Daily Fits ..........................................</td>
</tr>
<tr>
<td>5</td>
<td>Exp. 1 – Direct Suppression Model (Eq. 7) Fits .......................</td>
</tr>
<tr>
<td>6</td>
<td>Exp. 1 – Response Allocation Across Successive Reinforcers ........</td>
</tr>
<tr>
<td>7</td>
<td>Exp. 1 – Response Allocation Across Successive Reinforcers ........</td>
</tr>
<tr>
<td>8</td>
<td>Exp. 1 – Response Allocation Across Successive Shocks .............</td>
</tr>
<tr>
<td>9</td>
<td>Exp. 1 – Response Allocation Across Successive Shocks .............</td>
</tr>
<tr>
<td>10</td>
<td>Exp. 2 – Generalized Matching Law (Eq. 2) Fits .......................</td>
</tr>
<tr>
<td>11</td>
<td>Exp. 2 – GML (Eq. 2) Daily Fits ..........................................</td>
</tr>
<tr>
<td>12</td>
<td>Exp. 2 – Direct Suppression Model (Eq. 7) Fits .......................</td>
</tr>
<tr>
<td>13</td>
<td>Exp. 2 – Response Allocation Across Successive Reinforcers ........</td>
</tr>
<tr>
<td>14</td>
<td>Exp. 2 – Response Allocation Across Successive Reinforcers ........</td>
</tr>
</tbody>
</table>
Figure
15 Exp. 2 – Response Allocation Across Successive Shocks ..........................143
16 Exp. 2 – Response Allocation Across Successive Shocks ..........................144
Punishment is a critical behavior-regulation mechanism for learning to stop engaging in maladaptive behavior (Horner, 2002; Todorov, 2001, 2011; Vollmer, 2002). Punishment plays an important role in both programmed and natural contingencies and is effective in suppressing behavior of several species in both basic and applied settings (Lerman & Vorndran, 2002). Furthermore, punishment has important implications for behavior disorders. For example, increased resistance to punishment is one of the criteria for addiction (American Psychiatric Association - APA, 2013) because empirical evidence shows that addiction persists despite the aversive outcomes associated with drug use (Amlung & MacKillop, 2014; Deroche-Gamonet et al., 2004; Pelloux et al., 2007; Pelloux et al., 2015; Vanderschuren & Everitt, 2004). It is also known that individuals suffering from anxiety and depression have differential sensitivity to punishment, and aversive stimuli can precipitate such disorders (APA, 2013; Avcu et al., 2019; Fragale et al., 2017; Bar-Haim et al., 2007; Lewinsohn et al., 1973; Spielger et al., 2018).

Besides the implications of punishment for behavioral disorders, punishment also can be an effective tool in behavioral treatments (e.g., Lerman & Vorndran, 2002; Matson & Taras, 1989; Vollmer, 2002). For example, aversive consequences are known to play an important role in abstinence from drugs (Panlilio et al., 2003; Pelloux et al., 2007; Pelloux et al., 2015) and relapse (Amlung & MacKillop, 2014; Fontes et al., 2018; Nall & Shahan, 2020). Punishment also can be critical for treatment of severe problem behavior that requires fast suppression, or of problem behavior that has no identified function (Lerman & Vorndran, 2002; Vollmer et al., 2011). Furthermore, punishment mechanisms seem to underlie popular functional-based interventions (e.g., time out,
contingent reinforcement loss, response blocking), and punishment procedures have been effective in reducing problem behavior when used during behavioral interventions (e.g., Hagopian et al., 1998; Hanley et al., 2005; Matson & Taras, 1989; Risley, 1968; Thompson et al., 1999; Vollmer, 2002).

Although punishment has clear and important implications for behavioral disorders and treatments, little is known about how punishment works and its underlying mechanisms (Critchfield & Rasmussen, 2007; Johnston, 1991). Furthermore, interest in punishment research has been in visible decline over the past few decades (Lydon et al., 2015), leaving important empirical and theoretical gaps in the punishment literature (Critchfield & Rasmussen, 2007). Such gaps have limited our understanding of punishment effects and prevented the consistent incorporation of punishment procedures into behavioral interventions (Johnston, 1972, 1991; Lerman & Vorndran, 2002; Skiba & Deno, 1991). Because the need for more punishment research is evident, the overall goal of the present set of studies is to fill in some of the gaps in the punishment literature.

One of the reasons for the noticeable decline in punishment research may be attributed to the widespread preconceptions about how punishment works and its dangerous collateral effects (Johnston, 1972, 1991; Skiba & Deno, 1991). Thus, the first aim of the present studies is to demystify some of these preconceptions regarding punishment effects. The idea that punishment results in undesirable collateral effects is one of the major arguments against the use of punishment (Sidman, 1989/2000, 1993, 2006) and has influenced how punishment is viewed by behavior analysts. Among some of the putative undesirable effects of punishment are the temporary nature of response suppression produced by punishment, the dangers of conditioned punishment, the
increase in escape and avoidance responses, the occurrence of punishment-induced aggression, and the development of countercontrol. Although such arguments have echoed in the behavior analysis literature for several decades (e.g., Johnston, 1972, 1991; Bland et al., 2018), they seem to lack empirical support. Chapter 2 re-examines each of these putative shortcomings of punishment and side effects based on available empirical data and highlights the lack of strong empirical support for the notion that these potential side effects are necessarily ubiquitous, long-lasting, or specific to punishment.

The points discussed in Chapter 2 indicate the need for additional research and a well-grounded theory of punishment incorporating punishment into a more general theory of behavior. There currently are two matching-law based (Herrnstein, 1961, 1970) theories of punishment, each based on different assumptions about the mechanisms underlying punishment. While the Competitive-Suppression Model suggests that punishment suppresses responding by increasing allocation toward competing options (Deluty, 1976; Skinner, 1953), the Direct-Suppression Model suggests that punishment suppresses responding by decreasing the value of the reinforcer maintaining the punished response (de Villiers, 1980; Farley, 1980, Farley & Fantino, 1978). However, the number of studies on punishment and choice is limited and both models have failed to adequately account for punishment data (Critchfield et al., 2003; Klates et al., 2018; Rasmussen & Newland, 2008). Although empirical evidence suggests that aversive and reinforcing outcomes are not weighted equally by organisms (e.g., Kahneman & Tversky, 1979; Kubanek et al., 2015; Kuroda et al., 2018), the effects of parameters of punishment on response allocation are still unclear. Thus, it remains unknown how organisms make
trade-offs between reinforcers and punishers and how different consequences interact to control response allocation.

Therefore, the second aim of the present studies is to investigate the effects of different punishment parameters on choice. Chapter 3 focuses on the effects of punishment intensity and repeated exposure to punishment on choice. More specifically, Chapter 3 investigates the effects of increases in punishment intensity and previous experience with punishment on response allocation between a punished and an unpunished option. Although the effects of punishment intensity on response suppression have been extensively investigated with single schedules of reinforcement (e.g., Azrin, 1960; Azrin & Holz, 1961; Cohen, 1968; Hake et al., 1967), little is known about how punishment intensity affects choice. Thus, manipulations of punishment intensity in choice situations can provide crucial information for quantitative models of punishment because none of the current punishment models provide a way to account for other parameters of punishment besides rate.

Continuing the study of punishment effects on choice, Chapter 4 investigates the effects of changes in absolute and relative punishment rates on response allocation in a dynamic choice procedure where relative reinforcement rates change rapidly within session. The experiments described in Chapter 4 tested some important predictions of a modified version of the Direct-Suppression Model. Furthermore, the experiments also investigated the impact of punishment rates and ratios on the fitted parameters of the generalized matching law. The results of those experiments provided important information about the interaction between reinforcement and punishment in the control of behavior.
Together, these studies aimed to improve our understanding of punishment effects and underlying mechanisms. The science of behavior can never be complete without understanding how different consequences impacts behavior (Critchfield & Rasmussen, 2007). Thus, understanding how punishment works and establishing a more comprehensive theory of punishment is crucial for integrating it into a more general theory of behavior. Better understanding punishment effects and underlying mechanisms can also have important implications for applied behavior analysis. For example, understanding how punishment works can provide information about variables maintaining problem behavior. Additionally, a better understanding of punishment effects can also inform how to safely and effectively include punishment in behavioral treatments (Critchfield & Rasmussen, 2007; Johnston, 1991; Vollmer, 2002).
References


CHAPTER 2

PUNISHMENT AND ITS PUTATIVE FALLOUT: A REAPPRAISAL

Abstract

In his book *Coercion and Its Fallout* Murray Sidman argued against the use of punishment based on concerns about its shortcomings and side effects. Among his concerns were the temporary nature of response suppression produced by punishment, the dangers of conditioned punishment, increases in escape and avoidance responses, punishment-induced aggression, and the development of countercontrol. This paper revisits Sidman’s arguments about these putative shortcomings and side effects by examining the available data. Although Sidman’s concerns are reasonable and should be considered when using any form of behavioral control, there appears to be a lack of strong empirical support for the notion that these potential problems with punishment are necessarily ubiquitous, long-lasting, or specific to punishment. We describe the need for additional research on punishment in general, and especially on its putative shortcomings and side effects. We also suggest the need for more effective formal theories of punishment that provide a principled account of how, why, and when lasting effects of punishment and its potential side effects might be expected to occur or not. In addition to being necessary for a complete account of behavior, such data and theories might contribute to improved interventions for problems of human concern.

Introduction

Murray Sidman’s exceptional scientific contributions to the field of behavior analysis are widely recognized (e.g., Ahearn, 2011; Arntzen, 2010; Holth & Moore, 2010; Johnson et al., 2020; McIlvane, 2011). Among his many contributions, Sidman’s research has had a noteworthy impact on the understanding of aversive control (e.g., Sidman, 1953a, 1953b, 1966, 1989, 2000). Despite his extensive research in this area, Sidman firmly opposed the use of methods based on aversive control (i.e., coercion), advocating instead for the use of positive reinforcement (Delprato, 1995; Sidman, 1993, 2011). His opposition to the use of coercive methods was especially clear in his book Coercion and its fallout (Sidman, 1989/2000), where he referred to negative reinforcement and punishment as the two major categories of coercive control. According to Sidman (1989/2000), negative reinforcement and punishment work in a complementary manner because a stimulus punishing a response also should increase behavior removing or avoiding that stimulus (i.e., negative reinforcement, e.g., Crosbie, 1998). This interdependence between punishment and negative reinforcement was noted by Sidman as one disadvantage of the use of coercive control, with the other being the dangerous side effects of such practices.

It appears that Sidman’s opposition to the use of aversive control and, more specifically to the use of punishment, may have impacted how punishment is viewed and used by both basic and applied behavior analysts (e.g., Ahearn, 2011; Holth, 2010). There has been an apparent decrease in interest in studying punishment, leaving several empirical and theoretical gaps in the literature (see Critchfield & Rasmussen, 2007; Horner, 2002; Lerman & Vorndran, 2002; Lydon et al., 2015; Todorov, 2001, 2011).
However, a similar decrease has not necessarily been observed with negative reinforcement (e.g., Baron & Galizio, 2005, 2006; Magoon & Critchfield, 2008; Sidman, 2006; Thompson & Iwata, 2005).

Although *Coercion and its fallout* (Sidman, 1989/2000) was focused broadly on the coercive nature of both punishment and negative reinforcement, the present paper focuses on Sidman’s concerns about the use of punishment. Sidman questioned the effectiveness of punishment in controlling behavior based on the transitory nature of the response suppression produced and he alerted his readers to the side effects of its use. Among these side effects were the dangers of conditioned punishment, an increase in escape and avoidance responses during punishment, the occurrence of punishment-induced aggression, and the development of countercontrol strategies.

Sidman’s (1989/2000) concerns are reasonable and highlight important aspects to be considered when using punishment. Despite his concerns and critiques, Sidman did not deny the relevance of punishment research and the need for a better understanding of punishment effects (Holth, 2010). Accordingly, the goal of the present paper is to revisit Sidman’s arguments about the shortcomings and side effects of punishment and examine empirical data that corroborate or contradict these arguments. We hope that such a review improves our understanding of punishment and help to inform discussions about whether, when, and how punishment might be employed, and perhaps help to renew empirical and theoretical interest in punishment.

**What is punishment and how does it work?**

In *Coercion and its fallout* Sidman defines punishment as follows:
We define reinforcers, positive or negative, by their special effect on conduct; they increase the future likelihood of actions they follow. But we define punishment without appealing to any behavioral effect; punishment occurs whenever an action is followed either by a loss of positive or a gain of negative reinforcers. This definition says nothing about the effect of a punisher on the action that produces it. It says neither that punishment is the opposite of reinforcement nor that punishment reduces the future likelihood of punished actions. (Sidman, 1989/2000, p. 45)

This definition was first proposed by Thorndike (1932) and adopted by Skinner (1953). According to this definition, reinforcement and punishment are assumed to be inherently different. Punishment refers to a procedure, while reinforcement is functionally defined, referring to both the procedure and a behavioral process (e.g., Holth, 2010; Sidman, 1993, 2011).

Other underlying assumptions are included in this procedural definition of punishment. First, this definition assumes there is no symmetry between reinforcement and punishment, thus they affect behavior through different mechanisms (e.g., Carvalho Neto et al., 2017; Carvalho Neto & Mayer, 2011; Holth, 2005). Second, defining punishment as a procedure and not a process implies that punishment does not have a direct effect on behavior. Instead, the response suppression observed during punishment is assumed to result from other indirect processes, such as an increase in the frequency of other unpunished responses (i.e., escape and avoidance), or the occurrence of unconditioned emotional responses (e.g., freezing) that are incompatible with the punished response (Hineline, 1984; Schuster & Rachlin, 1968). Thus, punishment is only
effective in reducing behavior to the extent that it increases the frequency of competing unpunished responses (Dinsmoor, 1954; 1955; Hineline, 1984; Solomon, 1964).

A different definition of punishment was proposed by Azrin and Holz (1966), suggesting that punishment is a consequence (e.g., removal of an appetitive stimulus or presentation of an aversive stimulus) that reduces the probability of the behavior that produces it. This definition has been the most commonly used and accepted one (e.g., Hineline & Rosales-Ruiz, 2013; Holth, 2010; Lerman & Vorndran, 2002; Mallpress et al. 2012; see Sidman, 2006 for discussion). Here, punishment is defined functionally, similar to reinforcement, and punishment and reinforcement are considered symmetrical processes having similar effects on behavior, but in opposite directions (Hake & Azrin, 1965). Furthermore, the Azrin and Holz (1966) definition does not attribute the effects of punishment to any observable or hypothesized competing response (Carvalho Neto et al., 2017; Holth, 2005).

The widespread use and acceptance of the Azrin and Holz (1966) definition, however, is not an indication of this definition being superior to the one defended by Sidman (1989/2000, 2006, 1993, 2011). Although the functional definition has the advantage of acknowledging punishment as a behavioral process similar to reinforcement, attaching the definition of punishment to its effects on behavior is not necessarily indicative of any conceptual improvement (Holth, 2005, 2010). Functional definitions have been criticized for their circularity because the function of a stimulus is identified by its effects on behavior while simultaneously being used to at least implicitly explain the occurrence of that behavior (Holth, 2010; Sidman, 2006; Staddon, 1993). Conversely, concerns about the functional definition proposed by Azrin and Holz (1966)
are also not an indication that the procedural definition defended by Sidman (1989/2000) is superior. Defining punishment as a mere procedure that impacts behavior indirectly disregards it as a valid method of behavior control and can mistakenly confound the effects of punishment with the effects of negative reinforcement. However, the decision about the appropriate definition of punishment (and reinforcement) requires a deeper discussion about the conceptual framework upon which behavior analysis is built (e.g., Gallistel, 2005; Gallistel et al., 2001; Killeen, 1988; Shahan, 2017; Staddon, 1993; Timberlake, 1988), a discussion we will not take up here.

One potential way to place both the definition of punishment and the potential mechanisms by which it has its effects on firmer footing is via consideration of formal quantitative models of punishment. To be of any utility, such models must make explicit how punishment has its effects. The two different definitions of punishment described above roughly correspond to two separate quantitative models of punishment based on the matching law (Herrnstein, 1961, 1970). According to the competitive-suppression model (Deluty, 1976), punishers delivered for one option decrease allocation to that option by increasing the relative value of a competing option such that,

\[
\frac{B_1}{B_2} = \frac{(R_1 + P_2)}{(R_2 + P_1)}
\]  

(1)

where \(B_1\) and \(B_2\) are the response rates, \(R_1\) and \(R_2\) are the reinforcement rates, and \(P_1\) and \(P_2\) are the punishment rates for each of two options. By suggesting that the effects of punishment are mediated indirectly by its effects on other responses, such a model is conceptually akin to Sidman’s definition noted above. In contrast, the direct-suppression model (de Villiers, 1980; Farley, 1980) suggests that reinforcement and punishment are symmetrical processes and that punishers for one option decrease allocation to that option
by directly decreasing the relative value of the punished option in a manner that it is opposite in direction (i.e., sign) from reinforcement such that

\[
\frac{B_1}{B_2} = \frac{(R_1-P_1)}{(R_2-P_2)}
\]

where all terms are as in Equation 1.

Direct comparisons of these models have provided overwhelming empirical evidence in favor of the direct-suppression model (de Villiers, 1980; Farley, 1980; Farley & Fantino, 1978). Thus, punishment might best be understood and defined as having a direct suppressive effect on behavior that is the opposite of reinforcement, as suggested by Azrin and Holz (1966). However, both models sometimes fail to provide accurate quantitative predictions about the effects of punishment on behavior (e.g., Critchfield et al., 2003; Rasmussen & Newland, 2008). In addition, Klapes et al. (2018) have shown that modern versions of both models based on the generalized matching law (Baum, 1974) perform no better quantitatively with punishment data than simpler versions of the generalized matching law that completely omit any role for punishment. Thus, despite the superiority of the direct-suppression model in many circumstances, something remains amiss with its quantitative foundations. Resolving whatever it is that remains amiss with the direct-suppression model could have important implications for the definition of punishment and for our understanding of how it has its effects.

In short, there is still no clear or easy answer to the question of how punishment should be defined or how it works. There remains considerable room for debate about both the appropriate definition and the best conceptual/theoretical account of punishment. Our view is that an increase in empirical and theoretical effort directed at generating a more complete quantitative account of punishment is required before a robust and truly
acceptable definition of punishment will emerge. Nevertheless, an inability to provide an acceptable definition or clear understanding of how punishment works does not prevent assessing the putative shortcomings and side effects of the use of punishment with which Sidman was concerned.

**Putative Shortcomings and Side Effects of Punishment**

**Response Suppression is Temporary**

A shortcoming of the use of punishment-based interventions discussed by Sidman (1989/2000) is the fact that response suppression produced by punishment is transitory. According to Sidman:

> After a period of suppression, the activity gradually recovers; the animal ends up pressing the lever as rapidly as ever, even though it still gets shocked every time. [...] The longer the animal stopped, the hungrier it became; the positive reinforcement for pressing the lever eventually became more powerful than the punishment. (Sidman, 1989/2000, p. 72-73).

Indeed, response recovery after continuous exposure to constant punishment is a robust finding and has been demonstrated with different species such as pigeons (e.g., Azrin, 1960a), rats (e.g., Storms et al., 1963), squirrel monkeys (e.g., McMillan, 1967), and humans (e.g., Azrin, 1958); with different punishing stimuli such as shocks (e.g., Rachlin, 1966), bar slap (e.g., Skinner, 1938), and noise (e.g., Azrin, 1958); and with different punishment intensities (e.g., Azrin, 1960a; Azrin & Holz, 1961; Rachlin, 1966). Studies showing response recovery during punishment have reliably demonstrated that response suppression is greater when punishment is first introduced, and response rates typically recover following continuous exposure to a constant punishment intensity (e.g., Azrin,
Furthermore, response rates typically recover to baseline levels or higher with the removal of the punishment contingency (e.g., Azrin, 1960a; Azrin & Holz, 1961).

Conversely, complete response suppression without response recovery during punishment or after the suspension of the punishment also has been reported (e.g., Appel, 1961, 1963; Storms et al., 1962), suggesting that punishment can result in lasting response suppression. For example, response recovery is less likely with high punishment intensities than with low punishment intensities (Azrin, 1958, 1959b, 1960a; Azrin & Holz, 1961). Furthermore, Storms et al. (1962) demonstrated complete and persistent response suppression even when removal of the punishment contingency was followed by increases in food deprivation.

These contradictory results suggest that the degree of response suppression and the transitory effects of punishment can be impacted by other variables. Punishment intensity appears to be the main factor impacting response suppression and recovery during punishment (e.g., Azrin, 1958, 1960a; Azrin & Holz, 1966). Whenever the punishment intensity (e.g., shock voltage) is severe enough to suppress behavior completely, no response recovery is observed (e.g., Appel, 1961, 1963; Storms et al., 1962; Hake et al., 1967). The degree of response suppression and recovery also differs across strains (e.g., Storms et al., 1963), species (e.g., Appel, 1961, 1963; Azrin, 1959a, 1960a, Hake et al., 1967), and punishing stimuli. For example, shocks usually produce greater suppression than loud noise (e.g., Azrin, 1958; Azrin & Holz, 1966). Shocks also produce greater and faster response suppression than timeout from positive reinforcement (i.e., negative punishment; e.g., Holz et al., 1963; McMillan, 1967). Contrary to the
abrupt initial suppression commonly observed with punishment by shock, studies using timeout have shown that the frequency of the punished behavior increases upon the introduction of timeout, and gradually decreases with continuous exposure to the timeout condition (e.g., Brantner & Doherty, 1983; Bostow & Bailey, 1969; Harris, 1985; Holz et al., 1963; McMillan, 1967; Smith, 1981). Furthermore, direct comparisons have shown less response recovery and more lasting response suppression with timeout than with shocks (McMillan, 1967).

Although some of the variables impacting response suppression and response recovery during punishment have been identified, it remains unclear why response recovery happens in the first place. According to Sidman (1989/2000), response recovery occurs due to competition between punishment and reinforcement. Because decreases in response rates commonly are correlated with decreases in obtained reinforcement rates, the animal gets hungrier and the value of the reinforcer overcomes the aversiveness of the punisher. Although this not an unreasonable account of response recovery, it is limited for three reasons. First, the amount of hunger is typically controlled in such experiments with supplemental food after the session. Second, this explanation does not account for the results from Storms et al. (1962) showing no response recovery with increased deprivation; nor does it explain instances of a lack of recovery after the suspension of punishment (e.g., Appel, 1961, 1963; Hake et al., 1967) or the difference in the degree of recovery between different punishing stimuli (e.g., McMillan, 1967). Lastly, if changes in response suppression during punishment are to be explained by the competition between reinforcement and punishment, rules about how organisms make trade-offs between reinforcers and punishers are necessary. Without a quantitative description of how the
values of the reinforcers and punishers change over time, this explanation remains speculative.

A potential alternative explanation for response recovery during punishment is based on habituation. Habituation is defined as a reduction in responsiveness to a stimulus following repeated or prolonged exposure to that stimulus (Rankin et al., 2009; Thompson & Spencer, 1966). In the context of punishment, the reduction in responsiveness would refer to a decrease in the suppressive efficacy of a punishing stimulus. Studies of habituation suggest that higher rates of stimulus presentation and prolonged exposure to a constant stimulus can speed and enhance habituation to that stimulus (McSweeney et al., 1996; Thompson, 2009). Habituation also develops faster and is more pronounced in the presence of a weaker stimulus, and rarely occurs in the presence of a strong stimulus (Rankin et al., 2009; Thompson, 2009). Furthermore, the presentation of a new or stronger stimulus commonly results in recovery of the habituated response (i.e., dishabituation; Rankin, et al., 2009; Thompson, 2009).

The studies reviewed above share several characteristics that highlight the potential role of habituation in response recovery with punishment. First, punishment was presented at high rates (i.e., fixed ratio [FR] 1; Azrin, 1960a, 1960b; Azrin & Holz, 1961; Hake et al., 1967; Rachlin, 1966; Storms et al., 1963), which results in faster habituation. Second, response recovery was observed in the presence of weak punishment (e.g., Azrin, 1959a, 1960a, 1960b, Azrin & Holz, 1961), but not in the presence of intense punishment (e.g., Appel, 1963; Azrin, 1959b; Hake et al., 1967; Storms et al., 1962). Third, dishabituation (i.e., recovery of the habituated response) was observed when increases in punishment intensity following response recovery resulted in resuppression
of the punished response (e.g., Azrin, 1960a; Azrin & Holz, 1961; Rachlin, 1966). Further evidence of habituation to punishment also is provided by studies showing that preexposure to the punisher or gradual increases in punishment intensity increase resistance to punishment (e.g., Banks, 1966a, 1966b, 1976; Baron & Antonitis, 1961; Campbell & Cleveland, 1977; Cohen, 1968), and by studies showing that decreases in punishment efficacy are prevented by using varied rather than constant punishers (e.g., Charlop et al., 1988).

The notion that habituation might impact the degree of response suppression and response recovery during punishment could provide important insights about differences in punishment effects across species and stimuli. Because habituation is an adaptive learning mechanism, the speed of habituation, and the stimuli to which organisms habituate depend on the evolutionary history of the species (Eisenstein et al., 2001). Thus, different species and even different individuals within the same species can show different levels of responsiveness to the same stimulus (e.g., Biedenweg et al., 2011; Blumstein, 2016).

The idea that habituation can impact operant conditioning is not new. McSweeney and colleagues have argued that the reinforcing efficacy of a stimulus is affected by habituation to repeated presentation of that stimulus during the operant session, resulting in changes in response rates across session time (e.g., McSweeney et al., 1996; McSweeney & Murphy, 2009; McSweeney & Roll, 1998; McSweeney & Swindell, 2002). It seems reasonable that a similar process may occur during punishment. If the response recovery observed in punishment studies might result from habituation to the
punisher, the transitory effects of punishment should not be considered grounds for challenging the effectiveness of punishment in general.

In conclusion, the studies reviewed above suggest that the transitory effects of punishment noted by Sidman (1989/2000) are real. However, those transitory effects are likely not specific to punishment and depend on several aspects of the environment and of the contingency. This is true for all variables controlling behavior; thus, it should not be taken as an intrinsic disadvantage of punishment (Johnston, 1972).

**Conditioned Punishment**

One side effect of punishment discussed by Sidman (1989/2000) was conditioned punishment and its role in the generalization of punishment effects to the environment in which punishment is delivered. According to Sidman,

> Whenever we are punished, more and more elements of our environment become negative reinforcers and punishers. We come more and more under coercive control and we rely more and more on countercoercion to keep ourselves afloat. […] That is why conditioned punishment is a “toxic” side effect of punishment. Environments where we are punished become punishing themselves and we react to them as to natural punishers. (Sidman, 1989/2000, p. 89)

Studies of conditioned punishment have investigated the suppressive effects of stimuli associated with a punisher using two procedures: discriminated punishment and conditioned suppression (Church et al., 1970). In discriminated punishment experiments, only responses in the presence of a specific stimulus are followed by the punisher. Thus, the punisher is contingent on both the response and the antecedent stimulus (i.e., discriminative stimulus; Church et al., 1970). In conditioned suppression experiments, a
neutral stimulus is paired with an unconditioned aversive stimulus (i.e., punisher), thus acquiring aversive properties through classical conditioning. The suppressive effect of the conditioned stimulus is demonstrated when response-independent presentation of the conditioned stimulus results in suppression of an operant response that was never previously followed by the unconditioned aversive stimulus.

A discriminative stimulus associated with the delivery of response-dependent punishment can function as a conditioned punisher for other responses. For example, Davidson (1970) trained rats on a multiple schedule of reinforcement and showed that the discriminative stimulus associated with the punished component functioned both as a punisher and as a negative reinforcer when presented dependent on responding during the unpunished component. Using a similar procedure, Weisman (1975) demonstrated that the discriminative stimulus for the punished component functioned as a punisher, but only while it continued to be associated with the delivery of the unconditioned punisher. In a related study, Hake and Azrin (1965) demonstrated that the conditioned stimulus from a conditioned suppression procedure also can function as a conditioned punisher when presented dependent on a response. Pigeons were trained on a conditioned suppression procedure where a tone was paired with shock. When the tone was used as a conditioned punisher delivered contingent on key pecking, suppression of key pecking was then observed. The suppressive efficacy of the tone was a function of the intensity of the shock with which the tone was paired. Furthermore, the tone was only effective as a conditioned punisher while the tone-shock contingency was maintained. Taken together, the results of these studies suggest that discriminative or conditioned stimuli associated with punishment can in fact become punishers themselves. However, these stimuli are
only effective in suppressing operant responses while the contingency between the stimuli and the original punisher is maintained.

Generalization of the effects of conditioned punishers also has been investigated. For example, Honig and Silvka (1964) reinforced key pecking in the presence of seven different key colors and superimposed response-dependent punishment on the schedule of reinforcement for one of the colors. Punishment effects generalized to all colors initially; however, a U-shaped inhibitory generalization gradient developed with continued training. Furthermore, response rates returned to baseline levels on the removal of the punishment contingency (see Honig, 1966 and Carman, 1972 for similar findings). Brush et al. (1952) also trained pigeons to key peck in the presence of a discriminative stimulus and reported similar generalization gradients for pigeons tested after reinforcement only and for pigeons tested after key pecking was suppressed by punishment in the presence of the same discriminative stimulus. These results suggest that the generalization process for punishment is similar to that of positive reinforcement.

The suppressive effects of conditioned punishers have been compared between conditioned suppression and discriminated punishment procedures. For example, Orme-Johnson and Yarczower (1974) trained separate groups of pigeons on each procedure and reported greater response suppression with conditioned suppression than discriminated punishment. Furthermore, conditioned suppression effects generalized to stimuli associated with the unpunished baseline, while discriminated punishment effects did not (see Hunt & Brady, 1955 and Hoffman & Fleshler, 1965 for similar results; but see Hoffman & Fleshler, 1961 and Church et al., 1970 for different results). Additionally, greater resistance to extinction of punishment (Hoffman & Fleshler, 1965; Hunt & Brady,
1955) and greater emotional responses (Hunt & Brady, 1955) also have been demonstrated with conditioned suppression than with discriminated punishment.

These findings suggest that the suppressive effects of a conditioned punisher are directly related to the contingency between the conditioned and unconditioned punishers. Once this contingency is broken, the conditioned punisher loses its punishing efficacy. Furthermore, the contingency between the response and the delivery of the punisher (conditioned or unconditioned) also seems to play an important role in the degree of response suppression and generalization of the suppressive effects. This suggests that the "aversiveness" of the punishing stimulus, as measured by the degree of response suppression and emotional responses produced by the punisher, is impacted by the organism’s control of the punisher.

In conclusion, the studies reviewed above support Sidman’s argument that stimuli correlated with presentation of unconditioned punishers can become punishers themselves. However, those stimuli are only effective as conditioned punishers while correlated with unconditioned punishers, and do not necessarily acquire lasting effects of the unconditioned punishers with which they are associated. Furthermore, the generalizability of conditioned punishment effects is reduced with continued training, contradicting Sidman’s argument that more exposure to punishment results in greater generalization of response suppression. Thus, it appears that these concerns of Sidman are not supported by empirical evidence. Instead, the “toxicity” of the conditioned punishment side effect seems to be greatly impacted by the animal’s control of the punishment delivery and the information conditioned punishers provide about the contingency.
Furthermore, contrary to Sidman’s concerns about conditioned punishment, in applied settings the establishment of conditioned punishers commonly is described as a desirable side effect of punishment (e.g., Brantner & Doherty, 1983; Johnston, 1972; Lerman & Vorndran, 2002). However, few applied studies have addressed these effects. There is some evidence that verbal cues paired with the delivery of an unconditioned punisher can acquire conditioned punishing functions (e.g., Dorsey et al., 1980; Lovaas & Simmons, 1969), though it remains unclear under which conditions those conditioned punishers result in response suppression during treatment. For example, verbal warnings are usually presented before the imposition of response-dependent timeout from positive reinforcement (e.g., Harris, 1985; MacDonough & Forehand, 1973; Wilson & Lyman, 1983). However, the effectiveness of verbal warning or other stimuli associated with the onset of the timeout as a conditioned punisher has yet to be investigated (Brantner & Doherty, 1983; Everett et al., 2010; Harris, 1985).

Applied studies also have provided evidence that response suppression obtained with punishment-based interventions, such as timeout, can generalize to other nontarget undesirable behavior (e.g., Brantner & Doherty, 1983; Firestone, 1976; Lovaas & Simmons, 1969). Again, such effects are generally described as a desirable side effect. However, generalization of the suppressive effects of punishment to other, desirable behavior also has been reported (e.g., Lerman et al., 2003; Mayhew & Harris, 1978). These mixed results have prevented a clear understanding of the conditions under which desirable and undesirable generalization of punishment effects occur in applied settings, thus highlighting the importance of more research on this potential side effect of punishment (e.g., Lydon et al., 2015; Matson & Taras, 1989).
Increase in Escape and Avoidance Behavior

Sidman (1989/2000) considered punishment and negative reinforcement as complementary processes, suggesting that the difference between them relies on the temporal relation between the presentation of the aversive stimulus (i.e., negative reinforcer or punisher) and the occurrence of behavior. Given the intrinsic connection between punishment and negative reinforcement, the second side effect of punishment (and conditioned punishment) discussed by Sidman was an increase in escape and avoidance behavior. According to Sidman,

Punishers, whether things, places, events, or people, suppress actions that produce them but also generate escape as one of their side effects. A victim of punishment who can turn it off, or can somehow get out of the situation, will do so. (Sidman, 1989/2000, p. 93)

The notion that punishment increases escape and avoidance is directly related to the procedural definition of punishment discussed above (Sidman, 1989/2000) suggesting that punishment only has an indirect effect on behavior by increasing the frequency of competing responses (Carvalho Neto et al., 2017; Church, 1963; Holth, 2005). This competing response hypothesis states that behavior suppression observed during punishment is due to 1) unconditioned emotional responses elicited by the punisher that compete with the punished response (e.g., Estes, 1944, Estes & Skinner, 1941), or 2) increases in the frequency of operant responses that are negatively reinforced by the removal of the punisher or conditioned punishers (e.g., Dinsmoor, 1954, 1955, 1977, 2001; Millenson & MacMillan, 1975; Sidman, 1993, 2000).
The contribution of unconditioned emotional responses to response suppression during punishment has been demonstrated by studies on conditioned suppression and by experiments using response-independent punishers. Given the lack of dependency between the response and the punisher in both procedures, the obtained response suppression is attributed to emotional responses elicited by the punisher that compete with the positively reinforced operant behavior (e.g., Annau & Kamin, 1961; Estes & Skinner, 1941; Hunt & Brady, 1955; Orme-Johnson & Yarczower, 1974). This competing emotional response hypothesis has been challenged by studies showing greater response suppression with response-dependent than response-independent punishment (e.g., Azrin, 1956; Camp et al., 1967; Schuster & Rachlin, 1968). If unconditioned emotional responses were responsible for response suppression during punishment, equal suppression should occur in both conditions. The greater suppression obtained with response-dependent punishment thus suggests that punishment has a suppressive effect regardless of the occurrence of emotional responses (Church, 1963; Schuster & Rachlin, 1968).

The hypothesis that response suppression during punishment results from increases in competing operant responses (i.e., avoidance and escape) also has been extensively investigated. For example, Millenson and McMillan (1975) arranged reinforcement dependent on 10 s of bar holding with rats and showed that the average hold time was greater than 10 s during baseline but considerably shorter than 10 s when punishment was superimposed on the schedule of reinforcement. Failures to complete the response requirement (i.e., 10-s hold) during punishment were interpreted as avoidance responses that prevented the delivery of punishment.
Furthermore, Azrin, Hake, et al. (1965) and Arbuckle and Lattal (1987) investigated the effects of the availability of a specific avoidance response on behavior suppression during punishment with pigeons. In Azrin, Hake, et al., an FR1 punishment schedule was superimposed on different schedules of reinforcement for pecking the main key. Each peck on a second key (i.e., avoidance response) started an interval during which responses on the main key were not punished. Increases in punishment intensity increased the frequency of avoidance responses, and avoidance responses were maintained even when responding on the avoidance key decreased obtained reinforcement rates. Furthermore, more resistance to punishment occurred when the avoidance response was unavailable than when it was available. However, because the avoidance response allowed the animals to continue responding on the main key in the absence of punishment, increases in avoidance responding did not decrease the frequency of main-key responses, but only decreased the frequency of main-key responses that were punished. Thus, the relation between punishment and negative reinforcement in that study was not entirely clear.

Arbuckle and Lattal (1987) also superimposed punishment on a schedule of key- peck reinforcement. During some of the punishment conditions, the punisher could be avoided if responses were spaced by a minimum inter-response time (IRT). Response rates and shock rates were lower in all conditions in which the IRT avoidance contingency was in effect than in an initial no-avoidance condition. Furthermore, response rates decreased as the length of the IRT required to avoid shocks increased. In a subsequent re-exposure to the no-avoidance condition, response rates decreased even further than during any of the IRT conditions, and shock rates remained relatively low.
These results suggest that the effects of a punisher might be augmented indirectly by negative reinforcement, however, they do require interpreting the absence of responding (i.e., pausing) as an increase in active avoidance. Obviously, such an interpretation introduces some potential interpretive issues related to differentiating response rate decreases resulting from direct effects of punishment versus indirect effects of increases in pauses between the punished response.

Taken together, these findings suggest that negative reinforcement might play a role in response suppression during punishment, supporting the complementary relation between punishment and negative reinforcement. However, none of the studies reviewed above provided evidence that increases in escape and avoidance responses are necessary for punishment to effectively suppress behavior, as proposed by the competing response hypothesis (see Dunham, 1971; Rachlin & Herrnstein, 1969; Schuster & Rachlin, 1968 for discussion), unless one considers the lack of responding as an avoidance response. The competing response hypothesis also has been challenged by empirical data demonstrating suppression during punishment without increases in specific avoidance responses. For example, Leitenberg (1965a, 1967) compared the effects of punishment in the presence and absence of an escape response with rats and reported greater suppression when punishment was delivered in the absence of an escape response than when an escape response was available.

In application, the occurrence of escape and avoidance responses can be one of the main reasons for the inefficacy of punishment-based interventions (Nelson & Rutherford, 1983; Wilson & Lyman, 1983). For example, timeout from positive reinforcement has been shown ineffective in reducing problem behavior when escaping
from timeout is possible or other sources of reinforcement are available during the timeout (e.g., Solnick et al., 1977). Thus, establishing contingencies to prevent escape, such as blocking or return to timeout are commonly recommended (e.g., Donaldson & Vollmer; 2011; Quetsch et al., 2015; Riley et al., 2017).

Therefore, although the relation between punishment and negative reinforcement discussed by Sidman seems clear, there is not enough empirical evidence to confirm that punishment increases competing responses, nor that such an increase in competing responses is the mechanism underlying response suppression during punishment. Instead, the findings above suggest that the consequence of a response can impact how organisms allocate their behavior across other available options. Superimposing punishment on one of many available responses may impact how an organism weighs the consequences associated with all options and how it allocates its time across options (e.g., Baum, 1973, 2010, 2012; Baum & Rachlin, 1969). Thus, the relation between punishment and competing responses would be better understood by acknowledging that punishment may have both a suppressive effect on the punished response and a facilitative effect on other options (e.g., Carvalho Neto et al., 2017; Spradlin, 2002). Therefore, changes in response allocation during punishment would be more appropriately described as resulting from changes in the relative values of the options. Indeed, because they are based on the matching law, both quantitative models of punishment described above necessarily suggest that punishment impacts the relative values of both punished and non-punished options.

**Punishment-Induced Aggression**

Another side effect of punishment discussed by Sidman (1989/2000) was an
increase in aggressive behavior following the presentation of a punisher. As he stated, Coercive practices can bring counterattack against individuals and against the groups [...] It is easy to see how aggression could become a new way of life for the formerly subservient. The very success of the counteraggression can set into motion a self-perpetuating buildup of an aggressive way of life. (Sidman, 1989/2000, p. 211-212)

Sidman’s concern was not only with punishment-induced aggression, but also with the persistence and perpetuation of such responses. This was considered an especially dangerous side effect because the aggression may be misplaced toward an organism that is not the one imposing the punishment and trigger aggressive reactions in the attacked organism.

In fact, several studies have demonstrated that presentations of response-independent aversive stimuli do result in aggressive responses in the form of attack toward another animal (e.g., Azrin et al., 1963; Myer & Benninger, 1966; Ulrich & Azrin, 1962) or toward inanimate objects (e.g., Azrin, 1970; Azrin et al., 1964). Attack and fight responses have been demonstrated with response-independent presentation of different aversive stimuli, such as shocks (e.g., Azrin et al., 1967; Ulrich et al., 1964), preheated floor (e.g., Ulrich & Azrin, 1962), and tail pinches (e.g., Azrin, Hake & Hutchinson, 1965). Furthermore, this effect has been replicated with several species, such as rats (e.g., Myer & Benninger, 1966; Ulrich & Azrin, 1962), mice (Azrin, 1964; Ulrich, 1966), squirrel monkeys (e.g., Azrin et al., 1963), hamsters (e.g., Ulrich & Azrin, 1962), and cats (e.g., Ulrich et al., 1964).
Studies investigating aggressive responses with the presentation of response-independent aversive stimuli have shown that the probability of such responses depends on both environmental and organismic variables (Azrin, 1964; Ulrich, 1966). Among the environmental variables, the frequency of aggression increases with the frequency (Ulrich & Azrin, 1962), intensity (Ulrich & Azrin, 1962; Ulrich et al., 1964), and duration (Azrin, Ulrich, et al., 1964) of the aversive stimulus. However, this function is reversed at more extreme intensity and duration of shocks, and aggressive responses seem to decrease when shocks are severe enough to produce escape and physical reactions (Azrin, 1964; Azrin et al., 1964; Azrin, Ulrich, et al., 1964; Ulrich, 1966; Ulrich & Azrin, 1962). Aggressive responses also are more common immediately after the presentation of the aversive stimulus (e.g., Azrin et al., 1968; Azrin et al., 1964; Hutchinson et al., 1971); in smaller chambers where the animals were physically close compared to chambers with a larger floor area (e.g., Ulrich & Azrin, 1962), and among food-deprived animals compared to free-fed animals (e.g., Cahoon et al., 1971). Among organismic variables, aggressive responses elicited by response-independent aversive stimuli vary among different strains of the same species. For example, attack responses are observed less frequently with Wistar rats than with other rat strains (Ulrich, 1966; Ulrich & Azrin, 1962). Differences across species have also been reported. For example, no aggressive response is observed with guinea pigs (e.g., Azrin, 1964; Ulrich, 1966; Ulrich & Azrin, 1962). Furthermore, pigeons and monkeys typically attack an inanimate object in the absence of another living being, although rats rarely do so (Ulrich, 1966; Ulrich & Azrin, 1962). Variables such as castration, age, and social conditions in the
home cage also have been shown to impact the frequency of aggressive responses (e.g., Hutchinson et al., 1965; Ulrich, 1966).

In applied studies using response-dependent punishment, punishment-induced aggression has been reported with physical punishment (e.g., Mayhew & Harris, 1978) but not with some other punishers such as timeout from positive reinforcement (e.g., Bostow & Bailey, 1969; Risley, 1968). The occurrence of other emotional responses, such as crying and temper tantrums, have been reported anecdotally with the use of seclusion timeout (e.g., Azrin & Wesolowski, 1974; Sachs, 1973). However, reductions of such emotional responses also have been reported to accompany the reduction of the problem behavior during timeout and other punishment-based interventions (e.g., Matson & Taras, 1989; van Oorsouw et al., 2008).

Although elicitation of aggression by aversive stimuli is a robust and reliable finding (Azrin, 1964; Ulrich, 1966), the evidence just reviewed does not suggest that aggression is a necessary collateral effect of punishment. In the experiments reviewed above, the aversive stimulus was delivered response-independently, thus not meeting the definition of punishment as a procedure (i.e., presentation of an aversive stimulus following a specific response) or as a process (i.e., reduction of a response that produces an aversive stimulus).

**Countercontrol**

The final undesirable punishment side effect discussed by Sidman (1989/2000) was the development of countercontrol. Sidman stated that,

> If punishees are confined or restricted and cannot get away, the coercion will inevitably produce one of its most prominent side effects, countercontrol. If
people cannot escape or avoid, they will find another way to deflect punishments and threats of punishment; they will learn how to control their controllers.

(Sidman, 1989/2000, p. 214)

Countercontrol was extensively discussed by Skinner (1953, 1971, 1974) and is defined as operant behavior in response to social aversive control that results in extinction or punishment of the punishing agent’s behavior. Countercontrol can have different topographies such as overt aggression, passive resistance, or escape from the agent imposing the punishment. Therefore, countercontrol is considered a serious and socially relevant side effect of aversive control (Ornelas, 2018).

Basic research with nonhumans on countercontrol is nonexistent. According to Sidman (1989/2000), the lack of studies on countercontrol in laboratory research is a result of the highly controlled environments where such research is conducted. The isolation of the experimental setting prevents the animals from countercontrolling the experimenter. Thus, countercontrol has been discussed as an exclusively human side effect of social aversive control (e.g., Delprato, 2002; Mace, 1994; Miller, 1991; Sidman, 2000; Skinner, 1953, 1974).

Instances of countercontrol with humans in different social situations have been described in the literature. Carey and Bourbon (2004, 2006) described several examples of countercontrol by students observed in schools in several countries. The authors noted that some students described their behavior, such as cheating on an exam or missing class, as countercontrol against their teachers. Countercontrol has also been discussed during behavior modification as a form of resistance from the client to comply with the treatment (e.g., Mace, 1994; Miller, 1991; Seay et al., 1984). Examples of countercontrol
have also been described in experimental studies with humans. For example, Boren and Colman (1970), using a token economy with psychiatric patients, reported that when patients were fined a few tokens for staying in bed instead of attending a morning meeting, attendance dropped from 70% to 0%. During informal observations, the authors mentioned hearing some of the participants ordering others not to attend the morning meetings as a form of rebellion.

However, in all the situations described above, countercontrol was used as a post hoc explanation for unexpected conduct observed during investigation of other topics. Most work on countercontrol has been conceptual, and the variables that impact the probability, frequency, and topography of countercontrol have not been thoroughly investigated empirically. It is also unknown how countercontrol may affect the behavior of the punishing agent and the probability of punishment in the future (Mace, 1994). To the best of our knowledge, the only experimental study attempting to evoke countercontrol was conducted by Ornelas (2018) using a simulated work environment. During the experiment, aversive verbal statements were used to evoke countercontrol from the participants. However, the results were inconclusive about the relevant variables involved in countercontrol. First, the aversive statements were given at the beginning of the experimental session and no aversive stimulus was dependent on the participant’s behavior. Thus, the procedure did not meet the definition of countercontrol as a strategy to deflect punishment and control the punishing agent. Second, the results did not show any evidence of what the experimenter considered as countercontrol by the participants. In conclusion, nearly nothing is known about this potential side effect of punishment
Conclusion and future directions

Sidman’s opposition to the use of aversive control, and more specifically to the use of punishment, was clear in his writings (e.g., Sidman, 1993, 2000, 2011). Although his concerns are reasonable and highlight important aspects to be considered when using any form of behavior control, the literature reviewed above suggests a lack of strong empirical support for the notion that these shortcomings and side effects are ubiquitous, long-lasting, or specific to punishment. The transitory nature of response suppression produced by punishment does not appear to be an inherent issue with punishment and depends on many aspects of the environment and the contingency. In addition, although stimuli associated with unconditioned punishers can indeed become punishers themselves, such effects are not indiscriminately generalized to other stimuli present and do not necessarily persist once the contingency is suspended. Similarly, increases in escape and avoidance can be observed during punishment, but the occurrence of such responses is not necessary for punishment to suppress responding. Increases in aggressive behavior in the presence of aversive stimulation have also been shown to be a reliable effect; however, it is not necessarily or exclusively a result of punishment procedures. As with conditioned punishment effects, the occurrence of punishment-induced aggression seems to be impacted by the organism’s control of the punishment delivery. Lastly, although anecdotal examples of countercontrol have been described in the literature, countercontrol has not been empirically investigated and it remains unclear when or how such behavioral strategies might develop.

The lack of undesirable side effects associated with the use of punishment has also been noted in the applied literature (e.g., Brantner & Doherty, 1983; Harris, 1985;
Johnston, 1972; van Oorsouw et al., 2008). Indeed, the use of punishment-based interventions typically has been related to increases in positive behavior (e.g., Bostow & Bailey, 1969; Firestone, 1976; van Oorsouw et al., 2008; Risley, 1968). For example, Matson and Taras (1989) reviewed 382 applied studies employing different punishment procedures during interventions with individuals with developmental disabilities and concluded that the results reviewed did not provide evidence supporting the occurrence of undesirable side effects. Instead, the majority (93%) reported positive side effects during punishment interventions, such as increases in social behavior and responsiveness to the environment. Furthermore, the severity of the undesirable side effects, to the extent that they occur, was considered less harmful than the target behavior to be treated by punishment (Matson & Taras, 1989).

Given the considerations above, one wonders if opposition to the use of punishment might reflect a more general cultural tendency to regard its use as inherently bad. Such a view of punishment could be one of the reasons for the apparent decline in punishment research over the years (e.g., Bland et al., 2018; Johnston, 1991). Thus, the first step to renew the interest in punishment as a scientific topic is to acknowledge that aversiveness is not intrinsic to punishment but instead is contextually dependent (Leitenberg, 1965b; Perone, 2003). As noted by Perone (2003), the distinction between positive reinforcement and aversive control can be a matter of perspective, and every situation can be interpreted in terms of positive reinforcement or aversive control. As Sidman (1989/2000) noted, the use of deprivation to increase the efficacy of positive reinforcers might also be considered coercive. Thus, such concerns should not be taken as a reason to avoid seeking a better understanding of punishment (Vollmer, 2002).
Regardless of how one feels about Sidman’s (1993, 2000, 2011) and others’ (e.g., Skinner, 1953, 1974) view of punishment, punishment-based procedures are effective in reducing the behavior of several species, in both basic and applied settings (see Lerman & Vorndran, 2002 for a review). Indeed, punishment is a valuable method in the treatment of problem behavior, and is commonly used in such settings (e.g., Hagopian et al., 1998; Hanley et al., 2005; Lerman & Vorndran, 2002; Lydon et al., 2015; Matson & Taras, 1989; Risley, 1968; Thompson et al., 1999). However, much remains unknown about punishment and its potential side effects. These empirical and theoretical gaps emphasize the need for more research on punishment (e.g., Horner, 2002; Johnston, 1991; Todorov, 2001, 2011). The potential benefits of an increased understanding of punishment and its potential side effects could be manifold.

First, an improved understanding of punishment and its putative side effects could help shine an empirical light on preconceptions about the “dangerousness” of punishment. As noted above, there is a lack of strong empirical support for many of the putative shortcomings and side effects of punishment. In cases where those side effects do occur, many questions remain unanswered. For example, it is unclear under what circumstances punishment generalizes to other stimuli present during its presentation and if punishment effects generalize with unconditioned punishers besides shock. Much also remains unknown about the interactions between punishment and reinforcement. Better understanding such interactions could improve our understanding of decision-making processes more generally by providing information about how organisms make trade-offs between different types of consequences. Understanding such trade-offs could provide important information about potential side effects of punishment. As one example, it is
unknown if the availability of other sources of positive or negative reinforcement impacts the frequency of punishment-induced aggression. Lastly, the complete lack of research on countercontrol makes clear the need for additional research on this potential side effect of punishment before it is considered in arguments against the use of punishment.

Second, additional research on punishment could contribute to the development of a well-grounded quantitative theory of punishment. As discussed above, both the competitive-suppression and direct-suppression models have failed to adequately account for punishment data. Furthermore, to the extent that punishment side effects do occur, a good quantitative theory of punishment should provide a principled account of how, why, and when they occur. As just one example, response recovery is a robust and reliable phenomenon that needs to be accounted for by a quantitative model of punishment. If habituation indeed plays a role in response recovery during punishment, a theory of punishment will need to incorporate a formal account of habituation in order to predict the conditions under which recovery should be expected to occur.

Furthermore, a science of behavior cannot be complete without understanding how aversive consequences contribute to behavior control (e.g., Johnston, 1991; Magoon & Critchfield, 2007; Vollmer, 2002). Punishment is a biological, behavior-regulation mechanism critical for learning to stop engaging in maladaptive behavior (e.g., Todorov, 2011; Vollmer, 2002). Regardless of whether or not one believes that punishment should ever be a part of explicitly arranged contingencies, it will always be a part of natural ones. Thus, it is critical that punishment be effectively integrated into more general formal theories of behavior. But for that to happen, the amount of rigorous data related to punishment and its potential side effects needs to increase substantially. Not only would
such data and theories be valuable in their own right, but they could also meaningfully improve applications to problems of human concern.

Finally, our call for increased empirical and theoretical work on punishment should not be misconstrued as a disregard for concerns about the use of punishment on ethical and humanitarian grounds. Nor should this call for additional research be mistaken as an argument for more widespread use of punishment-based practices. Instead, our goal in highlighting empirical and theoretical gaps in the literature is to emphasize the need for a more complete understanding of punishment and its putative pitfalls before adopting or abandoning its use.
Acknowledgements

We have no known conflict of interest to disclose. The authors thank Kaitlyn Browning and Anthony Nist for their feedback on initial drafts of this paper. The preparation of this paper was supported in part by R21AA025604 (TAS) from the National Institute on Alcohol Abuse and Alcoholism.
References


Hineline, P. N., & Rosales-Ruiz, J. (2013). Behavior in relation to aversive events: Punishment and negative reinforcement. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, & K. A. Lattal (Eds.), *APA Handbooks in...


Wolf, M., Risley, T., Johnston, M., Harris, F., & Allen, E. (1967). Application of operant conditioning procedures to the behavior problems of an autistic child: A follow-

https://doi.org/10.1016/0005-7967(67)90004-6
CHAPTER 3

EFFECTS OF REPEATED EXPOSURE TO ESCALATING VERSUS CONSTANT PUNISHMENT INTENSITY ON RESPONSE ALLOCATION ²

Abstract

The present experiment investigated the effects of 1) repeated exposures to escalating punishment intensities and 2) repeated exposure to punishment after periods of vacation on response allocation between punished and unpunished responding in three groups of rats. The first group (intensity + vacation) experienced repeated exposures to escalating punishment intensities after a period of vacation (i.e., return to baseline) from punishment. The second group (intensity-only) experienced repeated exposures to escalating punishment intensities without vacation from punishment. The third group (vacation-only) experienced repeated exposures to a constant punishment intensity after a period of vacation from punishment. Results showed that superimposition of punishment on one of two concurrently available responses decreased allocation toward the punished response and increased allocation toward the unpunished response. Furthermore, greater changes in allocation were observed with the introduction of a moderate constant intensity than with the introduction of a low intensity that increased across sessions. Re-exposure to punishment had different effects between the groups. Although there was evidence that high shock intensities can enhance the efficacy of lower intensities to shift allocation away from the punished response and toward the unpunished response, there was little evidence of changes in response allocation with reintroduction of punishment after a period of vacation.

² Chapter 3 of this dissertation is under review in the Journal of Experimental Analysis of Behavior
Introduction

Punishment is defined as a consequence that decreases the probability of behavior upon which it is contingent (Azrin & Holz, 1966). The degree of response suppression during punishment is known to be impacted by punishment parameters and by the availability of other sources of reinforcement (e.g., Azrin et al., 1963; Azrin & Holz, 1966; Brethower & Reynolds, 1962). For example, studies manipulating punishment intensity have demonstrated that response suppression is a direct function of the intensity of the punisher (e.g., Azrin, 1960, Azrin & Holz, 1961; Cohen, 1968; Hake et al., 1967).

Furthermore, punishment results in a faster and greater response suppression when an alternative unpunished response is available than when the punished response is the only source of reinforcement (e.g., Herman & Azrin, 1964; Holz et al., 1963; Nall et al., 2019; Nall & Shahan, 2020; Pelloux et al., 2015; Rawson & Leitenberg, 1973).

Although it is known that increases in punishment intensity and the availability of an alternative unpunished response enhance response suppression when these variables are manipulated separately, it is unclear what effects the combination of these two variables may have. Only a few experiments have investigated the effects of increases in punishment intensity when an alternative unpunished response was available (Davison, 1970; Fontes et al., 2018; Rachlin, 1967; Reynolds, 1963). These studies demonstrated that increases in punishment intensity across sessions resulted in greater suppression of the punished response and increased allocation toward the unpunished response.

However, the unusual procedures employed in these studies complicate interpretations of how punishment intensity impacts response allocation between punished and unpunished options.
For example, Rachlin (1967, Exp. 1) and Reynolds (1963) investigated changes in response allocation between punished and unpunished responses in concurrent chains schedules. In both experiments, superimposition of punishment on the schedule of reinforcement for one of the terminal links impacted response allocation in the initial link and increased preference for the stimulus associated with the unpunished terminal link. Furthermore, the degree of shift in allocation in the initial link was a function of the punishment intensity in the terminal link. However, it is important to note that in both experiments the changes in allocation were between responses maintained by stimuli associated with the punished and unpunished terminal links (i.e., conditioned reinforcers). Thus, such responses were never directly punished or reinforced. Additionally, although Reynolds (1963) reported changes in response allocation in the initial links with introduction of punishment for one of the terminal links, responding in the punished terminal link remained unchanged.

In Davison (1970), reinforcers were delivered according to a fixed ratio (FR) 10 schedule that could be completed by either lever pressing ten times or completing ten 5-s periods without responding (i.e., differential reinforcement of other behavior – DRO), or any combination thereof totaling ten. During the punishment phase, the tenth response was followed by food and shock if completed by a lever press response or food only if completed by a DRO response. The results showed that the probability of completing the FR requirement by a DRO response (i.e., unpunished response) increased with increases in punishment intensity across sessions. However, because DRO schedules do not specify a response topography, the unpunished response was any response other than lever pressing. Thus, decreases in punished responding (i.e., lever press) increased unpunished
responding (i.e., DRO), necessarily. Additionally, DRO can be considered a form of negative punishment (i.e., emitting the target response results in reinforcement loss) and is a procedure typically used to reduce the frequency of a target response (e.g., Homer & Peterson, 1980; Mulick et al., 1976; Thompson & Iwata, 2005). Therefore, in Davison (1970) the target response (i.e., lever press) was associated with both positive (i.e., shocks) and negative (i.e., DRO) punishment.

Fontes et al. (2018) also reported increases in unpunished responses with increases in punishment intensity. However, in Fontes et al., changes in responding between punished and unpunished options were investigated during a study of resurgence induced by punishment of an alternative response. Thus, punished and unpunished responses were never concurrently reinforced. In that study, a target response (i.e., lever press) was reinforced according to a variable-interval (VI) 15-s schedule in phase 1 and then extinguished in phase 2 while an alternative response (i.e., nose poke) was reinforced according to a VI 15-s schedule. During phase 3, the target response remained on extinction and a punishment contingency was superimposed on the schedule of reinforcement for the alternative response such that each response produced a 50-ms foot shock with a $p = .50$. Shock intensity was then increased daily in 0.10 mA increments from 0.25 mA to 0.75 mA. Increases in punishment intensity during phase 3 resulted in greater suppression of the alternative response and increased the previously extinguished (but unpunished) target response (i.e., resurgence). After an extinction phase, subjects were then re-exposed to the three-phase sequence two more times. During replications of phase 3, a greater suppression of the alternative response (i.e., punished response) and a greater increase in the target response (i.e., unpunished response) was observed in the
presence of the lower punishment intensities compared to the first exposure to punishment.

However, in Fontes et al. (2018) it was unclear why increases in punishment intensity produced greater suppression of the alternative response and faster resurgence with repeated exposures to punishment. First, because target and alternative responses were never concurrently reinforced, it remains unclear whether similar changes in responding would have been observed in a typical concurrent schedule where both punished and unpunished responses are simultaneously available. Second, given that each alternative response had a 0.50 probability of being punished, changes in response and punishment rates were necessarily confounded. Thus, it is unknown if changes in obtained punishment rates across replications played a role in the degree of response suppression observed during each exposure to punishment. Lastly, as discussed by Fontes et al., it is unclear if the pattern of results was a result of the rats’ experience with escalating intensities or the repeated exposure to punishment after a period in the absence of punishment. Both effects have been noted in the punishment literature.

Experiments manipulating punishment intensity have demonstrated that the degree of suppression obtained with a new intensity is typically biased toward the last experienced intensity. That is, greater response suppression is obtained when a given intensity is preceded by a higher intensity than when preceded by a lower intensity (e.g., Azrin, 1960; Azrin & Holz, 1961; Cohen, 1968; de Souza et al., 1984; Hake et al., 1967). Thus, the suppressive effects of low punishment intensities can be enhanced when those low intensities are reintroduced after organisms are exposed to high intensities (e.g., Durand et al., 2021).
Previous studies also have shown that the efficacy of a punisher can be enhanced when punishment is reintroduced after what has been referred to as a “vacation” from punishment (i.e., period in the absence of punishment); thus, resulting in a greater suppression of the punished response than observed during the first exposure (e.g., Azrin, 1960; Azrin & Holz, 1966; Orme-Johnson, 1967; Rachlin, 1966). For example, Rachlin (1966) and Orme-Johnson (1967) exposed rats to a constant punishment intensity during one component of a multiple schedule and showed that responding was more suppressed when that same intensity was reintroduced after a vacation from punishment compared to the first exposure to that punishment intensity. In Fontes et al. (2018), rats were repeatedly exposed to escalating punishment intensities separated by a “vacation” period (i.e., phases 1 and 2) between exposures to punishment (i.e., phase 3), thus it is unclear which of these variables (or both) was responsible for the increased suppression during the replications of the punishment phase.

Therefore, the first goal of the present experiment was to systematically replicate the procedure used by Fontes et al. (2018) with a wider range of shock intensities (0.25 mA to 1.45 mA) and using more typical concurrent schedules where both punished and unpunished responses were concurrently available and reinforced throughout the entire experiment. Understanding how increases in punishment intensity affect responding on the punished and unpunished options could improve our understanding of the effects of punishment on choice. Although the effects of punishment intensity have been extensively investigated in single schedules (e.g., Azrin, 1959, 1960; Azrin & Holz, 1961; Azrin et al., 1963; Durand et al., 2021; Holz & Azrin, 1963; Cohen, 1968; Powel &
Azrin, 1968), few studies have manipulated punishment intensity in concurrent schedules.

The second goal was to separately investigate the role of 1) previous experience with high punishment intensities on increased response suppression during repeated exposures to lower intensities, and 2) repeated exposure to punishment after a period of vacation. Although both the intensity effect and the vacation effect have been observed in previous studies, they have never been investigated in conjunction. Furthermore, none of these effects have been investigated in concurrent schedules, thus it is unclear how they might impact the allocation of responding to punished and unpunished responses in a choice situation.

Therefore, in the present experiment three groups of rats were trained on a concurrent VI 30-s VI 30-s schedules during baseline, and punishment was superimposed on the schedule of reinforcement for one of the responses during the punishment phase. The first group (intensity + vacation) was a systematic replication of Fontes et al. (2018), and rats experienced repeated exposures to escalating punishment intensities after a period of vacation (i.e., return to baseline) from punishment. The second group (intensity-only) experienced repeated exposures to escalating punishment intensities without a period of vacation from punishment. The third group (vacation-only) experienced repeated exposures to a constant punishment intensity after a period of vacation from punishment. Furthermore, because the probability schedule used by Fontes et al. allowed for large variations in obtained punishment rates with changes in responding, in the present experiment punishment was delivered according to a VI schedule for all groups.
Methods

Subjects

Eighteen male Long Evans rats (Charles River, Portage, MI), approximately 8 months old at the beginning of the experiment were used. Rats were individually housed in a temperature-controlled colony room with a 12:12 light/dark cycle and water was available ad libitum in their home cages throughout the experiment. The rats were maintained at 80% of their free-feeding weight. All 18 rats had previous experience with single and concurrent schedules of reinforcement, but no experience with punishment. Subjects were split in three groups of six rats after the last day of baseline. Groups were matched based on absolute and relative response rates. Two rats (one from the intensity + vacation and one from the vacation-only group) were removed from the experiment because they experienced an equipment malfunction during one of the punishment phases. All procedures were conducted in accordance with the Utah State University Institutional Animal Care and Use Committee.

Apparatus

Six identical Med-Associates operant chambers (St. Albans, VT, USA) housed in sound-attenuating cubicles were used. Each chamber measured 30 cm long, 24 cm wide, and 21 cm high. The front door, ceiling, and rear wall were Plexiglas, and the two side walls were aluminum panels. The interior of each chamber consisted of a response panel with two retractable levers. The levers were arranged 13 cm apart, equidistantly from a centralized aperture where food pellets could be delivered. Above each response lever was a white stimulus light. The house light was located on the wall opposite to the response panel. Med-Associates ENV-414 cards connected to grid floor rods allowed the
administration of scrambled foot shock across the rods. All stimulus and response events were programmed and recorded using Med-Associates software and interfacing.

**Procedure**

Sessions were conducted 7 days per week at approximately the same time every day. Sessions lasted 30 min, not including time for reinforcement. During reinforcement, all lights in the chamber were turned off and the feeder light was illuminated for 3 s. Because the rats had previous experience with the operant chamber, no magazine or lever press training were necessary.

*Baseline:* Sessions began with the illumination of the house light and lever lights, and insertion of both levers. All lights remained on for the duration of the session, except during food delivery. Lever presses were reinforced according to VI 30-s VI 30-s concurrent schedules. All VI values throughout the experiment were selected from 10 intervals derived from Fleshler and Hoffman’s (1962) distribution. A changeover delay (COD) of 3 sec was arranged such that a response to either lever was not reinforced if a response to the opposite lever had been emitted in the previous 3 sec. The first baseline lasted 55 sessions for all groups.

*Punishment:* During this phase, all characteristics of the chamber remained the same as in baseline, except that for all three groups a punishment contingency was superimposed on the schedule of reinforcement for one of the responses (left-right counterbalanced across subjects). Responses to the punished lever produced a 50-ms foot shock according to a VI 30-s schedule. The COD also applied to shock deliveries such that responses to the punished lever did not produce shocks if a response to the unpunished lever had been emitted in the previous 3 s. For the *vacation + intensity* group
and for the *intensity-only* group, shock intensity increased from 0.25 mA to 1.45 mA across sessions in daily 0.10 mA increments. For the *vacation-only* group, shock intensity remained constant at 0.85 mA throughout the entire punishment phase. The intensity for the *vacation-only* groups was selected because it was the median value for the escalating groups. This phase lasted for 13 sessions for all groups.

*Replications:* Following the initial punishment phase, the *vacation + intensity* and the *vacation-only* groups were exposed to two additional baseline-punishment sequence replications. Thus, for these two groups the return to baseline functioned as the vacation from punishment. During each replication of the punishment phase, the *vacation + intensity* group experienced escalating shock intensities, as in the initial punishment phase for that group, and the *vacation-only* group experienced a constant shock intensity (i.e., 0.85 mA), as in the initial punishment phase for that group. During the replications, baseline lasted until responding had recovered and remained stable for at least three consecutive sessions. The length of each baseline replication (40 and 20 sessions for the first and second replications, respectively) was equal for both the *vacation+ intensity* and the *vacation-only* groups. For the *intensity-only* group, after the last session of the initial punishment phase, rats were exposed to the entire intensity sequence (i.e., 0.25 mA to 1.45 mA) two more times, without returning to baseline. During the replications, all procedures remained as previously described. All punishment phases lasted 13 sessions for all groups.

**Data Analysis**

The main dependent variables were response rates (responses/min) for punished and unpunished responses and response ratios across sessions and across different
exposures to punishment for each group. Response ratios were calculated by dividing punished response rates by unpunished response rates in each session for each individual rat.

Changes in response rates between phases and changes in response ratio across repeated exposures to punishment were separately analyzed using multilevel model (MLM) analyses because of the hierarchical structure of the data. The multilevel structure of the present data included two levels: sessions and response type nested in rats. The level 1 independent variables were Sessions, which included 13 sessions for each phase (i.e., baseline and punishment), and Response (i.e., punished and unpunished). The level 2 independent variables were Phase, Replication, and Group. In MLM, the effects of the independent variables can be simultaneously modeled for the group and individual subjects through the inclusion of fixed and random effects. While fixed effects model the effects of multiple predictors and their interactions on an outcome variable at the group level, random effects model individual subject differences accounting for the dependence across repeated within-subject observations (Hox et al., 2018).

There are many advantages to using MLM for data from within-subject designs over more common statistical tests, such analyses of variance (ANOVA; Caron, 2019; DeHart & Kaplan, 2019). First, while ANOVA is based on group averages, MLM accounts for individual differences by assigning different weights to different subjects depending on how close or far they are from the average (Caron, 2019; DeHart & Kaplan, 2019). Second, accounting for individual data also has the advantage of increasing the sample size by including all level 1 observations (e.g., repeated measures of the same subject), while analyses based on averages have the sample size reduced to
the number of level 2 units (e.g., groups; Hox et al., 2018). Third, ANOVA assumes independence between observations, which disregards the correlation between repeated measures from the same subject. This assumption is commonly violated in within-subject designs, and this violation may result in incorrectly identifying significant effects (Caron, 2019; DeHart & Kaplan, 2019). Because MLM is designed to model both group and individual effects, these analyses are more appropriate and powerful for data that lack independence across repeated measures (DeHart & Kaplan, 2019; Hox, 2010; Hox et al., 2018).

All statistical analyses were conducted in R (R Core Team, 2020) and the MLM analyses were conducted using the `lme4` package (Bates et al., 2015). The MLMs for response rates and response ratios were separately built in several steps, in a bottom-up approach (Hox et al., 2018). Initially, an intercept only model including a random intercept of rat was fit via the restricted maximum likelihood (REML) estimation method to assess the intra-class correlation (ICC). Then, the models were expanded by including fixed effects of Session, Phase, Replication, Group, and all possible interactions between these variables, after accounting for individual differences (i.e., the random intercept of rat). These intermediate models were fit using maximum likelihood (ML) estimation and the significance of parameter estimates were based on Satterthwaite's method degrees of freedom. The inclusion or exclusion of any fixed or random effects in the final model was determined using likelihood ratio tests. Follow ups for significant interactions were

---

3 Because the MLM for response ratios was restricted to the punishment phase, Phase was not included as a fixed effect on that model.
conducted on estimated marginal means from the final model and significance of pairwise differences were based on Kenward-Roger degrees of freedom (Lenth, 2018). Statistical significance was determined at $p < .05$. Bonferroni corrections were applied to all pairwise comparisons.

**Results**

Figure 1 shows mean punished and unpunished response rates for each group across the last three sessions of baseline and all sessions of punishment during each exposure to the baseline-punishment sequence. The MLM for response rates indicated that 20% of the variance in response rates can be attributed to individual differences between rats, ICC = .200. The final model included a significant three-way (Phase x Response x Group) fixed effects interaction, a random intercept of rat, and Session as a covariate. The significant interaction suggests that changes in response rates across phases differed between responses and between groups. Table 1 shows coefficients and standard errors for the final model for response rates.

Pairwise comparisons between responses indicated there were no significant differences in punished and unpunished response rates during baseline for the *intensity-only* group ($ps \geq .65$). However, for the *intensity + vacation* and *vacation-only* groups, punished response rates were significantly lower than unpunished response rates during baseline ($ps < .01$). Because groups were matched on response rates and response ratios after the first baseline, the difference between responses during baseline for the two groups were the result of differences in response rates during the returns to baseline after
punishment. During the punishment phase, punished response rates were significantly lower than unpunished response rates for all three groups ($p < .001$).

Despite the differences between responses during baseline for the intensity + vacation and vacation-only groups, pairwise comparisons between phases indicated that
punished response rates were significantly lower ($ps < .001$) and unpunished response rates were significantly higher ($ps < .001$) during punishment relative to baseline, for all three groups. Thus, exposure to either escalating or constant punishment intensities had a significant impact on both punished and unpunished responding compared to baseline, regardless of baseline punished and unpunished response rates.

The differences between responses during baseline for those two groups also did not produce differences in response rates between groups, as pairwise comparisons indicated no significant differences in punished or unpunished response rates during baseline between the three groups ($ps = .99$). There were also no significant differences in unpunished response rates during punishment between groups ($ps \geq .27$). Punished response rates, however, were significantly lower for the vacation-only group than for the intensity + vacation and intensity-only groups ($ps \leq .05$) during punishment. These results indicate that exposure to a constant moderate punishment intensity resulted in a greater suppression of the punished response than exposure to escalating intensities. Conversely, unpunished responding increased to the same degree in the presence of both constant and escalating intensities.

Figure 2 shows mean response ratios across sessions of punishment for all three groups during each exposure to the punishment phase. The MLM for response ratios indicated that 37.5% of the variance in response rates can be attributed to individual differences between rats, $ICC = .375$. The final model included three significant two-way fixed-effects interactions (Session x Replication, Session x Group, and Replication x Group), and a random intercept of rat. The significant interactions suggest that changes

\footnote{Three-way interaction was not significant.}
in response allocation differed across replications and between groups. Table 1 shows coefficients and standard errors for the final model for response ratios.

Analysis of simple slopes yielded negative values that were significantly different than zero for both intensity + vacation and intensity-only groups (ps < .001) during each exposure to punishment. For the vacation-only group the slope was not different than zero during any exposure to punishment (ps ≥ .06). These results indicate that for both groups experiencing escalating intensities, response allocation shifted away from the punished response with increases in punishment intensity across session. Conversely, for the group experiencing a constant intensity (i.e., vacation-only) there were no changes in response allocation across sessions of punishment.

Within-group pairwise comparisons across replications were conducted on the first and last sessions of punishment. For the intensity + vacation group, changes in
response allocation were more extreme (i.e., lower response ratio) during the first session of the third exposure relative to the first session of the first exposure to punishment ($p < .05$), but there were no significant differences in response allocation in the remaining comparisons. During the last session of punishment, changes in allocation were less extreme during the second exposure to punishment relative to the first ($ps \leq .01$), but more extreme during the third exposure relative to the second ($p < .01$). Thus, for the intensity + vacation group, there was no systematic effect of replication on response allocation in the presence of low or high intensities.

For the intensity-only group, changes in allocation were more extreme during the first session of both the second and third exposures to punishment, relative to the first session of the first exposure ($ps < .05$), but less extreme during the first session of the third exposure relative to the second exposure ($p < .01$). During the last session of punishment, there were no significant differences across replications for that group ($ps \geq .38$). Thus, for the intensity-only group, more extreme changes in allocation were observed in the presence of the low, but not of high, intensities during both replications of punishment relative to the first exposure to punishment. However, it is important to note that when the low intensities were reintroduced for the third time the shift in allocation was less extreme compared to the second time. These results suggest that although previous experience with high intensities may enhance the efficacy of lower intensities, such effect on response allocation is likely to decrease with repeated re-exposures to punishment.

For the vacation-only group there were no significant differences in response allocation during the first or last session of punishment across replications ($ps \geq .07$).
Table 1

Parameter Estimates for Multilevel Models (MLM) for separate outcome of response rate and response ratio

<table>
<thead>
<tr>
<th></th>
<th>MLM 1</th>
<th>MLM 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response Rate</td>
<td>Response Ratio</td>
</tr>
<tr>
<td>FIXED EFFECTS</td>
<td>b (SE)</td>
<td>p</td>
</tr>
<tr>
<td>Intercept</td>
<td>43.67 (5.46)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Phase, Ref = Baseline Punishment</td>
<td>-19.14 (1.38)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Response, Ref = Punished Unpunished</td>
<td>4.35 (1.38)</td>
<td>&lt;.01**</td>
</tr>
<tr>
<td>Group, Ref. = Int + Vac</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int-Only</td>
<td>3.45 (7.46)</td>
<td>.65</td>
</tr>
<tr>
<td>Vac-Only</td>
<td>0.98 (7.68)</td>
<td>.90</td>
</tr>
<tr>
<td>Session</td>
<td>-0.33 (0.08)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Phase X Response</td>
<td>23.88 (1.96)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Phase X Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pun &amp; Int-Only</td>
<td>-5.01 (2.26)</td>
<td>&lt;.05*</td>
</tr>
<tr>
<td>Pun &amp; Vac-Only</td>
<td>-22.58 (1.96)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Response X Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unpun &amp; Int-Only</td>
<td>-5.34 (2.59)</td>
<td>&lt;.05*</td>
</tr>
<tr>
<td>Unpun &amp; Vac-Only</td>
<td>-0.72 (1.96)</td>
<td>.71</td>
</tr>
<tr>
<td>Phase X Response X Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punishment &amp; Unpun &amp; Int-Only</td>
<td>10.59 (3.20)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Punishment &amp; Unpun &amp; Vac-Only</td>
<td>36.34 (2.77)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Replication, Ref = 1st</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>-0.15 (0.07)</td>
<td>&lt;.05*</td>
</tr>
<tr>
<td>3rd</td>
<td>-0.17 (0.07)</td>
<td>&lt;.05*</td>
</tr>
<tr>
<td>Session X Replication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>0.02 (0.01)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>3rd</td>
<td>0.01 (0.01)</td>
<td>.33</td>
</tr>
<tr>
<td>Session X Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Int-Only</td>
<td>0.00 (0.01)</td>
<td>.65</td>
</tr>
<tr>
<td>Vac-Only</td>
<td>0.06 (0.01)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Replication X Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd &amp; Int-Only</td>
<td>-0.24 (0.06)</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>3rd &amp; Int-Only</td>
<td>-0.01 (0.06)</td>
<td>.83</td>
</tr>
<tr>
<td>2nd &amp; Vac-Only</td>
<td>-0.02 (0.06)</td>
<td>.80</td>
</tr>
<tr>
<td>3rd &amp; Vac-Only</td>
<td>0.12 (0.06)</td>
<td>.07</td>
</tr>
<tr>
<td>RANDOM EFFECTS</td>
<td>VAR</td>
<td>p</td>
</tr>
<tr>
<td>Between-Rat (Intercepts)</td>
<td>142.70</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Within-Rat (Residuals)</td>
<td>186.64</td>
<td>.07</td>
</tr>
</tbody>
</table>

Note: MLM 1 fit on 2184 observations made on 16 individual rats. MLM 2 fit on 624 observations made on 16 individual rats. SE = standard error; Significance codes: *** p < .001; ** p < .01; * p < .05.
Thus, repeated exposure to a constant intensity did not impact response allocation. However, because punished responding was almost completely suppressed for that group during each exposure to punishment, significant changes in response allocation would likely be difficult to detect.

Lastly, between groups comparisons indicated that changes in response allocation during the first session of punishment were more extreme for the vacation-only group than for the intensity + vacation and intensity-only groups in all replications ($p > .001$). However, by the last session of punishment, there were no differences in response allocation between the groups during any of the replications ($p \geq .16$). These results suggest that the abrupt introduction of a moderate punishment intensity produced a greater shift in allocation away from the punished response than the introduction of a low intensity that increased across sessions. Furthermore, the degree of change in allocation observed at the highest intensity experienced by the groups exposed to escalation (i.e., 1.45 mA) was similar to the degree of change in allocation observed for the group exposed to a lower (0.85 mA) but constant intensity.

**Discussion**

The first goal of the present experiment was to systematically replicate the procedure used by Fontes et al. (2018) with a wider range of shock intensities (0.25 mA to 1.45 mA) and using a more typical concurrent schedule. As in Fontes et al., the intensity + vacation group was repeatedly exposed to both escalating shock intensities and vacations from punishment (i.e., return to baseline). Our results showed that increases in punishment intensity produced a greater suppression of the punished
response and increased responding toward an alternative unpunished response; thus, replicating one of the effects reported by Fontes et al.

However, the faster and greater suppression of the punished response and the faster increase in the unpunished response reported by Fontes et al. was not observed in the present experiment. Overall, there was no systematic effect of replication on response allocation for the \textit{vacation + intensity} group. For example, changes in response allocation were similar in the presence of lower intensities during the first and second exposures. However, in the presence of higher intensities, changes in response allocation were \textit{less} extreme during the second exposure relative to the first. Thus, the first reintroduction of escalating intensities had a smaller impact on allocation than when punishment was first introduced. Conversely, when escalating intensities were introduced for the third time, changes in response allocation in the presence of the lower intensities were \textit{more} extreme than during the first exposure. However, in the presence of higher intensities, changes in response allocation were more extreme during the third exposure compared to the second, but not different between the third and the first exposures. Therefore, there was no clear and consistent change in the suppressive efficacy of low or high intensities across repeated exposures to escalating intensities for that group. In summary, the faster and more extreme changes in response allocation with repeated exposures to escalating punishment observed by Fontes et al. in the resurgence study were not replicated with a typical concurrent schedules where punished and unpunished responses were simultaneously available.

One possible explanation for the differences in results between the studies could be the differences in response topography between experiments. In Fontes et al. (2018),
punished and unpunished responses differed in topography (nose poke and lever press, respectively), while in the present experiment both responses were lever press. The different topography between responses in the previous study might have made the punished response more discriminable, thus leading to a faster switch in allocation away from the punished response, and toward the historically unpunished response, once punishment was reintroduced after the vacation.

Another possible explanation for the differences in results could be differences in obtained punishment rates between studies. Because Fontes et al. (2018) used $p = .50$ schedule for punishment, obtained punishment rates were confounded with response rates. As a result, in Fontes et al., obtained punishment rates during the low intensities (when punished responding was less suppressed) were much higher than the obtained punishment rates at the same intensities in the present experiment. For example, in Fontes et al. the average punishment rate ranged from 13.65 shocks/min in the first session of punishment (i.e., 0.25 mA) to 0.22 shocks/min in the last session of punishment (i.e., 0.75 mA), while in the present experiment the average obtained punishment rate for the intensity + vacation group ranged from 2 shocks/min to 0.9 shocks/min during the first (0.25 mA) and last (1.45 mA) sessions of punishment, respectively. Because punishment rates were much lower in the present experiment, punished responding was suppressed to a lesser degree, even in the presence of intensities higher than the ones used by Fontes et al. (2018). For example, in Fontes et al., 0.75 mA was enough to produce a 90% decrease in punished response rates, relative to baseline, while in the present experiment this same intensity produced less than 50% decrease in punished responding. Indeed, in the present experiment, 90% decrease in punished response rates was not observed even at the
highest intensity (1.45 mA) for the escalating groups. Thus, the low obtained punishment rates seem to have decreased the suppressive efficacy of high shock intensities when the intensities were gradually increased across sessions.

The effects of experience with different punishment rates on punished responding is unknown. If punishment rate and intensity share the same properties, it would be reasonable to expect that previous experience with high punishment rates might impact the degree of suppression obtained with repeated exposures to punishment similar to what have been observed with previous experience with high intensities. If that is true, the experience with high punishment rates in Fontes et al. (2018) may have played a role in the increased suppression observed during the replications of the punishment phase. However, this explanation remains speculative without further empirical support.

The second goal of the present study was to separately investigate the effects of 1) previous experience with high punishment intensities on increased response suppression during repeated exposures to lower intensities, and 2) repeated exposure to punishment after a period of vacation. The groups experiencing escalating punishment intensities showed different patterns of changes in response allocation with repeated exposures to punishment. For the intensity-only group significant changes in allocation occurred in the presence of low, but not high, intensities. Thus, for that group, experience with higher intensities during the first exposure to punishment enhanced the suppressive efficacy of the low intensities when they were reintroduced, as it has been shown in previous experiments (e.g., Azrin, 1960; Durand et al., 2021). However, this increased shift in allocation during re-exposures to punishment was restricted to the initial intensities and did not persist throughout the entire escalation. Conversely, for the intensity + vacation
group, no systematic changes in response allocation were observed with repeated exposures to punishment.

One possible explanation for the differences between the two intensity groups might be the re-exposure to baseline between exposures to punishment for the intensity + vacation group. Although the periods of vacation from punishment were expected to enhance the suppressive efficacy of punishment, it is possible that instead, the temporary removal of punishment impacted the effect of experience with high intensities on the suppressive efficacy of the low intensities. Indeed, the effects of vacation from punishment on allocation between punished and unpunished responses were not clear in the present experiment. Although 0.85 mA did not produce complete response suppression for the groups experiencing the escalation, it was enough to completely suppress responding when abruptly introduced, resulting in a floor effect for the vacation-only group. Because responding was already completely suppressed during the first session of the first exposure to punishment, punished responding could not decrease any further during the replications. Thus, there is no evidence in the present experiment that vacations from punishment enhance the suppressive efficacy of punishment in a choice situation.

Overall, the present experiment partially replicated previous studies showing that experience with high shock intensities can enhance the suppressive effect of lower and previously ineffective intensities (Azrin, 1960; Durand et al., 2021). The results also corroborate previous studies showing that superimposition of punishment on one of two concurrently available responses increases responding on the unpunished response (Davison, 1970; Rachlin, 1967; Reynolds, 963). Furthermore, the present results showed
that the degree of suppression of the punished response and of increase of unpunished responding is a function of punishment intensity and the manner in which punishment is introduced. For example, greater shifts in allocation were observed for the group exposed to a moderate and constant intensity than for the groups that experienced gradual increases in shock intensity. Interestingly, the groups exposed to escalating punishment intensities did not show changes in response allocation as extreme as the group experiencing a constant intensity, even at the highest intensity experience for both intensity groups. These results provide another piece of evidence that responding is more resistant to punishment when the punisher is gradually introduced.

Although changes in response rates were followed by changes in obtained reinforcer rates, the decrease in reinforcer rates does not seem sufficient to explain the changes in responding. First, in most cases for the intensity + vacation and the intensity-only groups, decrease in responding occurred sooner than the changes in obtained reinforcer rates. Second, the changes in response allocation during punishment were more extreme than the changes in obtained reinforcer rates. Figure 3 shows response ratio as a function of reinforcer ratio for each group across sessions of punishment. Both intensity + vacation and intensity-only groups show overmatching. For the vacation-only group, however, more extreme changes in obtained reinforcer rates were observed given that punished responding was almost completely suppressed for that group.

Obtained punishment rates also decreased across sessions of punishment. However, decreases in obtained punishment rates also seem an unlikely explanation for the decreases in punished response. First, changes in responding also occurred sooner than changes in punishment rate. Second, previous studies manipulating punishment rate
have shown that responding is more suppressed with higher than with lower punishment rates (e.g., Azrin et al., 1963; Barker et al., 2010; Critchfield et al., 2003; Farley, 1980). Thus, it is unlikely that the lower punishment rates during higher punishment intensities were responsible for the greater response suppression observed at those intensities.

The present results could have important applied implications. For example, resistance to punishment has been an increasing concerned for the treatment of drug abuse because one of the criteria for addiction is continued drug use despite the aversive consequences associated with this behavior (APA, 2013; Deroche-Gamonet et al., 2004; Durand et al., 2021; Vanderschuren & Everitt, 2004). The present results provide evidence that choice of a punished behavior is more resistant to punishment when the intensity of the punisher is increased gradually than when a moderate intensity is abruptly introduced. Thus, one of the reasons for the increased persistence of drug abuse may be because the aversive consequences associated with such behavior were likely not so harmful or damaging when the individual first started using drugs and worsen with time.

Figure 3

Response ratio as a function of reinforcer ratio

![Graph showing response ratio as a function of reinforcer ratio](image)

**Note.** Response ratio (Punished/Unpunished) as a function of obtained reinforcer ratios (Punished/Unpunished) across sessions of punishment. Symbols represent different exposures to punishment. Note different axes for vacation-only group.
Importantly, in the present experiment responding was resistant to escalating punishment even with the availability of an alternative unpunished response, which could be analogous to the availability of other more adaptive sources of reinforcement in the context of addiction.

The present findings can also provide information on how to better incorporate punishment into behavioral treatments when such procedure is necessary. For example, our results indicate that punishing a target response (e.g., problem behavior) can increase allocation toward a more adaptive alternative and unpunished response (e.g., FCT). However, the degree and speed of the change in response allocation will depend on the intensity of the punisher and how punishment is introduced. Implementing a low and inefficient punisher initially may increase resistance to punishment and require much higher intensities to shift allocation toward an alternative unpunished response than would have been necessary if a moderate intensity was abruptly introduced. Furthermore, the present results indicate that repeated implementation of escalating punishment intensities during the treatment might have different impacts on response allocation if punishment for the problem behavior is temporarily removed and then reintroduced or if there are no breaks between repeated exposure to escalation. However, it is important to note that repeated exposure to escalation without vacation in between exposures might have a greater impact on response allocation the first time, but that effect is likely to diminish with continued re-exposures to escalating intensities. On the contrary, the suppressive efficacy of a constant moderate intensity does not seem to decrease with repeated exposures to punishment after periods of break between exposures.
In conclusion, the present findings provide important contributions to our current understanding of the effects of punishment on choice. With the noticeable decline in the number of studies on punishment over the last decades, there are important empirical and theoretical gaps in the literature that need to be addressed (Fontes & Shahan, 2021). Such gaps may be one of the reasons for many preconceptions about the use of punishment. Furthermore, although punishment has been extensively investigated in single schedules during the 60’s and 70’s (Azrin & Holz, 1966; Lerman & Vorndran, 2002), fewer studies have investigated the effects of punishment on choice (e.g., Deluty, 1976, de Villiers, 1980; Farley, 1980; Farley & Fantino, 1978; Critchfield et al., 2013; Rasmussen & Newland, 2008). A better understanding of the impacts of punishment parameters on response allocation can improve our understanding of how animals make trade-off between appetitive and aversive consequences and put us one step closer to a well-grounded theory of punishment.
Acknowledgements

We have no known conflict of interest to disclose. The authors thank Gabrielle Sutton for helping in conducting this research, and Jeremy Haynes and Sarah Schwartz for the feedback on the data analyses. This work was supported in part by R21AA025604 (TAS) from the National Institute on Alcohol Abuse and Alcoholism.
References


https://doi.org/10.1002/jeab.465

https://doi:10.1901/jeab.1967.10-95

https://doi.org/10.1901/jeab.1964.7-185

https://doi.org/10.1901/jeab.1963.6-399

https://doi.org/10.1901/jeab.1963.6-407

https://doi.org/10.1016/S0005-7894(80)80063-3


https://CRAN.R-project.org/package=emmeans


https://doi.org/10.1901/jaba.1968.1-63


https://doi.org/10.1126/science.139.3557.838

https://doi.org/10.1901/jaba.2005.176-03
CHAPTER 4
PUNISHMENT IN THE CONTEXT OF DYNAMIC CHOICE

Abstract

This study investigated the effects of punishment rate on response allocation when relative reinforcement rates change rapidly with session. Predictions from a modified version of the Direct-Suppression Model were tested in two separate experiments. In both experiments, sessions were composed of seven unsignaled components, each one programming a different reinforcement ratio. In Experiment 1, equal punishment rates were superimposed on the schedule of reinforcement for both responses, and absolute punishment rates increased across blocks of sessions. The results showed that punishment increased preference for the richer schedule of reinforcement. However, the increase in preference was not a function of absolute punishment rates. In Experiment 2, unequal punishment rates were superimposed on the schedules of reinforcement for both responses, and relative punishment rates changed across blocks of sessions. The results showed that response allocation shifted away from the richer punishment schedule creating a bias toward the option associated with less frequent shocks. The results of both experiments challenged the predictions of the Direct-Suppression Model. Furthermore, fits of the generalized matching law to the data indicated that superimposition of equal or unequal punishment schedules on responses maintained by unequal reinforcement schedules differentially impact the values of sensitivity and bias.
Introduction

It is well established that when organisms choose between two sources of reinforcement, behavior allocation is impacted by the rate and amount of the reinforcing stimuli (Baum & Rachlin, 1969). The relation between response allocation and parameters of reinforcement has been quantified and well described by the matching law (Herrnstein, 1970; McDowell, 1988). According to the matching law, response allocation matches the reinforcer allocation across options as described by Eq. 1, where $B_1$ and $B_2$ represents response rates and $R_1$ and $R_2$ represents reinforcement rates for each option respectively (Herrnstein, 1970).

$$\frac{B_1}{B_2} = \frac{R_1}{R_2} \tag{1}$$

However, in experiments with both humans and nonhumans, deviations from perfect matching are commonly observed (e.g., Baum, 1979). Contemporary versions of the matching law, such as the generalized matching law – GML (Baum, 1974), have been successful in accounting for these deviations through the sensitivity and bias parameters, as described in Eq. 2 where all terms remain as described in Eq. 1, $a$ represents sensitivity and log $b$ represents bias.

$$\log \left( \frac{B_1}{B_2} \right) = a \cdot \log \left( \frac{R_1}{R_2} \right) + \log b \tag{2}$$

Sensitivity accounts for changes in response allocation due to changes in relative reinforcement rates. When log response ratios are plotted as a function of log reinforcer ratios, sensitivity is measured through changes in the slope of the function (Baum, 1974, 1979; McDowell, 1989). When response allocation perfectly matches reinforcer allocation, $a = 1$. If response allocation is more extreme than expected based on reinforcer allocation, $a$ will be greater (i.e., overmatching). Conversely, if response
allocation is less extreme than expected based on reinforcer allocation, \( a \) will be less than one and the response ratios will approach indifference (i.e., undermatching). Undermatching is the most commonly reported type of deviation from matching in studies of choice (e.g., Baum, 1979; McDowell, 1989).

Bias represents a constant proportional allocation toward one option unrelated to variations in reinforcement ratio and is measured through changes in the intercept of the function (Baum, 1974, 1979). For example, bias accounts for asymmetry due to differences in response topography or effort required for each response, or due to differences in reinforcer quality between the options. When the options are symmetrical, \( \log b \) is equal to zero and the function crosses the origin. However, when the options are asymmetrical, bias can be used to measure additional preference above and beyond that expected based on relative reinforcement rates.

The GML has been successful in describing the relation between parameters of reinforcement and response allocation, and the deviations from matching commonly observed in choice studies (Baum, 1974, 1979). However, the GML fails to account for choice situations resulting in aversive outcomes (i.e., punishers). For example, it is known that options associated with punishers tend to be less preferred and avoided, and that both the parameters of reinforcement and the severity of punishment impact behavior allocation (Azrin & Holz, 1966). However, the effects of punishment on choice have been much less investigated, both empirically and theoretically, than the effects of reinforcement (Horner, 2002; Johnston, 1972, 1991; Todorov, 2001, 2011), and little is known about how reinforcement and punishment interact in the control of behavior (e.g., Lerman & Vorndan, 2002).
Previous studies have demonstrated that superimposition of a punishment contingency on either one or both responses of a concurrent schedule of reinforcement affects the way organisms distribute their behavior across the options (i.e., Farley, 1980; Farley & Fantino, 1978). When punishment is superimposed on only one of two available options, decreases in the rate of the punished response and increases in the rate of the unpunished response are commonly observed (e.g., Barker et al., 2010; Dunham, 1972, 1978; Rasmussen & Newland, 2008; Rawson & Leitenberg, 1973). However, the degree of the shift in allocation toward the unpunished response is a function of the parameters of reinforcement associated with the unpunished option, and the severity of the punisher applied to the punished option. For example, greater persistence of the punished response is observed when suppression of that response results in substantial loss of reinforcers (Herman & Azrin, 1964; Holz et al., 1963; Katz, 1973; Sears, 1964). Conversely, an increase in allocation toward the unpunished response is observed with increases in punishment rate or intensity (e.g., Crosbie et al., 1997; Davison, 1970; Fontes et al., 2018).

Parameters of punishment also affect behavior allocation when punishment is applied to both responses of a concurrent schedule. Deluty (1976) demonstrated that superimposing unequal punishment rates on responses maintained by equal reinforcement rates resulted in an increase in response allocation toward the leaner punishment rate. However, when equal punishment rates are superimposed on responses maintained by unequal reinforcement rates, increased preference for the richer reinforcement option (i.e., overmatching) is commonly observed (de Villiers, 1977; Farley, 1980). Furthermore, increases in the frequency of punishment result in even more extreme
changes in behavior allocation, thus increasing the degree of overmatching (Critchfield et al., 2003; de Villiers, 1977; Farley, 1980 – Exp. 2).

Although the empirical evidence just reviewed suggests that changes in punishment parameters impact behavior allocation, manipulations of punishment rate have only been investigated under limited reinforcement conditions (Critchfield et al., 2003; Deluty, 1976; Farley, 1980). Therefore, it is still unclear how punishment parameters interact with parameters of reinforcement in choice situations. It is also unclear whether organisms are equally sensitive to reinforcers and punishers.

Attempts to quantitatively summarize the effects of punishment on behavior allocation have resulted in two major punishment models (Deluty, 1976; de Villiers, 1980; Farley, 1980). Although both models are based on modified versions of Eq. 1, they differ in their assumptions regarding the mechanism underlying punishment. According to the Competitive-Suppression Model - CSM (Deluty, 1976), punishment for a given option impacts the relative reinforcement value for that option by increasing the value of a competing option (cf. Skinner, 1953). Thus, the decrease in response rates for the punished option is attributed to increases in the relative values of the competing option, as described by Eq. 3 (Deluty, 1976),

\[
\frac{B_1}{B_2} = \frac{(R_1 + cP_2)}{(R_2 + cP_1)}
\]  

(3).

Alternatively, the Direct-Suppression Model - DSM (de Villiers, 1980; Farley & Fantino, 1978) suggests that punishment for a given option directly decreases the reinforcement value for that option. Thus, according to DSM, response suppression during punishment is attributed to a reduction in the value of the reinforcer, as described by Eq. 4 (de Villiers, 1980),
In both Eq. 3 and 4 all terms remain as described in Eq. 1, and $P_1$ and $P_2$ represents the punishment rates for each option respectively. The parameter $c$ (constrained to be $\geq 0$) is a scaler for the punisher that converts units of punishers (e.g., shocks) into units of reinforcers (e.g., food; Farley, 1980; Farley & Fantino, 1978).

Although these models have been useful in providing qualitative (i.e., directional) predictions, they have failed to provide accurate quantitative predictions (Critchfield et al., 2003). Additionally, given that Eq. 3 and 4 are based on Eq. 1 (Herrnstein, 1970), both fail to account for the common deviations from matching often reported in the literature. Therefore, modified versions of the GML (Eq. 2) incorporating both CSM and DSM have been proposed by Critchfield et al. (2003) resulting in Eq. 5 and 6, where all the terms remain as previously described, $a$ is the sensitivity parameter and $\log b$ is bias,

$$ \log \left( \frac{B_1}{B_2} \right) = a \log \left( \frac{R_1+P_2}{R_2+P_1} \right) + \log b $$

(5)

$$ \log \left( \frac{B_1}{B_2} \right) = a \log \left( \frac{R_1-P_1}{R_2-P_2} \right) + \log b $$

(6).

Direct comparisons of the models in their original (Eq. 3 and 4) and generalized (Eq. 5 and 6) versions have provided near universal empirical evidence in favor of DSM (de Villiers, 1980; Farley, 1980; Farley & Fantino, 1978). For example, Critchfield et al. (2003) directly compared Eq. 5 and 6 in three experiments with humans, using money gain and loss as reinforcers and punishers, respectively. The authors concluded that Eq. 6 was qualitatively and quantitatively superior to Eq. 5 in describing punishment data. Furthermore, DSM is the only model to predict previously-observed overmatching with
superimposition of equal punishment rates to both responses of a concurrent schedule (Farley, 1980; Critchfield et al., 2003).

Klapes et al. (2018) reanalyzed Critchfield’s data and compared Eq. 5 and 6 to the previous versions of each model (e.g., Eq. 3 and 4) and to the original GML (Eq. 2). The authors corroborated Critchfield et al.’s conclusions that Eq. 6 outperformed Eq. 5 in accounting for the data. However, Eq. 6 did not outperform the fits obtained with Eq. 2 without any inclusion of punishment parameters. Klapes et al. (2018) reported that similar fits for both Eq. 2 and 6 were obtained when 1) sensitivity and bias assumed different values in each equation or 2) obtained reinforcement ratio was approximately equal to obtained punishment ratio.

However, it is not reasonable to consider that a model that does not include any parameters of punishment, like GML, could be a good descriptor of the punishment effects. The success of the GML in accounting for punishment data suggests that the effects of punishment on behavior allocation are being accounted for by other factors in the model. For example, punishment may result in changes in the values for sensitivity and bias, compared to the values obtained in the absence of punishment. Changes in the obtained values for sensitivity and bias with superimposition of punishment have been reported by Rasmussen and Newland (2008) and Kuroda et al. (2018). Thus, if GML is accounting for punishment data through changes in the values of the fitted parameters, principled rules for how punishment impacts those parameters are necessary to make GML an adequate punishment model.

The adequacy of Eq. 6 as a model of punishment is also still under debate. First, Eq. 6 has only been tested in experiments with humans that programmed reinforcers and
punishers on the same metric (e.g., monetary gain and loss, Critchfield et al., 2003; Kuroda et al., 2018). Because reinforcers and punishers were in the same metric, Eq. 6 was fitted to the data without a punisher scaler, $c$ (Farley, 1980). Thus, in all the comparisons punishers and reinforcers were assumed to have symmetrical effects on behavior. However, empirical evidence has demonstrated that gains and losses are not weighted equally by organisms (e.g., Kahneman & Tversky, 1979; Kubanek et al., 2015; Kuroda et al., 2018). For example, Rasmussen and Newland (2008) demonstrated that superimposing punishment on one of the responses of a concurrent schedule produced an increase in allocation toward the unpunished option greater than the expected increase based on the net gain (i.e., gains minus losses).

The absence of the scaling parameter in Eq. 6 further results in two major limitations for the model. First, without a scaling parameter for the punisher Eq. 6 cannot account for punishment data obtained with reinforcers and punishers that have different units (e.g., food and shock), as commonly used in laboratory research with nonhumans. Second, Eq. 6 cannot account for situations where punishment rates are higher than reinforcement rates. Therefore, the inclusion of the punisher scaler in Eq. 6 is a logical (although absent in the literature) extension of the model, resulting in Eq. 7:

$$
\log \left( \frac{B_1}{B_2} \right) = a \times \log \left( \frac{R_1 - cP_1}{R_2 - cP_2} \right) + \log b
$$

(7),

where all terms remain as previously described and $c$ is the scaler parameter for the punisher constrained such that $c \geq 0$. 
Because Eq. 7 has never been formally proposed as a model of punishment, tests of its predictions, and empirical evaluations of its adequacy are necessary. According to Eq. 7, superimposition of equal punishment rates on responses maintained by unequal reinforcement rates should have a greater impact on the value of the leaner than on the value of the richer schedule of reinforcement, thus increasing preference for the richer reinforcement option (i.e., overmatching). Furthermore, increases in absolute punishment rate that are equally applied to both options should result in more extreme preference for the richer reinforcement option, as shown in Figure 1. Eq. 7 also predicts that the superimposition of unequal punishment rates on responses maintained by unequal reinforcement rates should increase response allocation toward the richer schedule of reinforcement when the richer schedule is combined with the lower punishment rate. However, no changes in allocation should be observed when the richer schedule of reinforcement is combined with the higher punishment rate, as shown in Figure 2.

Some of the predictions based on Eq. 7 are in accordance with the effects of punishment described in the literature and reviewed above (e.g., Critchfield et al., 2003;
Deluty, 1976; de Villiers, 1977, 1980; Farley, 1980). However, the effects of changes in absolute and relative punishment rates have only been investigated under limited conditions, and punishment rates have only been manipulated in combination with equal or constant relative reinforcement rates (Critchfield et al., 2003; Deluty, 1976; Farley, 1980 – Exp. 2). Thus, it remains unclear if the same effects would be obtained with a different reinforcement distribution between the options.

Hence, testing the predictions of Eq. 7 would be an important step toward a better understanding of the interaction between reinforcement and punishment in the control of response allocation during choice. Testing these predictions can improve our understanding of how animals weigh reinforcement and punishment when choosing between options. Furthermore, such tests could provide important information about how the effects of punishment are accounted for by both Eq. 2 and Eq. 7. None of the predictions just described are supported by Eq. 2, given that this equation does not include any terms representing punishment.
Therefore, the only way Eq. 2 could account for the effects of those manipulations is through changes in the values obtained for sensitivity and bias. In that case, a theory of how the parameters of punishment affect sensitivity and bias will be required. Therefore, empirical evidence about the degree of asymmetry between reinforcers and punishers could provide a means to identify differences in sensitivity to relative reinforcement and punishment rates that could be informative to both models.

Factors affecting sensitivity and bias have been extensively examined in the absence of punishment. (e.g., Aparicio et al., 2016; Baum & Rachlin, 1969). Procedures using within-session dynamic changes of relative reinforcement rates (hereafter referred to as the dynamic-choice procedure) have proven a successful way to investigate how relative reinforcement rates affect behavior allocation (e.g., Aparicio & Baum, 2006, 2009; Davison & Baum, 2000, 2003). In such procedures, each session consists of seven unsignaled components programming different relative reinforcement rates. This procedure has been used to investigate variables affecting response allocation and sensitivity to relative reinforcement rates, such as the duration of exposure to each condition (e.g., Baum & Davison, 2004; Davison & Baum, 2000), amount of reinforcement (e.g., Aparicio & Baum, 2009; Davison & Baum, 2003), and changeover response requirements (e.g., Jimenez & Aparicio, 2014).

Despite the widespread use of this dynamic-choice procedure in the choice literature, this procedure has not been extensively used to investigate the impact of punishment on choice. Such an investigation could improve our understanding of the dynamics of choice in punishment situations. Another advantage of investigating
punishment effects on a dynamic-choice procedure is that a wider range of reinforcement and punishment rates can be investigated in a shorter time.

To the best of our knowledge, Klapes (2020 – Exp. 3) is the only study that has examined the effects of punishment on response allocation in the context of dynamic choice. Klapes’ procedure was a computerized adaptation of Davison & Baum’s (2000) procedure to investigate the effects of punishment on choice with human subjects. However, in that experiment only the effects of constant and equal punishment rates or punishment rates that were proportional to the reinforcement rates were investigated in combination with varying reinforcement schedules. Conversely, varying punishment rates were only investigated in combination with equal and constant reinforcement schedules. Furthermore, in Klapes (2020) all the conditions programmed overall punishment rates that were leaner than the overall reinforcement rates. Lastly, although the author used the results to test the adequacy of different quantitative models of punishment, there were no clear description of the effects of punishment schedules on response allocation across conditions.

Because it remains unclear how parameters of punishment and reinforcement interact in the control of choice, the goal of the present experiments was to investigate how changes in absolute and relative punishment rates impact response allocation across a range of relative reinforcement rates in the context of dynamic-choice. Experiment 1 investigated the effects of absolute punishment rates on behavior allocation by increasing the overall punishment rate equally applied to both responses. Experiment 2 investigated the effects of relative punishment rates on response allocation by superimposing varying punishment ratios on both responses.
Experiment 1

Subjects

Eight naïve male Long Evans rats, approximately 3 months old, were used. Rats were maintained at 80% of their free-feeding weight and were individually housed in a temperature-controlled colony room with a 12:12 light/dark cycle with water available ad libitum in their home cages throughout the experiment. All procedures were conducted following the Utah State University Institutional Animal Care and Use Committee guidelines.

Apparatus

Six identical Med-Associates operant chambers (St. Albans, VT, USA) were used. Each chamber was 30 cm long, 24 cm wide and 21 cm high. The interior of each chamber consisted of a response panel with two retractable levers. The levers were arranged 13 cm apart, equidistantly from a centralized aperture where food pellets could be delivered. Above each response lever was a white stimulus light. The front door, ceiling, and rear wall were Plexiglas, and the entire chamber was housed in a sound and light-attenuating cubicle. Med-Associates ENV-414 cards connected to grid floor rods allowed the administration of foot shock scrambled across the rods. All stimulus and response events were programmed and recorded using Med-Associates software and interfacing.

Procedure

Sessions were conducted 7 days per week at approximately the same time every day. During reinforcement delivery, all lights in the chamber were turned off, the feeder was illuminated, and all timers were paused for 3 s. All variable-time (VT) and variable-
interval (VI) schedules were selected from 10 intervals derived from Fleshler and Hoffman’s (1962) distribution.

*Magazine Training:* Each rat received three sessions of magazine training during which a food pellet was delivered on a VT 60-s schedule. During magazine training, all lights in the chamber were off and the levers retracted. Each session lasted 30 minutes. Lever press training started the day following the last day of magazine training.

*Lever press training:* Sessions began with illumination of the house light, insertion of one of the levers, and illumination of the stimulus light above the lever. All the lights remained on during the session, except during food delivery. During this phase, lever pressing was reinforced on a fixed ratio (FR) 1 schedule. Each rat received two sessions of lever press training, with each lever presented in separate sessions. The order of presentation of the levers (left-right) was counterbalanced across subjects. Each session lasted for 30 minutes or 100 reinforcers, whichever occurred first. Baseline started following the last session of lever press training.

*Baseline:* Sessions began with the illumination of the house light and stimuli above the levers, and the insertion of both levers. All lights remained on during the entire session, except during food delivery. Each session was composed of seven unsignaled components programming different reinforcement ratios (1:8, 1:4, 1:2, 1:1, 2:1, 4:1, 8:1). The overall reinforcement rate for each component was kept at 6 reinforcers/min. A changeover delay (COD) of 5-s was arranged such that a response to either lever was not reinforced if a response to the opposite lever had been emitted in the previous 5-s. Each component lasted until 20 reinforcers were obtained from the two levers combined. The components were separated by a 120-s blackout, during which all the lights were turned
off and both levers were retracted. The order of the components was randomly selected without replacement during each session. Sessions ended after all seven components were completed or 90-min had elapsed, whichever occurred first. Sessions finished by time (i.e., sessions with incomplete components) were excluded from the analysis. Baseline lasted until the values obtained for sensitivity and bias remained stable for 5 consecutive days, which took 44 sessions. Values for sensitivity and bias were based on Eq. 2 fits.

**Punishment:** During the punishment phase, sessions remained the same as during baseline, except that responses on both levers produced 50-ms foot shocks according to a constant 1:1 ratio. During the first session of punishment, shock intensity was set at 0.5 mA to avoid complete response suppression. During the second session and throughout the punishment phase, shock intensity was set at 0.7 mA. If a rat failed to finish all seven components during the first session with 0.7 mA, the shock intensity was reduced back to 0.5 mA for that rat during the following session, and then decreased by 0.1 mA across sessions until it reached an intensity where all components were completed. That intensity was then maintained for that rat for the remaining of the punishment phase. Table 1 shows the shock intensity for each rat during Experiment 1. Sessions with incomplete components (i.e., finished by time) were excluded from the analysis, and the rat remained on the same punishment condition until there were five complete sessions in

<table>
<thead>
<tr>
<th>Rat</th>
<th>Shock Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RD101</td>
<td>0.7 mA</td>
</tr>
<tr>
<td>RD202</td>
<td>0.7 mA</td>
</tr>
<tr>
<td>RD203</td>
<td>0.7 mA</td>
</tr>
<tr>
<td>RD204</td>
<td>0.5 mA</td>
</tr>
<tr>
<td>RD205</td>
<td>0.5 mA</td>
</tr>
<tr>
<td>RD206</td>
<td>0.7 mA</td>
</tr>
<tr>
<td>RD207</td>
<td>0.7 mA</td>
</tr>
<tr>
<td>RD208</td>
<td>0.5 mA</td>
</tr>
</tbody>
</table>

*Note.* Shock intensity for each rat in Experiment 1.

---

5 In the absence of punishment, Eq. 7 reduces to Eq. 2.
that condition. The overall punishment rate started at 2 punishers/min and was increased every 5 sessions in the following order: 4, 6, 8, 10 and 2 punishers/min for all rats.

**Baseline:** Baseline was reestablished after the last day of punishment to evaluate whether the changes in behavior allocation in the presence of punishment persisted after punishment was removed. This phase lasted until all rats had completed at least 30 sessions.

**Data Analyses**

All fits were conducted using the sum of the number of responses, reinforcers, and punishers for each option during the last 5 sessions of each condition across all rats. Eq. 2 was separately fitted to data from 1) each condition (7 data points per condition; fitted parameters allowed to vary across conditions), 2) each session (7 data points per session; fitted parameters allowed to vary across sessions), and 3) all conditions simultaneously (56 data points; one set of fitted parameters shared across conditions). Eq. 7 was only fitted to data from all conditions simultaneously (56 data points; one set of fitted parameters shared across conditions). The functions were fitted using least-squares regression (Microsoft Excel Solver). Comparisons of parameter (i.e., sensitivity and bias) values and proportion of variance accounted for by each model were based on fits of each model to all conditions simultaneously. The Akaike Information Criteria and the Bayesian Information Criteria (BIC) were used to compare the goodness of fit between models (Klapes et al., 2018). For the separate fits of Eq. 2 to each condition, an extra sum-of-squares F test was used to evaluate if a single curve or different curves better described data from different conditions, and to determine significant differences in the obtained values of $a$ and $\log b$ across conditions. Statistical significance was determined
using $\alpha = .05$. Response allocation was also analyzed across successive reinforcer and punisher deliveries in each component across conditions. For this analysis, all responses from the beginning of the component to the first reinforcer (or shock) delivered, from the first reinforcer (or shock) to the second, and so on, were pooled across all presentations of that component for all rats during the last 5 sessions of each condition.

**Results and Discussion**

Figure 3 shows log response ratio ($B_1/B_2$) as a function of obtained log reinforcer ratio ($R_1/R_2$) during the first baseline\(^6\) and all punishment phases for each punishment condition (i.e., different punishment rates). To assess changes in the free parameters of Eq. 2 between baseline and punishment, Eq. 2 was fit to data from each condition separately allowing for changes in the fitted values for sensitivity and bias across conditions. Fits of Eq. 2 to the first baseline yielded low values for sensitivity, which is commonly reported in experiments using this dynamic procedure (e.g., Aparicio & Baum, 2006, 2009; Davison & Baum, 2000, 2003). Despite the low sensitivity during baseline, fits of Eq. 2 to each punishment condition showed increases in sensitivity (i.e., increases in slope) relative to baseline. An extra sum-of-squares F test comparing fits to baseline and to each punishment condition indicated that sensitivity (i.e., slope) was significantly higher than baseline ($ps < .05$) during all but one of the punishment

\(^6\) Note that baseline data and fits are replotted across graphs.
These results indicate that the superimposition of equal punishment rates on both options increased preference for the richer schedule of reinforcement.

To track more closely the changes in sensitivity and bias across conditions, Figure 4 shows the parameter values obtained from fits of Eq. 2 to the daily data. Similar to the fits across conditions (Fig. 3), the daily fits also showed increases in sensitivity during punishment relative to baseline. However, there was no systematic increase in sensitivity across sessions with increases in absolute punishment rate. There were also no systematic changes in bias across the conditions.
Despite the significant differences in slope between baseline and punishment, there were no significant differences in slope across punishment conditions (Fig. 3 and 4). The similar slopes across the different punishment conditions provide another piece of evidence that the degree of preference for the richer schedule of reinforcement was not impacted by overall punishment rate. Additionally, significant changes in bias were only observed between the first exposure to the 2/min condition (i.e., when punishment was first introduced) and the remaining punishment conditions ($p < .05$; except 10/min and 2/min, $p_s > .08$).

The increase in slope during punishment relative to baseline confirms the prediction of Eq. 7 that punishment increases preference for the richer schedule of reinforcement. However, contrary to the predictions (see Fig. 1), changes in response allocation during punishment were not a function of the absolute punishment rates, and preference for the richer schedule of reinforcement did not become more extreme with increases in overall punishment rate. Figure 5 shows fits of Eq. 7 to data from all conditions simultaneously. In the fits, the parameter values were shared (i.e., not allowed to vary) across all conditions. Although Eq. 7 accounted for 97% of the variance in the
data, there were no differences between the fitted lines for baseline and punishment phases (note that lines overlap in the graphs). Thus, the model did not capture differences in responding between baseline and punishment.

The similarity between baseline and punishment in Eq. 7 fits was likely due to the small value assumed by $c$ ($c = 0.05$), which decreased the impact of punishment rates on the model. Indeed, when Eq. 2 was fitted to data from all conditions simultaneously (i.e., same parameter values across conditions), values for sensitivity and bias were similar to
the ones obtained by Eq. 7 fits\(^7\). Additionally, both models accounted for the same amount of variance in the data, as indexed by the similar \(R^2\).

The models were compared using information criteria. Eq. 2 had the smallest AIC and BIC value and was used as the best model in the comparisons. However, the comparisons yielded \(\Delta IC < 10\) and Eq. 2 did not outperform Eq. 7 (\(\Delta AIC = 1.21\), \(\Delta BIC = 3.0\)). Thus, the modified version of the direct suppression model depicted in Eq. 7 with the inclusion of the scaler for punishment (i.e., \(c\)) better accounted for the effects of equal punishment rates on response allocation than the generalized matching law (Eq. 2) that does not include any punishment parameters. These results contradict the conclusions from previous studies comparing Eq. 2 to other versions of DSM that did not include the scaler parameter (e.g., Klapes et al., 2018).

Changes in response allocation were also analyzed across successive reinforcer deliveries to assess the impact of punishment on response allocation within each component. Figure 6 shows log response ratio across successive reinforcers for each component during baseline and punishment conditions where the overall punishment rate was lower than the overall reinforcement rate (i.e., < 6 shocks/min). Figure 7 shows log response ratio across successive reinforcers during baseline\(^8\) and punishment conditions where the overall punishment rate was higher than the overall reinforcement rate (i.e., \(\geq 6\) shocks/min). Experiments using dynamic-choice procedures have shown that, in the

\(^7\) Eq. 2: \(a = 0.54\), \(\log b = -0.10\), \(R^2 = .97\).

Eq. 7: \(a = 0.52\), \(\log b = -0.10\), \(c = 0.05\), \(R^2 = .97\).  

\(^8\) Note that baseline data are replotted across figures.
absence of punishment, response allocation increase toward the richer schedule of reinforcement with increasing number of reinforcers. In the present experiment, this pattern of responding was also observed during baseline and preference for the richer schedule of reinforcement increased with consecutive reinforcer deliveries in each component. Furthermore, the degree of preference was a function of the programmed

**Figure 6**

*Exp. 1- Response allocation across successive reinforcers*

![Graph showing response allocation across successive reinforcers](image)

*Note.* Log response ratios following successive reinforcer deliveries in each of the seven components, across conditions. Panels show data from each component that programmed different reinforcer ratio indicated at the top of each panel. Thick solid line shows response allocation during baseline. Symbols show data from punishment conditions where overall punishment rate was lower than overall reinforcement rate (i.e., < 6 shocks/min). Leaner overall punishment rate (i.e., 2/min) highlighted. The value of 0 on the x axis indicates response ratios prior to the first reinforcer delivery.
reinforcer ratio, and more extreme changes in allocation were observed in the
components programing more extreme reinforcer ratios (e.g., 8:1, 1:8) than in the
components programing less extreme ratios (e.g., 2:1, 1:2). With the introduction of
punishment, allocation shifted more quickly and preference for the richer option became

![Figure 7](image)

*Figure 7*

*Exp. 1- Response allocation across successive reinforcers*

*Note.* Log response ratios following successive reinforcer deliveries in each of the seven
components, across conditions. Panels show data from each component that programmed
different reinforcer ratio indicated at the top of each panel. Thick solid line shows response
allocation during baseline. Symbols show data from punishment conditions where overall
punishment rate was higher than overall reinforcement rate (i.e., > 6 shocks/min). Richer
overall punishment rate (i.e., 10/min) highlighted. The value of 0 on the x axis indicates
response ratios prior to the first reinforcer delivery.
more extreme with each reinforcer than during baseline. However, changes in response allocation across reinforcer deliveries were not a function of the overall punishment rate, and similar changes in allocation were observed with low (Fig. 6) and high (Fig. 7) punishment rates.

A similar analysis was conducted to evaluate the impact of successive shock deliveries on response allocation. Figure 8 shows response allocation across successive shock deliveries in each of the seven components during punishment conditions where the overall punishment rate was lower than the overall reinforcement rate (i.e., < 6 shocks/min). Figure 9 shows response allocation across successive shock deliveries in each of the seven components during punishment conditions where the overall punishment rate was higher than the overall reinforcement rate (i.e., ≥ 6 shocks/min). Because there were no criteria for the number of shocks received in each component, the number of shocks delivered varied across components, sessions, and rats. Thus, the shock-by-shock analysis only included the maximum number of consecutive shocks delivered to all rats in each component during each condition and later shocks experienced by some of the rats in each component were excluded from the analyses. The number of shocks delivery in each component was directly impacted by the programmed punishment schedule. Thus, more shocks were delivered in the conditions programming richer punishment schedules (Fig. 9) than leaner punishment schedules (Fig. 8). During the conditions programming leaner punishment schedules (Fig. 8) there was no clear increase in preference for the richer schedule of reinforcement and allocation remained stable across consecutive shock deliveries. Conversely, during the conditions programming richer punishment rates (Fig. 9) there was a gradual increase in preference
for the richer schedule of reinforcenement with consecutive shocks in the components programing more extreme reinforcement ratios. However, preference for the richer schedule of reinforcement was not as extreme across successive shocks as across successive reinforcers and allocation changed directions more frequently with increasing shock deliveries than with increasing reinforcer deliveries (Fig. 6 and 7).

Figure 8

Exp. 1 - Response allocation across successive shocks

Note: Log response ratios following each successive shock delivered in each of the seven components across conditions. Panels show data from each component that programmed different reinforcer ratios indicated at the top of each panel. Lines show data from punishment conditions where overall punishment rate was lower than overall reinforcement rate (i.e., < 6 shocks/min). Leaner overall punishment rate (i.e., 2/min) highlighted. The value of 0 on the x axis indicates response ratios prior to the first shock delivery.
Once punishment was removed (i.e., return to baseline) preference for the richer schedule of reinforcement decreased, as indexed by the decrease in sensitivity (see Fig. 4). However, during the second baseline, response allocation did not fully recover to its pre-punishment levels and sensitivity remained higher than during the first baseline. An extra sum-of-squares F test indicated significant differences in slope between the first and

*Figure 9*

*Exp. 1 - Response allocation across successive shocks*

*Note:* Log response ratios following each successive shock delivered in each of the seven components across conditions. Panels show data from each component that programmed different reinforcer ratios indicated at the top of each panel. Lines show data from punishment conditions where overall punishment rate was higher than overall reinforcement rate (i.e., > 6 shocks/min). Richer overall punishment rate (i.e., 10/min) highlighted. The value of 0 on the x axis indicates response ratios prior to the first shock delivery.
second baseline ($p < .05$). These results suggest that the increase in sensitivity to reinforcer allocation and preference for the richer schedule of reinforcement observed during punishment may not be reversible.

In summary, the present experiment showed increases in allocation toward the richer schedule of reinforcement with superimposition of equal punishment rates on unequal schedules of reinforcement in every level of analysis conducted, i.e., across conditions (Fig. 3), sessions (Fig. 4), and successive reinforcers (Fig. 6 and 7). However, contrary to Eq. 7 predictions, the degree of preference for the richer schedule of reinforcement was not a function of punishment rate and preference did not become more extreme with increases in absolute punishment rates. Although the predictions of Eq. 7 regarding changes in punishment rates were not confirmed, the model accounted well for the variance in the data and outperformed Eq. 2. Additionally, the present results indicate that the way Eq. 2 may account for the effects of equal punishment rates on choice without including any punishment parameter is through changes in the values of its free parameters, more specifically in sensitivity (i.e., slope), in the presence and absence of punishment.

**Experiment 2**

Eq. 7 predicts that superimposing unequal punishment rates on responses maintained by unequal reinforcement rates should increase response allocation toward the richer schedule of reinforcement when the richer schedule is combined with the lower punishment rate. However, response allocation should remain unchanged when the richer schedule of reinforcement is combined with the higher punishment rate (see Figure 2). Experiment 2 tested these predictions and investigated the effects of changes in relative
punishment rates on response allocation across a range of relative reinforcement rates. As in Experiment 1, reinforcer ratios varied within session and punishment ratios varied across sessions.

Subjects

Eight naïve male Long Evans rats, approximately 3 months old, were used. Rats were maintained at 80% of their free-feeding weight and were individually housed in a temperature-controlled colony room with a 12:12 light/dark cycle with water available ad libitum in their home cages throughout the experiment. All procedures were conducted following the Utah State University Institutional Animal Care and Use Committee guidelines.

Apparatus

Same as described for Experiment 1.

Procedure

All procedures were as described in Experiment 1, except during the punishment phase. The first baseline lasted for 41 days.

Punishment: During the punishment phase, sessions remained the same as baseline, except that responses to both levers were punished with 50-ms foot shocks according to unequal punishment rates. Overall punishment rate was kept at 6 shocks/min and punishment was delivered according to seven different ratios (1:8, 1:4, 1:2, 1:1, 2:1, 4:1, 8:1). Each ratio was equally applied to all components during the session and lasted for 5 sessions. Thus, reinforcement ratios continued to vary within session and punishment ratios varied across sessions. The order of exposure to the punishment ratios
was randomly selected without replacement for each subject. This phase lasted until all subjects had experienced all punishment ratios. As in Experiment 1, shock intensity was set at 0.5 mA during the first session of punishment to avoid complete response suppression. During the second session and throughout the punishment phase, shock intensity was set at 0.7 mA. If a rat failed to finish all seven components during the first session with 0.7 mA, the shock intensity was reduced back to 0.5 mA for that rat during the following session, and then decreased by 0.1 mA across sessions until it reached an intensity where all components were completed. That intensity was then maintained for that rat for the remaining of the punishment phase. Table 2 shows the shock intensity and the order of punishment ratios for each rat during Experiment 2. Sessions with incomplete components (i.e., finished by time) were excluded from the analysis, and the rat remained on the same punishment condition until there were five complete sessions in that condition.

**Baseline:** Baseline was reestablished after the last day of punishment to evaluate whether the changes in response allocation in the presence of punishment persisted after punishment was removed. This phase lasted until all rats had completed at least 30 sessions.

<table>
<thead>
<tr>
<th>Rat</th>
<th>Shock Intensity</th>
<th>Order of Punishment Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>RD201</td>
<td>0.7 mA</td>
<td>2-5-4-6-7-3-1</td>
</tr>
<tr>
<td>RD102</td>
<td>0.5 mA</td>
<td>5-2-6-3-4-7-1</td>
</tr>
<tr>
<td>RD103</td>
<td>0.4 mA</td>
<td>2-5-6-4-1-3-7</td>
</tr>
<tr>
<td>RD104</td>
<td>0.4 mA</td>
<td>1-7-4-2-6-3-5</td>
</tr>
<tr>
<td>RD105</td>
<td>0.5 mA</td>
<td>5-6-3-4-2-7-1</td>
</tr>
<tr>
<td>RD106</td>
<td>0.7 mA</td>
<td>1-2-6-7-5-3-4</td>
</tr>
<tr>
<td>RD107</td>
<td>0.4 mA</td>
<td>7-1-5-2-4-6-3</td>
</tr>
<tr>
<td>RD108</td>
<td>0.7 mA</td>
<td>3-4-5-2-7-1-6</td>
</tr>
</tbody>
</table>

*Note. Shock intensity and order of punishment ratios for each rat in Experiment 2. Ratio code: 1 = 8:1; 2 = 4:1; 3 = 2:1; 4 = 1:1; 5 = 1:2; 6 = 1:4; 7 = 1:8.*
Data Analyses

Data analyses were conducted as described in Experiment 1.

Results and Discussion

Figure 10 shows log response ratio (B1/B2) as a function of obtained log reinforcer ratio (R1/R2) during the first baseline\(^9\) and all punishment phases for each punishment condition (i.e., different punishment ratios). To assess changes in the free parameters of Eq. 2 between baseline and punishment, Eq. 2 was fitted to the data from each condition separately to allow changes in the values for sensitivity and bias across conditions. As in Experiment 1, sensitivity was low during baseline. The extra sum-of-

---

\(^9\) Note that baseline data are replotted across graphs. 

---

Note. Fits of Eq. 2 to each condition separately. Dashed lines show fits to the first baseline and solid lines show fits to each punishment condition. Note that baseline data and fits are replotted across graphs. Parameter values and R\(^2\) refer to fits to the punishment phase.

Asterisks indicate parameter values that were significantly different from baseline. Baseline fit: \(a = 0.46, \log(b) = -0.07, R^2 = 0.96\).

Significance codes: \(*** p < 0.001; ** p < 0.01; * p < 0.05\).
squares F test indicated no significant difference in slope between baseline and any of the punishment conditions, or across punishment conditions. Conversely, changes in bias were observed between baseline and punishment phases, as the introduction of unequal punishment rates shifted responding toward the leaner punishment schedule. The extra sum-of-squares F test indicated that bias was significantly different between baseline and four of the seven punishment conditions (i.e., 8:1, 1:2, 1:4, 1:8, \( p_s \leq .03 \)).

To track more clearly the changes in sensitivity and bias across conditions, Figure 11 shows the parameter values obtained from the fits of Eq. 2 to the daily data. Similar to the fits across conditions, the daily fits showed no systematic changes in sensitivity during punishment relative to baseline. Conversely, the introduction of different punishment ratios shifted responding away from the richer punishment schedule and toward the response associated with less frequent shock delivery, producing systematic changes in bias. Furthermore, the degree of the shift in allocation was a function of the punishment ratio, and more extreme ratios produced a greater bias. Besides the differences in bias between baseline and punishment, bias was also significantly different.

**Figure 11**

*Exp. 2 - GML (Eq. 2) daily fits*

*Note. Obtained values for sensitivity (\( a \)) and bias (\( \log b \)) from fits of Eq. 2 to data from each session. \( R^2 \) ranged from .82 to .99.*
across punishment conditions ($p < .01$; except that 2:1 was not different that 4:1 and 1:1, $p > .06$, and 1:4 was not different than 1:2 and 1:8, $p > .16$).

The increase in allocation toward the leaner punishment schedule was not predicted by Eq. 7 (see Fig. 2). According to Eq. 7, superimposition of unequal punishment ratios on unequal reinforcement ratios should only impact response allocation when the richer schedule of reinforcement was combined with the leaner punishment schedule. Figure 12 shows fits of Eq. 7 to data from all conditions simultaneously. In the fits, the parameter values were shared (i.e., not allowed to vary) across all conditions. Although Eq. 7 accounted for 92% of the variance in the data, the fits did not capture the

**Figure 12**

*Exp. 2 - Direct Suppression Model (Eq. 7) Fits*

*Note.* Eq. 7 global fit plotted by condition. Dashed lines show fit to the first baseline and solid lines show fits to punishment condition. Note that baseline data and fits is reprojected across graphs. Fitted parameters: $a = 0.47, \log b = -0.04, c = 0.16, R^2 = .92$.
increase in preference for the leaner punishment schedule (i.e., bias) as such change in allocation is not predicted by the model.

Fits of Eq. 2 and 7 were also compared. When Eq. 2 was fitted to data from all conditions simultaneously (i.e., same parameter values across conditions), values for sensitivity were different than obtained by Eq. 7 fits\textsuperscript{10}, but values for bias were similar between the models. The difference in sensitivity between the models was likely due to the higher value assumed by $c$ ($c = 0.16$) in Eq. 7. Because of the interdependency between sensitivity ($a$) and $c$ in Eq. 7, changes in the value of $c$ have a direct impact on the slope of the function, but not on the intercept (i.e., bias). In the model comparison using information criteria, Eq. 7 had the smallest AIC and BIC value and was used as the best model in the comparisons. The comparisons yielded $\Delta$IC $> 10$ and Eq. 7 outperformed Eq. 2 ($\Delta$AIC = 15.72, ER = $2.59 \times 10^3$; $\Delta$BIC = 13.78, ER = $9.82 \times 10^2$).

Thus, the modified version of the direct suppression model with the inclusion of the scaler for punishment (i.e., $c$) depicted in Eq. 7 better accounted for the effects of unequal punishment rates on response allocation than the generalized matching law (Eq. 2) that does not include any punishment parameters.

Changes in response allocation were also analyzed across successive reinforcer deliveries to assess the impact of unequal punishment rates on response allocation within each component. Figure 13 shows log response ratio across successive reinforcers for each component during baseline and punishment conditions where the richer punishment

\textsuperscript{10} Eq. 2: $a = 0.55$, $\log b = -0.04$, $R^2 = .89$.

Eq. 7: $a = 0.47$, $\log b = -0.04$, $c = 0.16$, $R^2 = .92$. 
schedule was programmed for B1. Figure 14 shows log response ratio across successive reinforcers for each component during baseline\textsuperscript{11} and punishment conditions where the richer punishment schedule was programmed for B1. The value of 0 on the x axis indicates response ratios prior to the first reinforcer delivery.

\textsuperscript{11} Note that baseline data are replotted across figures.
ratio. During punishment, response allocation shifted in the direction of the leaner punishment schedule. When the leaner punishment schedule was combined with the richer schedule of reinforcement (bottom panels in Fig. 13 and top panels in Fig 14), preference for the richer schedule of reinforcement started sooner and was more extreme than during baseline. Conversely, when the richer punishment schedule was combined with the richer schedule of reinforcement (top panels in Fig. 13 and bottom panels in Fig. 14), response allocation either remained the same or became less extreme than during baseline.

**Figure 14**

_Exp. 2 - Response allocation across successive reinforcers_

*Note.* Log response ratio following each successive reinforcer delivery in each of the seven components across conditions. Panels show data from each component that programmed different reinforcer ratio, as indicated at the top of each panel. Lines show data from baseline and punishment conditions where the richer punishment schedule was programmed for B2. The value of 0 on the x axis indicates response ratios prior to the first reinforcer delivery.
baseline, moving in the direction of the leaner punishment schedule. Furthermore, the degree of shift toward the leaner punishment schedule was a function of the punishment ratio, and more extreme punishment ratios produced more extreme changes in allocation relative to baseline.

A similar analysis was conducted to evaluate the impact of successive shock deliveries on response allocation. Because there were no criteria for the number of shocks received in each component, the number of shocks delivered varied across components, sessions, and rats. Thus, the shock-by-shock analysis only included the maximum number of consecutive shocks delivered to all rats in each component during each condition and later shocks experienced by some of the rats in each component were excluded from the analyses. Figure 15 shows log response ratio across successive shock for each component during the punishment conditions where the richer punishment schedule was programmed for B1. Figure 16 shows log response ratio across successive shocks for each component during the punishment conditions where the richer punishment schedule was programmed for B2. When the leaner punishment schedule was combined with the richer schedule of reinforcement (bottom panels in Fig. 15 and top panels in Fig. 16), preference for the leaner punishment option remained stable across shock deliveries, which decreased the number of shocks received in these components. Conversely, when the richer punishment schedule was combined with the richer reinforcement schedule (top panels in Fig. 15 and bottom panels in Fig. 16), preference for the richer reinforcement schedule gradually increased with each shock delivery, increasing the number of shocks received in these components. However, preference for the richer reinforcement options was not as
Once punishment was discontinued, response allocation recovered to its pre-punishment baseline (Fig. 11). Although there was a greater bias toward B2 during the second baseline than during the first, an extra sum-of-squares F test indicated no extreme with increasing shock deliveries as with increasing reinforcer deliveries (Fig. 13 and 14).

Note: Log response ratios following each successive shock delivered in each of the seven components across conditions. Panels show data from each component that programmed different reinforcer ratios indicated at the top of each panel. Lines show data from different punishment conditions where the richer punishment schedule was programmed for B1. More extreme punishment ratio (i.e., 1:8) is highlighted. The value of 0 on the x axis indicates response ratios prior to the first shock delivery.
significant differences in slope or intercept between the first and second baselines ($ps > .05$). Thus, there were no changes in sensitivity (i.e., slope) across conditions, and the changes in bias produced by unequal punishment rates did not persist once punishment was removed.
In summary, the present experiment showed that superimposition of unequal punishment rates on unequal schedules of reinforcement shifted allocation away from the richer punishment schedule, increasing bias toward the option associated with less frequent shocks. The results also showed that the degree of bias towards the leaner punishment schedule was a direct function of the punishment ratio and responding quickly adjusted to the changes in relative punishment rates. Although Eq. 7 outperformed Eq. 2 in the model comparisons, the changes in response allocation observed in the present experiment did not conform to the predictions of Eq. 7. According to Eq. 7, superimposing unequal punishment ratios on unequal reinforcement ratios should only impact allocation when the richer schedule of reinforcement is combined with the leaner punishment schedule (see Fig. 2). However, in the present experiment the introduction of unequal punishment rates shifted allocation toward the leaner punishment schedule, regardless of the reinforcement ratio, and the degree of the shift was a function of the punishment ratio. These results provide another piece of evidence of the limitations of Eq. 7 as a model of punishment. The results also demonstrated that the changes in response allocation during punishment are accounted by Eq. 2 through systematic changes in its free parameters. More specifically, the superimposition of unequal punishment rates on unequal reinforcement rates impacted the values obtained for bias in the fits of Eq. 2.

**GENERAL DISCUSSION**

The goal of the present experiments was to test the predictions of a modified version of the Direct-Suppression Model (i.e., Eq. 7) using a dynamic-choice procedure that allowed for a wide range of reinforcement and punishment schedules to be
investigated in a short period of time. According to the model, superimposing equal
punishment rates on responding maintained by unequal reinforcement rates should
increase preference for the richer schedule of reinforcement. Furthermore, preference
should become more extreme with increases in absolute punishment rate. This prediction
was tested in Experiment 1 and the results showed that the introduction of equal
punishment rates indeed increased preference for the richer schedule of reinforcement.
These results replicate previous experiments showing overmatching with the
superimposition of equal punishment rates on unequal reinforcement rates (e.g.,
Critchfield et al., 2003; Farley, 1980). However, in Experiment 1 the degree of preference
was not impacted by increases in absolute punishment rates. These results contradict both
the model predictions and previous studies showing more extreme preference for the
richer schedule of reinforcement with increases in absolute punishment rates (Critchfield
et al., 2003; Farley, 1980).

It is important to note, however, that a much wider range of reinforcement and
punishment rates were used in the present experiment than in the previous studies. For
example, both Critchfield et al. (2003 – Exp. 2) and Farley (1980 – Exp. 2) increased
absolute punishment rate while keeping relative reinforcement rates constant at a 1:1 or
1:2 ratio. Conversely, the present experiment programmed seven different reinforcement
ratios that varied from a 1:1 to a 1:8 ratio. The range of punishment rates investigated in
the previous studies was also smaller (2, 4 or 8 punishers/min). Thus, it is possible that
more extreme preference for the richer schedule of reinforcement with increases in
absolute punishment rates is only observed under specific combinations of reinforcement
and punishment rates. Furthermore, in previous experiments the schedules of
reinforcement and punishment changed across phases, and subjects were exposed to each condition for longer periods of time. For example, in Farley (1980) subjects were exposed to each combination of reinforcement and punishment ratios for at least 6 sessions. In Critchfield et al. (2003), subjects were exposed to each combination of reinforcement and punishment ratios until stability, with a maximum of 15 sessions. Conversely, in the present experiment the entire range of reinforcer ratios was experienced within session and punishment rates changed within few sessions. Thus, the short exposure to each combination of reinforcement and punishment rates might have impacted the magnitude of the effect and it is possible that greater changes in allocation would have been observed with longer exposures to punishment.

The lack of systematic changes in allocation with increases in overall punishment rates also could have resulted from the manner with which punishment rates increased across conditions. Rats were exposed to an initial low punishment rate (2/min) that increased across sessions in small steps. Previous studies have demonstrated that responding is more resistant to punishment when the severity of the punisher is gradually increased than abruptly introduced (e.g., Azrin & Holz, 1966; Chapter 3 of this dissertation). Thus, it is possible that changes in absolute punishment would have a greater impact on response allocation if the different rates were presented in a random order.

Although preference for the richer schedule of reinforcement decreased with the removal of punishment in Experiment 1, response allocation did not fully recover to its pre-punishment levels. The lack of reversibility during the reinstatement of baseline puts in question if the changes in response allocation and increase in sensitivity observed in
Experiment 1 was produced by the superimposition of punishment. For example, it is possible that the increase in sensitivity resulted from the continued exposure to the procedure. However, this explanation seems unlikely. If the change in allocation resulted solely from additional exposure to the procedure, a systematic increase in sensitivity should have been observed across sessions, regardless of the conditions. Instead, sensitivity fluctuated across punishment conditions and slightly decreased once punishment was removed. Furthermore, a similar increase in sensitivity was not observed with continued exposure to the procedure in Experiment 2.

Evidence of the reversibility of punishment effects on choice are scarce and few experiments have reestablished baseline after punishment (de Villiers, 1980; Kuroda et al., 2018; Rasmussen & Newland, 2008). For example, Rasmussen and Newland (2008) showed that sensitivity and bias returned to pre-punishment levels once punishment was removed. However, in de Villiers (1980) and Kuroda et al. (2018) response allocation did not return to its pre-punishment levels when punishment was discontinued. One possible explanation for such effect is that punishment might result in more optimal allocation across components. If we consider sensitivity as a measure of discriminability between the two options (Davison & Jenkins, 1985; Nevin et al., 2016), the introduction of equal punishment rates might increase the discriminability of the programmed reinforcement ratios, thus improving performance across components. If the rats do learn to better discriminate the options during punishment, they might continue to do so once punishment is removed, thus maintaining a more optimal response allocation. However, additional empirical support is necessary to sustain such explanation.
The model depicted in Eq. 7 also predicted that superimposing unequal punishment rates on unequal reinforcement rates should increase response allocation toward the richer schedule of reinforcement when the richer schedule is combined with the lower punishment rate. However, no change in allocation should be observed when the richer schedule of reinforcement is combined with the higher punishment rate, according to the model. Experiment 2 tested and refuted this prediction. Instead, the results of Experiment 2 showed that the introduction of unequal punishment rates shifted response allocation away from the richer punishment schedule creating a bias toward the option associated with leaner punishment rates. Furthermore, the degree of change in response allocation was a direct function of the punishment ratio and greater shifts in allocation were observed with more extreme ratios than with less extreme ones.

Similar results were reported by Deluty (1976). In that experiment, responding in a concurrent schedule was maintained by equal reinforcement rates and punished with unequal punishment rates. There too response allocation shifted away from the richer punishment schedule with changes in relative punishment rates. Rasmussen and Newland (2008) also reported a change in bias with the superimposition of punishment to one of two available responses maintained by equal or unequal schedules of reinforcement. Thus, the present results replicate and extend the previous findings by demonstrating changes in bias across a wide range of reinforcer and punisher ratios.

Another goal of the present study was to compare the GML (i.e., Eq. 2) to Eq. 7 and investigate how punishment impact the free parameters in Eq. 2. Although GML has previously outperformed the DSM in accounting for punishment data (Klapes et al., 2018), it is not reasonable that the best model to account for punishment effects on choice
is one that does not include any punishment parameters. Thus, the only way Eq. 2 could account for punishment effects is through unaccounted for changes in the values of its free parameters. Indeed, in both experiments the introduction of punishment produced significant changes in the fitted parameters of Eq. 2. While the introduction of equal punishment rates significantly increased sensitivity (Exp. 1), the introduction of unequal punishment rates had a greater impact on bias (Exp. 2). Furthermore, the changes in sensitivity and bias during punishment were observed both across phases and across sessions. Thus, for the GML to be considered as a punishment model, principled rules of how parameters of punishment impact sensitivity and bias must be established.

Additionally, although GML has outperformed DSM in previous comparisons (Klapes et al., 2018), in the present experiments GML failed to outperform a modified version of DSM (Eq. 7) that included a scaler parameter for punishment (i.e., $c$). The improvement of the model with the inclusion of the scaler parameter suggests that reinforcers and punishers are not equally weighted in choice situations. Furthermore, the greater value assumed by $c$ in Experiment 2 compared to Experiment 1 suggests that punishers that are delivered according to unequal rates might be weighted more heavily than punishers that are delivered according to equal rates between the options. Importantly, even though Eq. 7 outperformed Eq. 2 in the present experiments, the changes in response allocation observed with the introduction of punishment in both experiments did not conform to the predictions (see Fig. 1 and 2). Thus, although Eq. 7 might perform better than previous versions of the DSM, it still is a limited and flawed punishment model.
In summary, the present study adds to the punishment literature on choice by providing a more extensive data set with a wider range of reinforcement and punishment rates. Besides the limited number of studies investigating the effects of punishment on choice, those studies were done using a limited number of reinforcement and punishment ratios. Additionally, no studies have investigated the effects of punishment in a dynamic-choice procedure with nonhuman animals. Besides the advantage of increasing the number of conditions investigated in a short period of time, the dynamic procedure demonstrated that punishment produce quick changes in response allocation and responding adapted fast to changes in punishment parameters. Lastly, the present study tested relevant predictions from a modified version of the DSM and demonstrated that although the model outperformed the GML (Eq. 2), it did not fully account for the effects of punishment on choice. Furthermore, the results provided evidence that the one of the ways Eq. 2 can account for punishment data is through changes in its free parameters. Thus, for Eq. 2 to be considered as an adequate punishment model, rules of how punishment impacts sensitivity and bias must be established.
Acknowledgements

We have no known conflict of interest to disclose. The authors thank Gabrielle Sutton and Jack Van Allsburg for helping in conducting this research. This work was supported by Utah State University dissertation grant funding (RMF).


https://doi.org/10.1007/BF03392569


Punishment has been an understudied topic in behavior analysis and has been surrounded by misconceptions (Horner, 2002; Todorov, 2001, 2011). As a result, there has been little advance in our understanding of punishment effects in the past few decades. Regardless of one agreeing or not with the inclusion of punishment procedures in programmed contingencies, punishment is part of nature and has important implications for our daily lives (Critchfield, 2014; Vollmer, 2002). A better understanding of how punishment impacts behavior can improve our understanding of socially relevant problems, such as addiction. For example, resistance to punishment has been an increasing concerned for the treatment of drug abuse and identifying ways to increase the efficacy of punishers in suppressing behavior can have important implications for the treatment of addiction (Deroche-Gamonet et al., 2004; Durand et al., 2021; Vanderschuren & Everitt, 2004). Despite the evident relevance of punishment studies, little is known about how punishment works. Therefore, the goal of the present series of studies was to fill in some of the theoretical and empirical gaps in the punishment literature.

The first study (Chapter 2) explored some of the undesirable side effects commonly attributed to punishment procedures. A thorough review of the empirical data concerning each of the putative punishment side effects indicated that most of the arguments against the use of punishment lack strong empirical support. For example, some of the putative side effects attributed to punishment, such as increase in aggressive behavior, have only been observed with noncontingent presentation of aversive stimuli.
The transitory nature of punishment effects is also not inherent to punishment and depends on several aspects of the contingency. Furthermore, there is little evidence of the occurrence of punishment side effects in the applied literature, and some of those effects have been reported as desirable (e.g., conditioned punishers) in applied contexts. Thus, the review presented in Chapter 2 highlighted some of the misconceptions regarding punishment effects and emphasized the need for more empirical and theoretical research on punishment. Filling in the gaps in the punishment literature will contribute for the establishment of a well-grounded theory of punishment that incorporates punishment into a more general theory of behavior.

One of the evident gaps in the punishment literature refers to the effects of punishment on choice. Despite the richness of punishment studies using single schedules in the 60’s and 70’s, the number of punishment studies using concurrent schedules is more limited. Therefore, Chapters 3 and 4 focused on the effects of different parameters of punishment on response allocation in choice situations.

More specifically, Chapter 3 investigated the effects the effects of repeated exposures to escalating punishment intensities and repeated exposure to punishment after periods of vacation on response allocation between punished and unpunished responses. The results showed that superimposition of punishment on one of two available responses decreased allocation toward the punished option and increased allocation toward the unpunished options. Furthermore, the degree of change in allocation was a direct function of punishment intensity, and more abrupt changes in allocation were observed with the abrupt introduction of a mild constant intensity then with the introduction of low intensities that were gradually increased. These results replicated previous findings
showing that punishment for one response increases allocation toward competing options (Davison, 1970; Fontes et al., 2018; Rachlin, 1967; Reynolds, 1963), and that the degree of shift in allocation is a function of punishment intensity and the way punishment is introduced (Azrin, 1960; Azrin & Holz, 1966).

The results presented in Chapter 3 contribute to a better understanding of the effects of punishment intensity on choice. Understanding how changes in punishment intensity impacts response allocation is a necessary step to the establishment of a well-grounded theory of punishment. Currently, none of the punishment models include intensity as an independent variable, and ways to incorporate intensity into the available quantitative models have not been proposed. For example, it is unclear whether punishment intensity is analogous to reinforcer magnitude and whether punishment intensity and punishment rate impact response allocation in a similar way. Thus, more studies investigating the impacts of punishment intensity on choice are required before we have a comprehensive understanding of the mechanisms underlying punishment.

Chapter 4 focused on the effects of absolute and relative punishment rates on response allocation to test relevant predictions of a modified version of the Direct-Suppression Model - DSM (de Villiers, 1980; Farley, 1980) that included a scaler parameter for punishment. Furthermore, in previous studies, DSM has failed to outperform the generalized matching law - GML (Baum, 1974) in accounting for punishment data (Klapes et al., 2018), even though GML does not include any punishment parameters. Therefore, the experiments described in Chapter 4 also provided an extensive data set for the comparison of the two models.
According to DSM, superimposing equal punishment rates on unequal reinforcement rates should increase preference for the richer schedule of reinforcement, and preference should become more extreme with increases in absolute punishment rates. Although increases in preference for the richer schedule were observed during punishment in Experiment 1, the degree of preference was not a function of absolute punishment rate. The model also predicted that superimposing unequal punishment rates on unequal reinforcement rates should only change response allocation when the richer schedule of reinforcement is combined with the leaner punishment schedule. However, the results from Experiment 2 showed that unequal punishment rates biased allocation toward the leaner punishment schedule.

Although the results did not confirm the model predictions, the model did outperform GML. The improvement of the model with the inclusion of the scaler parameter suggests that reinforcers and punishers are not equally weighted in choice situations. However, the model is clearly still limited as it failed to provide accurate quantitative predictions. The results also showed that the way GML accounts for punishment data is through changes in its free parameters. More specifically, the introduction of equal punishment rates produced significant changes in sensitivity, while the introduction of unequal punishment rates produced significant changes in bias. Thus, for GML to be considered as a punishment model principled rules of how punishment impacts sensitivity and bias must be established.

Overall, the present set of studies provide important contributions to the punishment literature. At the theoretical level, the present studies helped to demystify some of the preconceptions regarding the use of punishment by showing the lack of
empirical evidence for punishment putative side effects (Chapter 2). The studies also provided important information about the impact of punishment parameters on response allocation (Chapter 3 and 4) that will contribute to the development of a well-grounded quantitative model of punishment. At the empirical level, the present studies add to the punishment literature on choice. The effects of punishment on choice have not been extensively investigated and a better understanding of such effects will improve our comprehension of the mechanisms underlying punishment.
References


https://doi.org/10.1901/jeab.1960.3-123


https://doi.org/10.1007/s40617-014-0005-4


https://doi.org/10.1901/jeab.1980.33-15


https://doi.org/10.1002/jeab.465


https://doi.org/10.1126/science.139.3557.838


https://doi.org/10.1901/jaba.2002.35-469
CURRICULUM VITAE
Rafaela M. Fontes
December 2021

CONTACT
Psychology Department, USU
2810 Old Main Hill, 84322
Logan, UT
Phone (cell): (435) 890-7710
Email: rafaela.fontes@usu.edu
rafaelafontes@hotmail.com

EDUCATION
2018 - Ph.D., Psychology with emphasis in Behavior Analysis
Psychology Department, Utah State University
Dissertation: Punishment and choice
Dissertation Chair: Timothy A. Shahan, Ph.D.

2015-2018 Ph.D., Behavioral Sciences
Department of Basic Psychological Process, Institute of Psychology,
University of Brasília, Brasília/DF, Brazil
Dissertation: Efeitos da punição de uma resposta alternative na
ressurgência de uma resposta alvo previamente extinta (The
effects of punishment for an alternative response on resurgence of
a previously extinguished target response)
Dissertation Chair: João Claudio Todorov, Ph.D.

2013-2015 M.S., Behavioral Sciences
Department of Basic Psychological Process, Institute of Psychology,
University of Brasília, Brasília/DF, Brazil
Thesis: Controle de estímulos em metacontingência (Stimulus
control in metacontingency)
Dissertation Chair: João Claudio Todorov, Ph.D.

2008-2013 B.A., Psychology
Department of Psychology, Federal University of Sergipe, Aracaju/SE,
Brazil
Thesis: Estudos sobre a adequação de uma interpretação
analítico-comportamental do teste de Rorschach (Studies on the
fitness of a behavior-analytical interpretation of the Rorschach
Test)
Chairperson: Diogo Conque Seco, Ph.D.

PROFESSIONAL INTERESTS
Experimental Analysis of Behavior
Punishment of operant behavior
Choice Behavior
Quantitative analysis of behavior
MEMBERSHIPS
2017- Member, Association for Behavior Analysis International
2017- Member, Society for the Quantitative Analyses of Behavior
2014- Member, Brazilian Association for Behavior Analysis

HONORS AND AWARDS
2021-2022  Anthony LaPray Scholarship
            Department of Psychology, Utah State University
2020-2021  Walter R. Borg Scholarship and Research Productivity Award
            Department of Psychology, Utah State University
2019-2020  Ray Alvord Scholarship
            Department of Psychology, Utah State University
2018       SABA Senior Student Presenter Grant
            Society for the Advancement of Behavior Analysis
2016-2017  Fulbright Doctoral Dissertation Research Award
            Fulbright & CAPES Brazil

POSITIONS HELD

RESEARCH POSITIONS
2021  Research Assistant
      Department of Psychology, Utah State University
      Punishment and choice – Departmental Grant
      PI: Timothy Shahan, Ph.D.

2020  Research Assistant
      Department of Psychology, Utah State University
      Quantitative models of punishment – Departmental Grant
      PI: Timothy Shahan, Ph.D.

2019  Research Assistant
      Alcohol seeking and resurgence following escalating negative
      consequences - (R21AA025604) National Institute on Alcohol
      Abuse and Alcoholism
      PI: Timothy Shahan, Ph.D.

2016-2017  Fulbright Visiting Researcher
            Department of Psychology, Utah State University
            Faculty Supervisor: Timothy Shahan, Ph.D.

TEACHING POSITIONS
2021 (Fall)  Guest Lecturer
            Center of Behavioral Theory and Research, Federal University of
            Pará
            Course: Methods in Experimental Analysis of Behavior

2021 (Fall)  Instructor
            Department of Psychology, Utah State University, Logan/UT
            Courses: 1400 - Analysis of Behavior: Basic Principles
2018-2021  Graduate Teaching Assistant  
Department of Psychology, Utah State University, Logan/UT  
Course: 1410 – Laboratory Accompanying Psychology 1400

2019-2020  Graduate Teaching Assistant  
Department of Psychology, Utah State University, Logan/UT  
Course: 3400 – Advanced Behavior Analysis

2018-2019  Graduate Teaching Assistant  
Department of Psychology, Utah State University, Logan/UT  
Course: 1400 - Analysis of Behavior: Basic Principles

2017-2018  Lecturer  
Department of Psychology, IESGO College, Formosa/GO, Brazil  
Course: Learning Psychology; Behavioral Theories

2017-2018  Lecturer, Supervisor  
Brasilia Institute for Behavior Analysis, Brasilia/DF, Brazil  
Courses: Experimental Analysis of Behavior I; Experimental Analysis of Behavior II; Verbal Behavior; Evaluation, Measurement and Research; Fundamentals of Radical Behaviorism; Epistemological Bases of Radical Behaviorism

2017  Instructor  
Department of Basic Psychological Process, Institute of Psychology, University of Brasilia, Brasilia/DF, Brazil  
Course: Introduction to Psychology  
Faculty Supervisor: Raquel Aló, Ph.D

2016  Lecturer  
Department of Psychology, IESGO College, Formosa/GO, Brazil  
Course: Behavioral Theories; Basic Psychological Processes

2015  Instructor  
Department of Basic Psychological Process, Institute of Psychology, University of Brasilia  
Course: Introduction to Psychology  
Faculty Supervisor: Raquel Aló, Ph.D

2013  Instructor  
Department of Basic Psychological Process, Institute of Psychology, University of Brasilia Brasilia/DF, Brazil  
Course: Introduction to Psychology  
Faculty Supervisor: Eileen Pfeiffer Flores, Ph.D

2010-2011  Undergraduate Teaching Assistant  
Department of Psychology, Sergipe Federal University, Aracaju/SE, Brazil
Courses: Theories and Systems in Psychology II, Laboratory Practice in Experimental Psychology
Faculty Supervisor: Diogo Conque Seco, Ph.D.

PUBLICATIONS

BOOK CHAPTERS


PEER-REVIEWED PUBLICATIONS


MANUSCRIPTS UNDER REVIEW


*Last name changed from Azevedo to Fontes in 2017

PRESENTATIONS

ORAL PRESENTATIONS


3) **Todorov, J. C., Henriques, M. B., Vasconcelos, I., & Azevedo, R. M. F.** (2015). Contingências e metacointingências no terceiro nível de seleção por consequências. Symposium conducted at the XXIV Brazilian Meeting for Behavioral Psychology and Medicine, São Paulo/SP, Brazil.


**POSTERS**


**PROFESSIONAL SERVICES**

**Editorial Activities**

Ad Hoc reviewer for:
- *The Psychological Record*
- *Journal of Applied Behavior Analysis*
- *Acta Comportamentalia: Revista Latina de Análisis del Comportamiento*
- *Revista Brasileira de Análise do Comportamento (Brazilian Journal of Behavior Analysis)*
- *Revista Interação em Psicologia*

**Organizational Service**

2021 **Symposium Chair**
47th Annual Convention of the Association for Behavior Analysis International

2021 **Poster discussant**
47th Annual Convention of the Association for Behavior Analysis International
SUPERVISION AND MENTORSHIP

Committee

2021  Elisama dos Santos Almeida, Undergraduate Thesis
      College of Psychology, Pará Federal University (Brazil)

Mentorship

2021-Present  Sara Sousa da Fonseca, Master's Thesis (co-advisor)
               Center of Behavioral Theory and Research, Federal University of Pará

2018  Juliana D. Sales, Specialization Thesis
       Certification Training in Clinical Behavior Analysis, Brasilia Institute for Behavior Analysis

2018  Milene V. Soares, Specialization Thesis
       Certification Training in Clinical Behavior Analysis, Brasilia Institute for Behavior Analysis

2018  Maria da Consolação Teixeira, Specialization Thesis
       Certification Training in Behavior Analysis with Emphasis in Autistic and Neurodiverse People – Brasilia Institute for Behavior Analysis

2018  Paula M. do A. Guimarães, Specialization Thesis
       Certification Training in Behavior Analysis with Emphasis in Autistic and Neurodiverse People – Brasilia Institute for Behavior Analysis

2017  Lara R. Queiroz, Specialization Thesis
       Certification Training in Behavior Analysis with Emphasis in Autistic and Neurodiverse People – Brasilia Institute for Behavior Analysis

ONGOING RESEARCH PROJECTS

*Effects of punishment on choice in a dynamic change procedure
  Faculty Investigator: Timothy Shahan, Ph.D.

*Effects of punishing one of two available alternatives on response allocation
  Faculty Investigator: Timothy Shahan, Ph.D.

*Lead student investigator