July 2017

The Effects of Predictability on Stereotypic Behavior in Nonclinical Adult Humans (Homo sapiens) and Rhesus Macaques (Macaca mulatta)

Amy Ryan

Recommended Citation

Ryan, Amy, "The Effects of Predictability on Stereotypic Behavior in Nonclinical Adult Humans (Homo sapiens) and Rhesus Macaques (Macaca mulatta)" (2017). *Doctoral Dissertations*. 972.
https://scholarworks.umass.edu/dissertations_2/972

THE EFFECTS OF PREDICTABILITY ON STEREOTYPIC BEHAVIOR IN NONCLINICAL ADULT HUMANS (*HOMO SAPIENS*) AND RHESUS MACAQUES (*MACACA MULATTA*)

A Dissertation Presented

by

AMY M. RYAN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

MAY 2017

Neuroscience and Behavior

THE EFFECTS OF PREDICTABILITY ON STEREOTYPIC BEHAVIOR IN NONCLINICAL ADULT HUMANS (*HOMO SAPIENS*) AND RHESUS MACAQUES (*MACACA MULATTA*)

A Dissertation Presented

by

AMY M. RYAN

Approved as to style and content by:

Melinda A. Novak, Co-Chair

Brian Lickel, Co-Chair

Elizabeth M. Jakob, Member

Jerrold S. Meyer, Member

Lynnette Leidy Sievert, Member

Elizabeth R. Dumont, Director,
Neuroscience and Behavior Graduate
Program

DEDICATION

To the Monkeys

May you find this dissertation easy to tear apart.

ACKNOWLEDGMENTS

How does a Jersey girl, meatball, daughter of historians, dropped in a middle of the suburbs, grow up to be a real monkey scholar? Sure, I got a lot farther by working at lot harder, by being a lot smarter, by being a self-starter. But I never would have made it without the support of my parents, Dennis and Joanne. Thank you for seeing that there are a million things that I haven't done, but you just waited... through my nine years of graduate school. Because of all of your help through the years, I didn't have to throw away my shot.

I am also grateful for the steady, patient, and supportive mentorship of my adviser, Dr. Melinda Novak. I truly feel that I learned from the best about rhesus macaques, animal behavior, and how to do science. Thank you for agreeing to take me on as a graduate student. I still remember how nervous I was when I asked you and the relief when you said yes! I am also thankful for Dr. Brian Lickel, who agreed to try out an idea from a 5th-year graduate student that he didn't even know. Learning how to study humans could not have been a more enjoyable experience with fascinating results without Brian's guidance. I am also thankful for Dr. Beth Jakob, Dr. Jerry Meyer, Dr. Lynnette Sievert, and Dr. Sarah Partan for their support and guidance throughout my PhD experience.

I am also grateful to my mentors that supported me throughout my career. Dr. Chris Kuhar got me into this research business. Dr. Mark Hauber, Dr. Jessica Rothman, Dr. John Fraser, and Dr. Sheila Chase provided me with opportunities that have shaped my career trajectory to this day. Knowing that my mentors from the Hunter ABC program believed in me was a source of confidence and strength through my PhD experience.

My fellow graduate students and my fellow sisters in Graduate Women in STEM. You inspired me because I must have been somewhat decent if I could even be in your company as a fellow graduate student. You got me through the tough times. By talking about yours, I felt like I was not alone in my struggles. Thank you.

Finally, I have to thank all of the animals in my life that have inspired me to spend my life in service to animals and their welfare. They keep me guessing, and I would not have it any other way. Teddy and Snickers were on the front lines as PhD Support Staff, making sure that every paper got laid on and pen knocked onto the floor. There was also Dusty, Monkey, Snooky, and Rocky holding up on the NJ Front. I can’t forget Pooky, Maggy, Frisky, Pansie, Huey, Blackie, Bunbun, TimTim, and Moogie. There was also Grandpa the spider monkey, Spike the gorilla, and Matara the hamadryas baboon. And Omar the betta fish and all of the crayfish that sassed their way into my heart. And the black and white colobus monkey that I had to run to keep up with as he migrated from one group to another in Kibale. And the wild chimpanzees of Kanyanchu that watched me as I watched them during one of the best days of my life. But most important of all, I have to thank my other PhD advisers: Ivan, Coby, Taz, Violet, Lily, Kayla, Zoey, Little G, Bailey, Linus, Emmitt, Fozzie, Friday, and Nigel. My work was only done in a partnership with them. They were indifferent to my learning and even more indifferent to my experiments, but teachers and partners nonetheless.

ABSTRACT

THE EFFECTS OF PREDICTABILITY ON STEREOTYPIC BEHAVIOR IN NONCLINICAL ADULT HUMANS (*HOMO SAPIENS*) AND RHESUS MACAQUES (*MACACA MULATTA*)

MAY 2017

AMY M. RYAN, B.S., THE COLLEGE OF WILLIAM AND MARY

M.A., HUNTER COLLEGE OF THE CITY UNIVERSITY OF NEW YORK

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professors Melinda Novak & Brian Lickel

Stereotypies, or repetitive and purposeless behaviors, are observed in both humans and other animals. They have been primarily studied in captive animal and clinical human populations with comparably little research devoted to understanding less severe levels of stereotypies observed in nonclinical populations of adult humans and in most captive animals. As these behaviors are sometimes associated with routine events, I explored the relationship between the predictability of anticipated events and mild stereotypies. I studied this relationship in captive rhesus macaques and a novel comparison group of adult humans from a nonclinical population. I designed two experimental paradigms, a wait paradigm and a task paradigm, to elicit stereotypic behavior in both species. I also provided participants with questionnaires about their current emotional state and individual trait differences. I found that while my manipulations of predictability did not spur differences in stereotypic behavior, both monkeys and humans performed stereotypic behavior in both the wait and task paradigms. Humans performed similar amounts of stereotypic behavior between the two paradigms and individual amounts of stereotypic behavior were positively correlated between paradigms. Yet, the rhesus macaques performed significantly more behaviors during the wait paradigm than in the task paradigm and their stereotypic behaviors between paradigms were not positively correlated, which suggests that they responded differently to the two scenarios. I then compared monkey and human stereotypic behavior during the wait paradigm that was a 5-minute uninterrupted period for both species. The human participants performed significantly more stereotypic behavior than the captive rhesus macaques—a highly unexpected result given that there has been little research devoted to stereotypies in nonclinical adult humans. One reason for this difference may be differences in typical stimulation levels between species as participants who reported feeling more bored performed more stereotypies. My results suggest that while stereotypies in captive animals are typically considered abnormal pathological behaviors that warrant intervention and mitigation, they may serve a function in response to the current environment that is retained across two species of primates. As intervention and mitigation are typically not proposed for mild levels of stereotypic behavior in nonclinical populations of humans, the results in this dissertation suggest that captive animal managers may need to reexamine management strategies for captive animals that perform mild levels of stereotypic behavior.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS v

ABSTRACT vii

LIST OF TABLES xv

LIST OF FIGURES xvi

CHAPTER

1. INTRODUCTION 1

    1.1 Stereotypic Behavior: Definition, Prevalence, and Function....................1
    1.2 Mild Stereotypies....................4
    1.3 Predictability: Operational Definitions and Embedded Concepts 5
    1.4 Predictability and Stereotypic Behavior 7

2. THE EFFECTS OF PREDICTABILITY ON BEHAVIOR IN ADULT HUMANS FROM NONCLINICAL POPULATIONS....................9

    2.1 Overview....................9
    2.2 Introduction....................10

        2.2.1 Stereotypies in Clinical Human Populations 10
        2.2.2 Stereotypies in Nonclinical Human Populations 12

            2.2.2.1 Prevalence of Stereotypic Behavior....................13
            2.2.2.2 Factors that Affect Stereotypic Behavior....................13

                2.2.2.2.1 Individual risk factors 13
                2.2.2.2.2 Environmental Risk Factors: Long-term 14
                2.2.2.2.3 Environmental Risk Factors: Short-term 14
                2.2.2.2.4 Environmental Risk Factors: Predictability 16

        2.2.3 Experiment 1: the effects of predictability on behavior in adult humans from nonclinical populations....................17
        2.2.4 Hypothesis & Predictions 18

    2.3 Methods....................19

        2.3.1 Participants....................19
        2.3.2 Experiment Room 20

2.3.3 Experimental Design 22
2.3.4 Experimental Procedures 23

2.3.4.1 Behavioral Paradigm 1: Experimenter Wait Paradigm 23
2.3.4.2 Behavioral Paradigm 2: Delayed Non-Matching to Sample task 24
2.3.4.3 Use of deception with participants....................................27
2.3.4.3.1 Participants' Understanding of the Purpose of the Study 28

2.3.5 Video Coding of Participant Behavior..29

2.3.5.1 Types of Stereotypic Behaviors Performed 31

2.3.6 Questionnaire materials 32

2.3.6.1 Presentation to participants 32
2.3.6.2 Generation of Test Variables from Questionnaires 34
2.3.6.3 Current Emotional State...34

2.3.6.3.1 Positive and Negative Affect Scale-Expanded (PANAS-X) (Watson and Clark, 1994) 34
2.3.6.3.2 Multidimensional State Boredom Scale (MSBS) (Fahlman, et al. 2013)..............................35
2.3.6.3.3 Generation of Combined Boredom measure 36

2.3.6.4 Individual Differences 36

2.3.6.4.1 Ten-Item Personality Inventory (TIPI) (Gosling, Rentfrow, & Swann, 2003) 37
2.3.6.4.2 Shortened Intolerance of Uncertainty Scale (Carleton, Norton, & Asmundson, 2007) 38
2.3.6.4.3 ADHD Current Symptoms Scale—Self-Report Form: Adults (Barkley and Murphy, 1998) 38
2.3.6.4.4 Behavior Inhibition/Behavioral Activation (BIS/BAS) Scale (Carver and White, 1994)..........39

2.3.6.5 Other Factors from Questionnaires....................................40

2.3.6.5.1 Mental Health check 40
2.3.6.5.2 Self-Report of Stereotypic Behavior..................41

2.3.7 Statistical Analyses 42

2.3.7.1 The effects of predictability on stereotypic behavior 42
2.3.7.2 The effects of emotional state and individual differences on stereotypic behavior 43
2.3.7.3 Mediation analyses of the effects of emotional state and individual differences on stereotypic behavior 43

2.4 Results..44

2.4.1 Characteristics of stereotypic behavior performance........................44
2.4.2 Correlations between paradigms..45
2.4.3 The effects of predictability on Stereotypic Behavior 46
2.4.4 Correlations between stereotypic behavior, emotional state, and individual differences..48
2.4.5 Mediation analyses on individual differences and emotion on stereotypic behavior...49

2.4.5.1 The effect of Behavioral Inhibition and Boredom on stereotypic behavior 49
2.4.5.2 The effects of Intolerance of Uncertainty and Boredom on stereotypic behavior 50

2.5 Discussion..51

2.5.1 The effects of predictability on stereotypic behavior 51
2.5.2 Characterizations of Human Stereotypic Behavior...........................57
2.5.3 Conclusion 58

3. THE EFFECTS OF PREDICTABILITY ON BEHAVIOR IN ADULT RHESUS MACAQUES 59

3.1 Overview...59
3.2 Introduction...60

3.2.1 Rhesus Macaque Natural History 60
3.2.2 Types of Stereotypic Behavior..61
3.2.3 Prevalence of Stereotypic Behavior..62
3.2.4 Hypothesized Functions of Stereotypic Behavior.............................63
3.2.5 Factors that Affect Stereotypic Behavior..64

3.2.5.1 Individual Factors 64
3.2.5.2 Environment...67

3.2.5.2.1 Long-term risk factors..67

3.2.5.2.2 Short-term risk factors 69
3.2.5.2.3 Predictability...70

3.3 Experiment 2: The effects of predictability of delays in a delayed response task on behavior in adult rhesus macaques.................................71

3.3.1 Hypothesis and Predictions..72
3.3.2 Methods..72

3.3.2.1 Subjects..72
3.3.2.2 Housing and Experimental Setting 73
3.3.2.3 Apparatus 74
3.3.2.4 DRT Procedure 75
3.3.2.5 Training...76
3.3.2.6 Data Collection Procedure 78
3.3.2.7 Video Coding of Macaque Behavior 80
3.3.2.8 Statistical Analyses 81

3.3.2.8.1 Assessment of potential sex and housing effects on behavior 81

3.4 Results...82

3.4.1 The effects of Predictable and Unpredictable conditions on behavior..82
3.4.2 Individual differences in DRT response 84
3.4.3 The effects of sex and housing factors on behavior in the DRT Experiment...85
3.4.4 Rare but Notable Behaviors..86

3.5 Discussion...87

3.5.1 Conclusions...89

3.6 Experiment 3: The effects of predictability of a waiting period on behavior in adult rhesus macaques 90

3.6.1 Hypothesis and Predictions..90
3.6.2 Methods..91

3.6.2.1 Subjects..91
3.6.2.2 Training...91
3.6.2.3 Wait Experiment Procedure...91
3.6.2.4 Baseline Morning Data Collection.....................................93
3.6.2.5 Sampling and Coding of Macaque Behavior 94

3.6.2.5.1 Assessment of Behavior during the Wait Experiment 94
3.6.2.5.2 Comparison with AM Baseline Data 95

3.6.2.6 Statistical Analyses 95

3.6.2.6.1 Assessment of Predictability of Waiting period on Behavior 95
3.6.2.6.2 The effects of habitation on responses to predictability 96
3.6.2.6.3 Comparison of Wait Experiment Conditions and baseline AM Data 96
3.6.2.6.4 Assessment of potential sex and housing effects on behavior 96

3.7 Results 97

3.7.1 The effects of Predictable and Unpredictable conditions on behavior 97

3.7.1.1 The effects of Predictability on behavior: duration measures 97
3.7.1.2 The effects of Predictability on behavior: modified frequency measures 98

3.7.2 The effects of Wait Experiment conditions on AM behavior 99
3.7.3 The effects of sex and housing factors on behavior in the Wait Experiment and AM Data 102

3.7.3.1 Duration Measures and effects on the two Wait conditions 102
3.7.3.2 Modified Frequency measures and effects on the two Wait conditions and AM Data 103

3.8 Comparison of Stereotypic Behavior across Experiments 105

3.8.1 Stereotypic Behavior: Comparison of macaque stereotypic behavior in Experiments 2 & 3 105
3.8.2 Human and Monkey Stereotypic Behavior 109

3.9 Discussion 111
3.9.1 The effect of predictability on stereotypic behavior in the Wait Experiment 111
3.9.2 Conclusion 114

4. GENERAL DISCUSSION 115

4.1 Complexities of Stereotypic Behaviors in Captive Animals 123
4.2 Conclusion 125

APPENDICES

A. MATERIALS FOR RECRUITMENT, CONSENT, DEBRIEFING FORM, AND QUESTIONNAIRES FOR PARTICIPANTS IN EXPERIMENT 1...................127
B. ANALYSIS OF PARTICIPANT SELF-REPORT OF STEREOTYPIC BEHAVIOR...145
C. ANALYSIS OF THE EFFECTS OF NON-PREDICTABILITY FACTORS ON STEREOTYPIC BEHAVIOR 150
D. PARTICIPANT DISTRIBUTION BY CONDITION AND GENDER 154
E. ETHOGRAM FOR EXPERIMENT 1 PARTICIPANT STEREOTYPIC BEHAVIOR...155
F. UMASS RHESUS MACAQUE DEMOGRAPHIC INFORMATION.......................157
G. ETHOGRAMS FOR RHESUS MACAQUE BEHAVIOR.......................................158
H. TESTS OF NORMALITY FOR EXPERIMENT 2 AND 3 161
I. THE EFFECTS OF SEX AND HOUSING ON BEHAVIOR IN EXPERIMENTS 2 & 3...162
J. ANALYSIS OF PREDICTABLE DELAY LENGTH IN PREDICTABLE AND UNPREDICTABLE SESSIONS IN EXPERIMENT 2.......................................172
K. ANALYSIS OF POTENITAL HABITUATION IN EXPERIMENT 3.....................174

REFERENCES 176

LIST OF TABLES

Table Page

2.1 Stages and conditions of Experiment 1....................23

2.2 The Wait paradigm conditions and what was manipulated between conditions. 23

2.3 The DNMS paradigm conditions and what was manipulated between conditions. 26

2.4 The order and purpose of administered questions and scales to participants. 33

2.5 Mean and Range for Stereotypic behavior measures in the Wait and DNMS Paradigms....................45

2.6 Pearson correlation coefficient (r) for Stereotypic behavior, Emotion, and Individual Difference measures. Bolded values represent p<0.05. *p<0.05, **p<0.01 48

3.1 Wilcoxon Signed-Ranks Tests for behaviors during DRT. Bolded values represent significant. *p<0.05; +p<0.06. 84

3.2 Characteristics of stereotypic behavior observed in macaques in Experiments 2 and 3....................107

3.3 Types of stereotypic behaviors and frequencies of these behaviors observed in macaques in Experiments 2 and 3....................108

4.1 Examples of anthropological accounts of stereotypic behaviors across cultures.....................122

LIST OF FIGURES

Figure Page

2.1 The room used in Experiment 1. The full desk, top of the desk, and behind the desk views are shown..22

2.2 Example stimuli used in the DNMS task. One type of stimulus consisted of four basic shapes with gradations of similar colors (a) and the other type was a grid of 9 squares of different colors (b). In the test condition, the participants had to identify the alternative stimulus. 25

2.3 The prevalence (total duration divided by number of participants who performed the behaviors) of stereotypic behaviors observed both in the Wait and DNMS paradigms combined. The numbers above the bars represent the number of participants who engaged in these behaviors............32

2.4 Total duration of stereotypic behavior observed in the Wait paradigm, by condition. Bars represent ±1 SEM. 47

2.5 Total duration of stereotypic behavior observed in the DNMS paradigm, by the two DNMS conditions (a), and with the conditions of the Wait paradigm included, thus creating four conditions (b). Bars represent ±1 SEM. 47

3.1 The apparatus used in the Delayed Response Task (DRT) for Experiments 2 and 3. The photos display the front (left) view, back (center), and side (right) views of the apparatus. The center image is Friday (N01) using the apparatus in an Allentown cage and the right image is Coby (V43) using the apparatus in a pen. 75

3.2 The logistics of the DRT visualized on the apparatus. The treat (a peanut, pictured) was baited into the center treat well (a). The yellow block was then slid over the treat (b). The white occluder (made of polypropylene plastic sheet) was then placed in front of the apparatus for the duration of the delay (c). When the delay ended, the occluder was removed and the monkey slid the yellow block back and obtained the treat..76

3.3 (a,b) Average frequencies of behaviors observed in Experiment 2, by condition (a). Average behavior duration observed in Experiment 2, by condition (b). Bars represent ±1 SEM. 83

3.4 Average duration of behaviors performed, by monkey. 85

3.5 (a,b) Frequency of yawn and scratch behaviors, by sex (a). Average bout duration of yawn and scratch behaviors, by sex (b). += $0.06>p>0.05$, **=$p<0.01$. Error bars represent ±1 SEM. 86

3.6 Main effect of sex in duration of cage shake and threat behaviors. 86

3.7 The total duration of behaviors observed, averaged across sessions, by condition. Error bars represent ±1 SEM. 98

3.8 The modified frequency count of behaviors observed, by condition. Error bars represent ±1 SEM. 99

3.9 The modified frequency count of behaviors observed, by the two Wait conditions and AM Data. Error bars represent ±1 SEM. 100

3.10 The rate and range of behaviors observed, by the two Wait conditions and AM Data. Error bars represent ±1 SEM. 102

3.11 (a,b) The main effect of sex on yawn and scratch behaviors (a). The interaction between sex and cage shake and threat behaviors. The post-hoc t-tests with a Bonferroni correction applied did not reveal a significant direction for the interaction (b).. 103

3.12 Main effect of sex on yawn and scratch behaviors. * $p<0.05$. Error bars represent ±1 SEM. 104

3.13 Main effect of housing on cage shake and threat behaviors (a) and on tactile and oral exploration behaviors (b). * $p<0.05$.. 105

3.14 The interaction between housing and condition for visual explore behaviors. The post-hoc t-tests with a Bonferroni correction applied did not reveal a significant direction for the interaction. Error bars represent ±1 SEM. 105

3.15 Total duration of stereotypic behavior observed in macaques in Experiments 2 (Wait) and 3 (DRT). The total duration time in the DRT was multiplied by (5/6, 0.833) in order to correct for the time difference between experiments. *$p<0.05$. Error bars represent ±1 SEM. 107

3.16 Stereotypic behavior observed in the Wait paradigm of Experiment 1 in humans and Wait experiment of Experiment 3 in rhesus macaques. Data were log transformed for analyses but are presented not-transformed in seconds. 111

4.1 Image of Fidget Cube from its website. 121

CHAPTER 1

INTRODUCTION

From an evolutionary perspective, much of animal behavior has been shaped and selected for the acquisition of resources such as food, shelter, and access to conspecifics. Captive animals and humans living in industrialized and human-managed environments have many of these needs provided for, yet many encounter scenarios in which access to these resources are managed by humans such as caretakers. From a psychological perspective, a largely unanswered question is how do animals spend their time while waiting for these desired resources? And, secondly, does information about the arrival of the resource affect psychological processes and behaviors while waiting? The following dissertation is an examination of the effects of predictability of the delay for an anticipated event on stereotypic behavior (or *stereotypies*) in human and nonhuman primates.

1.1 Stereotypic Behavior: Definition, Prevalence, and Function

Stereotypies are motor actions that are repeated continuously for a period of time and do not serve an apparent purpose (Edwards, Lang, & Bhatia, 2012; Mason & Latham, 2004). In this respect, the motor actions are not clearly detrimental to the animal, as opposed to self-injurious behaviors such as self-biting or cutting, yet also do not serve a clear goal of the animal such as attaining resources. Examples of stereotypies in humans (*Homo sapiens*) include flexion–extension of legs, tapping of limbs against a surface or one’s own body, repetitive object manipulation, and rocking. In captive animals, one of the most prevalent forms of stereotypy is pacing but other forms include swinging, rocking, and hair pulling, as observed in rhesus macaques (*Macaca mulatta*). These behaviors are sometimes considered an indicator of poor welfare for captive animal managers, although this notion is controversial given that humans also engage in these activities.

While stereotypies are characterized as purposeless behaviors, these behaviors most likely serve a psychological purpose. For captive nonhuman primates, herein referred to as *primates,* there are four leading hypotheses for why animals perform these behaviors and these may also be applicable to humans (Mason & Latham, 2004). The first two are related and considered divergent responses to the current environment. One hypothesis is that an animal may perform stereotypies in order to increase stimulation in response to an under-stimulating environment, whereas another hypothesis is that an animal may perform stereotypies in order to cope with a stressful or otherwise challenging environment (Mason & Latham, 2004). There is no *a priori* reason to conclude that one explanation is better than the other and indeed stereotypic behavior may serve different functions in different animals.

The final two hypotheses are not related to the current environment. The third hypothesis is that stereotypies reflect a previously developed habit rather than a response to the current environment, although it may reflect previous exposure to an under or over-stimulating environment (Mason & Latham, 2004). Finally, stereotypies may have an underlying physiological cause rather than environmental. For example, psychostimulants such as cocaine and methamphetamines increase stereotypic behavior (cocaine: Fowler, Covington, & Miczek, 2007; methamphetamines: Kitanaka et al., 2009) and the increase may persist after the cessation of drug use (Twohig & Varra, 2006). Genetic disorders seen in humans such as Fragile X syndrome (Newman, Leader, Chen, & Mannion, 2015) and disorders with abnormalities in brain development such as Autism Spectrum Disorder (ASD) (Ecker, 2016) also can be associated with increased stereotypic behavior.

Much of what we know about stereotypies comes from two populations. The first population comprises humans diagnosed with clinical conditions such as neurodevelopmental disorders, ASD, and obsessive compulsive disorder (OCD) (reviews of stereotypic behavior in ASD: Leekam, Prior, & Uljarevic, 2014; OCD: Stein et al., 2009). The second population is captive non-human animals. Stereotypies are prevalent and observed across species and captive settings, from pacing in zoo-housed carnivores to crib-biting in horses (biting and chewing of wood), and back-flipping in laboratory rodents. Mason and Latham (2004) estimated that approximately 85 million animals housed in farms, laboratories, and zoos display stereotypies.

Stereotypies can range in severity based on how disruptive the behavior is to an individual's typical behavior repertoire. In captive primates, stereotypic behaviors range from mild stereotypies that do not disrupt basic biological processes through severe stereotypic behavior in which the animal cannot be interrupted while performing these behaviors (Novak, Kelly, Bayne, & Meyer, 2012). Currently, stereotypic behavior that can be considered severe has been observed in captive animals. Yet, because of differences between species and housing arrangements, it is difficult to systematically classify severe stereotypies in captive animals. For humans, a severity scale is used only for people diagnosed with ASD. In this scale, stereotypic behavior is rated for severity yet all ranges on the scale are for stereotypic behavior levels that interfere with the ability to perform other activities (5th ed., DSM-V, American Psychiatric Association, 2013).

1.2 Mild Stereotypies

Severe stereotypic behavior has largely been the focus of scientific inquiry because it is highly disruptive and usually part of a suite of symptoms associated with neurodevelopmental disorders. However, it is widely observed that many if not most humans and captive animals engage in mild levels of these behaviors. Much less is known about the environmental triggers or function of nonclinical or mild stereotypic behavior performance in both humans and captive animals.

Understanding the function of stereotypies can shed light on common human behaviors that are currently given various names such as nervous habits, mannerisms, rituals, and fidgeting and can inform management strategies of captive animals. It is currently unclear whether mild stereotypies observed in captive animals are a constructive response to the environment or represent a problematic behavior that is simply not expressed at levels that warrant mitigation. I thus investigated the potential functions of stereotypic behaviors using a comparative approach by examining mild stereotypic behavior in nonclinical human populations and laboratory-housed rhesus macaques.

I explored the possible functions of stereotypies through an aspect of the environment that may vary in stimulation and is ecologically relevant to both humans and captive animals: the predictability of anticipated events. For captive animals, much of their day is characterized by the routine of animal husbandry events such as feeding, cleaning, and the distribution of enrichment. When not experiencing these events, it is possible that animals devote psychological resources to anticipating the event that will occur next. Provided with the prevalence of stereotypies across species and settings, understanding how captive animals respond to the predictability of anticipated events thus has the potential to inform captive animal management and improve captive animal welfare.

1.3 Predictability: Operational Definitions and Embedded Concepts

As an independent variable, predictability has the potential to be confounded with psychological constructs of certainty and control. The predictability of an event is an environmental factor. An event may have preceding cues in the environment, and thus be predictable, or occur randomly, and thus be unpredictable. The predictability of an event can be altered in two different ways, either by manipulating whether the event occurs or not or by varying the time at which the event will occur.

The construct of certainty pertains to how an individual interprets the environment and appraises how likely it is for the event to occur (Lazarus & Folkman, 1984). The probabilistic assessment of a situation through certainty is different than appraisal of expectations for the outcome, for people may have similar expectations about what will happen but have different beliefs about the likelihood of these expectations (Dickhäuser, Reinhard, & Englert, 2011; Schindler, Reinhard, & Dickhäuser, 2016). Notably, certainty may not be a solely human phenomenon. During a matching-to-sample cognitive task, macaques sought out more information about the sample or its comparisons when provided with an opportunity to do so, possibly indicating a motivation to increase certainty (Beran & Smith, 2011). Secondly, when provided with an option of an “uncertain” response during cognitive tasks that allowed the macaque to move on to another trial, macaques used the uncertain response in the more difficult trials (Beran, Smith, Redford, & Washburn, 2006). However, as certainty is a psychological construct, it cannot be assumed that a macaque that selected an uncertain response was feeling uncertain, for it could be using an alternate strategy in the cognitive task.

Finally, the construct of controllability refers to the sense of agency that an individual feels that he or she has in order to influence a situation. The influence can either be in terms of changing the environment itself or changing one’s response to the environment. The construct of control is strongly related to the predictability of the environment. For example, if an individual can predict the occurrence of an event, then he or she is more likely to feel in control of the situation because a response can be generated in anticipation of the event. While the relationship with predictable conditions and control is clear, unpredictable conditions can provide a varying sense of control. Encountering both unpredictable and perceived uncontrollable conditions contribute to anxious responses and in the longer term can contribute to the development of anxiety disorders (Zvolensky, Lejuez, & Eifert, 2000).

In this dissertation, the predictability of the environment was manipulated in terms of the timing of the event. Human participants were told that an event would happen, thus leaving no uncertainty about its occurrence, but I manipulated whether the participant was able to anticipate when the event will happen. Predictable conditions were defined as situations in which the participant knew exactly when the anticipated event would occur. Unpredictable conditions were defined as situations in which the participant did not know exactly when the anticipated event would occur. Temporal certainty thus varied between the predictable condition and the unpredictable condition. Rhesus macaques received signals that were as equivalent as possible about the timing of events although manipulations of certainty could not be so clearly assumed. Finally, in both predictable and unpredictable conditions, the participants and monkeys were unable to control when the anticipated event would occur. The sense of control was not expressly manipulated between conditions; however, participants may have felt more in control in the predictable condition as they could anticipate when the event would occur.

1.4 Predictability and Stereotypic Behavior

Stereotypic behavior is prevalent in both humans and other animals such as rhesus macaques. However, not much is known about the environmental triggers or function of these behaviors because for most humans and other animals, these behaviors are not disruptive to other activities. While considered purposeless behaviors, the results from previous studies suggest that stereotypic behavior may either serve a regulatory role in current environments or is a habit that reflects previous exposure to stressful environments. My overarching hypothesis for this dissertation was that stereotypic behavior is associated with changes in the predictability of events.

Importantly, a potentially revealing group for understanding the function of stereotypic behavior has not been studied within the context of environmental predictability: adult humans in *nonclinical* populations. I suggest that one reason for this is a discrepancy in how questions are approached in human and animal studies. While animal behavior is commonly studied, much of the work on humans in nonclinical populations is in the form of cognitive responses such as preferences, emotions, and other thoughts related to the events. In this respect, studies reveal that people largely prefer predictable events, but little is known about how people behave while they are waiting for events and whether this is affected by information about the event such as its predictability.

In my dissertation, I directly compared stereotypic behavior in adult humans and macaques using similar experimental paradigms. Analyzing stereotypic behavior in adult humans in this context will contribute to the sparse literature on human behavior. Additionally, using human participants allowed me to explore variables that are generally unavailable with rhesus macaques. Using a variety of self-report measures of internal state in humans, I investigated a set of psychological variables that possibly mediate (e.g. self-reports of boredom) or moderate (e.g. individual differences in temperament) the relationship between the predictability of environmental conditions and the performance of stereotypies. In captive macaques, behavior can be readily measured but the actual function of these behaviors (i.e. boredom or hyperstimulation) is much harder to assess. By including humans in a similar experimental paradigm, I could ask humans for the rationale behind their behavior and assess the value of this information for understanding stereotypic behavior in monkeys.

In these studies, I tested a number of questions. First of all, I assessed the response to predictability in both humans and macaques, especially with respect to stereotypic behavior. For both species, the question of whether stereotypic behavior can be differentially triggered by environmental conditions was analyzed. In addition, I compared stereotypic behavior observed between species in terms of the types of behaviors performed as well as the frequency and duration of behaviors expressed in similar conditions. Finally, the administration of questionnaires offered opportunities for participants to answer questions about how current emotional states and differences in individual traits may relate to the expression of stereotypic behavior in humans and possibly rhesus macaques.

CHAPTER 2

THE EFFECTS OF PREDICTABILITY ON BEHAVIOR IN ADULT HUMANS FROM NONCLINICAL POPULATIONS

2.1 Overview

The experiment with nonclinical adult humans had four goals. First, using an experimental manipulation in a laboratory experiment, I tested the effects of predictability on participants' stereotypic behavior. Second, working with nonclinical adult human participants allowed me to not only observe behavior, but to also ask participants questions about themselves and their behavior. My next two goals were related to these questionnaires, from which I sought to (1) assess how current emotional state may affect stereotypic behavior and (2) explore whether individual difference variables can account for variability in stereotypic behavior. Together, these goals allowed me to generate comparative data that, together with data from the macaque studies, provides a basis for some inferences about stereotypic behavior across primate species.

Because of the paucity of research on the immediate environmental triggers for stereotypic behavior performance in nonclinical adult humans, I aimed to design a paradigm that could elicit these behaviors in participants. In addition, I wanted the paradigm to be reasonably analogous to scenarios and methods that can be carried out with rhesus macaques (see Chapter 3). Ultimately, two paradigms were combined into a single experimental session for each participant. I refer to the first paradigm as the *Experimenter Wait paradigm* and the second paradigm as the *Delayed Non-Matching to Sample task*. Following these paradigms, each participant completed a set of questionnaires, which may provide useful information about how individual traits and psychological states affect the performance of behavior observed in the experiments.

2.2 Introduction

Repetitive behaviors in humans can be categorized as tics, mannerisms, nervous habits, rituals, fidgeting, compulsions, or stereotypies depending on the frequency or context of the behaviors (Singer, 2009; Edwards et al., 2012). In order to remain consistent between macaque and human experiments, in this dissertation, repetitive (3 or more times) and purposeless motor movements are considered stereotypies. However, it is possible that non-repetitive behaviors such as fidgeting and other habits may serve a similar psychological function as stereotypic behavior.

Most of the work on stereotypies in humans has pertained to humans with clinical diagnoses. I will first briefly review what is known about stereotypies in clinical populations. Since there are many avenues of research on this topic in clinical populations, I will focus on those that suggest a possible relationship between stereotypies and responses to predictability. Then, I will review the much sparser literature on stereotypies in the population of interest in this experiment, adult humans without clinical diagnoses who perform mild levels of stereotypic behaviors.

2.2.1 Stereotypies in Clinical Human Populations

Stereotypic Movement Disorder is considered a disorder in the DSM-V (2013). This disorder is diagnosed when stereotypic behavior occurs at levels in which the behaviors are disruptive to social, academic, or other activities. However, most studies of stereotypies in humans are in people who display these behaviors as one component of multi-faceted disorders such as ASD and other moderate to severe intellectual disabilities. Much of this research is concerned with the cause, maintenance, and treatment of stereotypies through pharmacological (reviewed by Rapp & Vollmer, 2005) or behavioral means. As there are many approaches to treatment, I will focus on ones that suggest environmental predictability may play a role in treating stereotypies.

ASD is a neurodevelopmental disorder in which the two of the leading criteria for diagnosis are repetitive behavior and an insistence on sameness in the environment (DSM-V, 2013). Insistence on sameness describes how people with ASD either strongly desire routine or ritual or have trouble with transitions or changes from their known routine (DSM-V, 2013). Behavioral interventions for people with ASD demonstrate that the creation of more predictable conditions can reduce the frequencies of stereotypic behavior. For example, Tustin (1995) reported a case in which an adult man with ASD performed fewer repetitive behaviors when he had 2-minute notice that he was going to change tasks as compared with conditions in which the change in tasks was sudden. Secondly, a common approach for increasing engagement in activities is through the use of visual activity schedules. This schedule is a system in which a sequence of events is depicted and consulted before a change in activity is initiated with a person with intellectual disabilities. When compared with baseline conditions without a schedule, visual activity schedules reduced the frequencies of stereotypic behaviors performed both in a home setting with parents creating the schedule (Krantz, MacDuff, & McClannahan, 1993) and in the classroom environment (Bennett, Reichow, & Wolery, 2011). These results suggest that predictable conditions such as a visual display of when events will occur may reduce stereotypic behavior. In my experiment, the predictable condition will be akin to these conditions using clocks to convey information about when an anticipated event will occur.

However, these relationships may not relate to people who do not have ASD as there are neuroanatomical brain abnormalities associated with ASD (Ecker, 2016). These abnormalities are suggested to contribute to reduced abilities to detect novelty as tested in an auditory event-related fMRI paradigm (Gomot et al., 2006). It is currently unclear whether these abnormalities are involved only in the interpretation of the environment or in both the interpretation and subsequent behavioral response of stereotypic behavior. People with ASD may be more sensitive to unpredictability. Yet, while the expression of stereotypic behavior is much greater and more disruptive in people with ASD, it is possible that the general response of increasing stereotypic behavior in response to unpredictability is a similar phenomenon in people both in clinical and nonclinical populations.

2.2.2 Stereotypies in Nonclinical Human Populations

Many adults in nonclinical populations engage in mild levels of stereotypic behaviors that do not disrupt other activities. However, possibly because intervention or treatment is not sought for these behaviors, not much is known about the trigger for or function of mild stereotypic behavior in humans. I will review what is currently known about prevalence and risk factors for the performance of mild levels of stereotypic behavior.

2.2.2.1 Prevalence of Stereotypic Behavior

Researchers have consistently found that adult humans in nonclinical populations (college students) largely engage in stereotypic behavior when queried through questionnaires, although these behaviors are usually combined with fidgeting and other nervous habits. In Hansen, Tishelmian, Hawkins, & Doepke's (1990) survey, all 286 participants reported having habits. However, people may not perform these behaviors very often, or at least do not perceive frequently performing these behaviors. When behaviors were required to occur multiple times a week in order to be considered a habit, behaviors such as touching the face or leg shaking were found in 37.5% and 47.2% of the queried population, respectively (Woods, Miltenberger, & Flach, 1996), and 26% of another sample reported engaging in body rocking (Rafeli-Mor, Foster, & Berkson, 1999).

2.2.2.2 Factors that Affect Stereotypic Behavior

2.2.2.2.1 Individual risk factors

There is not much research on individual factors that may contribute to stereotypic behavior in nonclinical adult humans. With respect to sex differences, self-reports of body rocking demonstrated that body rocking was more prevalent in males as 36% of males reported body-rocking as compared to 20% of females (Rafeli-Mor et al., 1996). Other studies have not found a consistent sex difference, for there were no differences found in self-reports of fidgeting behavior (Mehrabian & Friedman, 1986) and differences between the sexes were inconsistent between years of behavioral observations of habits in college students (Young, 1947).

2.2.2.2.2 Environmental Risk Factors: Long-term

Human and captive animal researchers focus on different factors in the environment with respect to the relationship between the environment and stereotypic behavior. Much of the captive animal research pertains to long-term variables such as the social, developmental, and housing backgrounds of animals and how they relate to the development of stereotypic behavior in an animal’s repertoire. In contrast, human research mostly encompasses short-term factors and how they contribute to the immediate expression of stereotypies.

Yet, opportunities have occurred for researchers to study the relationship between long-term environmental variables and the development of stereotypic behavior in humans. The long-term social environment appears to influence human stereotypic behavior as it does captive animals. When children were naturally subjected to similar rearing conditions as nursery-reared laboratory macaques, stereotypies were more likely to develop, as also observed in macaques. When Romanian children raised in orphanages with minimal adult contact were studied, 84% of adoptees displayed stereotypies when living in adopted homes in Canada, whereas Canadian children who were not adopted or institutionalized displayed no stereotypies (Fisher et al., 1997). Beyond the social environment, how the long-term housing environment or any other long-term environmental factors may relate to stereotypic behavior has been studied in rhesus macaques but not in humans.

2.2.2.2.3 Environmental Risk Factors: Short-term

The few human studies on the relationship between environmental factors and stereotypic behavior have been mainly through short-term environmental manipulations that contributed to an increase in stereotypic behavior. In humans, observations and experimental manipulations suggest that stereotypies increase in potentially challenging situations such as in public-speaking situations or waiting to visit the dentist. The researchers in these studies tend to label these scenarios as being *anxious* or *anxiety-inducing* but levels of these states were not specifically measured or assessed. Three studies represent the current knowledge of environmental triggers to adult human stereotypic behavior.

Barash (1974) investigated stereotypic and fidgeting behaviors in patients and escorting nonpatients in a dentist's waiting room. Barash quantified repetitive stereotypes (frequency of finger, foot, or forehead tapping), rate of magazine page turning, and fidget frequency. He found that both patients and nonpatients displayed stereotypies but that dental patients performed more of these behaviors than nonpatients. He suggested that these increased activities in patients represented fear of the ensuing dentist visit.

In a direct experimental manipulation, Woods and Miltenberger (1996) quantified the nervous habits of undergraduate students in three experimental conditions called *bored*, *anxious*, and *neutral*. The conditions were 10 minutes of neutral (watching a video), anxiety (told that he or she was going to give a presentation of an article) and bored (participant sat and asked to do nothing) in a within-subjects design. Overall, hair and face manipulations were higher in the anxiety than in the bored condition, whereas object manipulation was higher in the bored than anxiety condition. Notably, these results are consistent with the potential self-regulatory function of stereotypic behavior as operating both when an individual is in an under or over-stimulating environment.

Finally, Lang et al. (2015) investigated the effects of challenging situations on what they called *spontaneous ritualized behavior* in an experimental manipulation. Participants were assigned to conditions call *high anxiety* or *low anxiety*. Consistent with Woods and Miltenberger's paradigm (1996), the preparation of a public presentation was considered a challenging situation. In this case, the high-anxiety group was instructed to prepare a public presentation on a decorative object for an art expert whereas the low-anxiety group was instructed to think about the decorative object. After the manipulation, participants were told to clean the object with a wet cloth. The researchers then analyzed the time spent cleaning the object and the nature of the hand movements used in order to characterize redundant and repetitive movements. Overall, there were more repetitive cleaning movements observed in participants in the high-anxiety condition than in the low -anxiety condition. Additionally, participants who reported high levels of anxiety during the task made more redundant movements while cleaning the object than people who felt less anxious (Lang et al., 2015). Along with the previously discussed studies, these results suggest that stereotypic behavior may be related to the need to cope with low or high stimulation in the environment.

2.2.2.2.4 Environmental Risk Factors: Predictability

In contrast to captive animals, there are no studies to date on the relationship between stereotypic behavior and the predictability of events in the general human population. Yet, people frequently encounter similar situations as captive animals in terms of experiencing temporal uncertainty for an anticipated event. For example, this scenario is common when waiting in lines or waiting for the arrival of a train or bus.

Although consumer preference in these situations is frequently studied, I know of no studies on how people behave while they are waiting.

Consumer preferences and self-reports may help gain insight into how people respond to predictability. Researchers in the consumer behavior and transportation fields have found that people largely have a preference for predictable or certain wait conditions as opposed to conditions that are unpredictable. Customers actively seek out information for their wait duration (Pamies, Ryan, & Valverde, 2016) and have considered a wait more acceptable if they have duration information (Hui & Zhou, 1996) or are progressing in their position in a line (Munichor & Rafeli, 2007).

Additionally, the ability to predict when a bus or train will arrive can be considered a naturalistic scenario of humans facing predictable or unpredictable conditions in anticipation of a specific event. Recent developments in technology create opportunities for real-time information on public transportation via countdown clocks posted in a station or available via smartphone applications (apps). Smartphone apps increased ridership and satisfaction with the service (Brakewood, Macfarlane, & Watkins, 2015; Watkins, Ferris, Borning, Rutherford, & Layton, 2011). While preference is a useful metric, an assessment of potential behavioral differences in response to predictable or unpredictable conditions may elucidate how nonclinical adults respond to these commonly encountered situations.

2.2.3 Experiment 1: the effects of predictability on behavior in adult humans from nonclinical populations

The procedure for the Wait paradigm, described in more detail in the Methods section, entailed video recording the participant’s behavior as he or she waited alone in a room at a desk for an experimenter to arrive. Barash (1974) observed people in a dentist waiting room and recorded stereotypic behaviors in both patients and nonpatients, so I created a similar waiting room scenario in order to increase the likelihood of eliciting stereotypic behavior from participants. Each participant waited for the same amount of time but I manipulated the information about the wait in terms of whether the participant could predict when the anticipated event of the arrival of the experimenter will happen.

The second test used the delay in a delayed non-matching to sample task (DNMS). The DNMS is a visual recognition test in which a stimulus (sample) is initially presented and then removed. Following a delay, two stimuli are presented—the sample stimulus and the novel stimulus. For a correct response, the participant must select the novel stimulus. As with the wait paradigm, the lengths of the delays in this task were the same across participants and what was manipulated was information about when the anticipated event of the two test stimuli would be presented.

2.2.4 Hypothesis & Predictions

Previous literature on human response to uncertainty suggests that humans prefer predictable situations to ones that are less predictable (Hui & Zhou, 1996; Munichor & Rafeli, 2007; Watkins et al., 2011; Brakewood et al., 2015). I hypothesized that the frequency of stereotypic behavior would be related to the predictability of the event. More specifically, I predicted that as people prefer being in predictable conditions, participants would find the unpredictable condition more challenging and perform more stereotypic behavior than in a predictable one for both the Wait and DNMS paradigms.

In addition to investigating the relationship between stereotypies and predictability, human participants provided an opportunity to learn more about trait differences and how immediate self-reported emotional states related to stereotypic behavior. With respect to self-reported emotional states, hypothesized functions of stereotypic behavior entail either using the behaviors to increase stimulation in an under-stimulating environment or decrease stimulation in a challenging or over-stimulating environment (Mason & Latham, 2004). Because stereotypic behavior can be a response to the current environment, I hypothesized that there will be emotional states related to stereotypic behavior. Furthermore, because the responses to the current environment are about increasing or decreasing stimulation, I predicted that emotional states related to boredom would be the emotions with the strongest relationship to stereotypic behaviors. I thus conducted an exploratory analysis on potentially relevant individual differences but did not have firm *a priori* predictions for how they would relate to stereotypic behavior.

2.3 Methods

2.3.1 Participants

The participants in this study were undergraduate students from a nonclinical population at the University of Massachusetts Amherst who were enrolled in undergraduate-level Psychology classes in the Spring 2016 semester. Participants were recruited via the SONA system in which they participated in exchange for experimental research credit that was applied to certain Psychology classes. I invited qualifying students who completed the Spring 2016 Prescreen Questionnaire to participate in this study by email (see Appendix A). In order to participate in the study, students needed to have 20/20 or corrected-to-normal vision and not be color blind in order to complete the DNMS task.

Eighty participants completed the study, and ultimately, 68 were included in the data analysis. The first six participants were excluded because I tweaked the methods after running them, as participants unexpectedly retrieved their mobile phones that were out of the testing space but still in the testing room. Five more participants were further eliminated because of methodological errors during their testing session. One participant was eliminated because she was the only participant outside of the 18-22 years old age range of the rest of the participants (age = 35).

Of the participants included in the subsequent analyses, 75% (n=51) identified as female and 25% (n=17) identified as male. The age range was 18-22 years old and the average age was 20 years old. The prescreen questionnaire that the participants took prior to the experiment asked demographic questions beyond what was included in my study questionnaire. One potentially relevant question pertained to the participants' racial identity. The majority of the participants (55.9%, n=38) identified as *White* and 14.7% identified either as *Asian* or *Black or African American* (n=10 for both racial identities). An additional 11.8% (n=8) identified as having more than one race, and 2 participants did not provide an answer to the question. Appendix D displays how the 68 participants were distributed by experimental condition and according to gender.

2.3.2 Experiment Room

The participants completed the consent process (see Appendix A for Consent forms) and left their belongings in an adjacent room. The experimental area of the testing room (Figure 2.1) was blocked off from the rest of the room with an opaque black curtain partition. The section of the room used in this experiment was 8'x 6'.

The furniture in the experimental area consisted of a desktop computer on a desk (Figure 2.1). The participant was seated in a chair at the desk. The chair was stable and not designed to rock or swivel. As for the desk, the computer's tower was on the floor and the surface of the desk had the computer monitor, mouse, speakers, and keyboard. All of these components were functional and served their respective purposes during the computer portion of the study in which participants had to use the mouse to select answers, the keyboard to type responses to other answers, and the speakers provided auditory feedback during the DNMS task. In addition to the computer components, there was a ceramic cup that contained pens on the desk. As there was no paper in the room and there was no writing involved in the study, this cup and its pens served no purpose during the study. Finally, a small digital clock was also on the desk. Depending on the condition for the Wait paradigm, this clock was either turned off and displayed a blank screen (Unpredictable) or turned on and displayed the correct time (Predictable).

There were three cameras used to record a participant's behavior. All three were clearly visible to the participant. The first camera was a GoPro Hero 3 positioned above the computer monitor and second camera was a GoPro Hero 3 positioned behind the participant in order to capture full-body movements. The third camera was a Sony HandyCam camcorder positioned behind the participant to record the computer screen.

Figure 2.1 The room used in Experiment 1. The full desk, top of the desk, and behind the desk views are shown.

2.3.3 Experimental Design

The procedure described below was approved by the University of Massachusetts Amherst Institutional Review Board (IRB). The experiment consisted of a between-subjects design with two behavioral paradigms. Each participant was exposed to one condition in each paradigm within a 1-hour session. For both paradigms, the conditions manipulated how an anticipated event was presented. There were two conditions for the Wait paradigm and four in the DNMS test as the condition encountered during the Wait paradigm may have affected the response to conditions in the DNMS task. Following the experimental manipulations, each participant completed the same set of questionnaires. Each participant thus received one of the four possible pairings of the two experimental paradigms outlined in Table 2.1.

Table 2.1 Stages and conditions of Experiment 1.

1. Wait Paradigm	2. DNMS Paradigm	3. Questionnaire
Predictable	Predictable with Clock	Same questionnaire
	Unpredictable with Clock	Same questionnaire
Unpredictable	Predictable without Clock	Same questionnaire
	Unpredictable without Clock	Same questionnaire

2.3.4 Experimental Procedures

2.3.4.1 Behavioral Paradigm 1: Experimenter Wait Paradigm

Following completion of the consent process, participants were brought over to the testing room that already had cameras on and recording. The experimenter told the participant that "*I am currently training another research assistant to help me run this study, so I was hoping that he would be here to help me start the study with you. He seems to be running late*." The experimenter was holding her phone in order to create the belief in the participant that the experimenter present in the room was in communication with the other experimenter who was running late. In actuality, the experimenter was using the phone's stopwatch function to initiate and track the 5-minute delay for the Wait experiment. What the experimenter said next about waiting for the other experimenter depended on the predetermined condition and is described in Table 2.2.

Table 2.2 The Wait paradigm conditions and what was manipulated between conditions.

Wait Paradigm	Predictable	Unpredictable
Digital Clock status	ON	OFF
Experimenter Statement	*looks at clock in room* *"*It's *whatever time it is* now so let's give him until *add 5 minutes* to show up."	"I'd like to give him about 5 more minutes to show up."

The experimenter received confirmation from the participant that it was ok to wait for the second experimenter. No participant indicated a problem with this scenario. The experimenter then said that she would wait for the second experimenter outside and left the room. The participant was left alone in the room with the door nearly closed for 5 minutes and the experimenter returned to the room across the hall with the participant's belongings and consent forms.

At the end of the 5-minute waiting period, the experimenter reentered the testing room and said "*Thank you for your patience. Sorry but my research assistant still hasn't arrived, so instead I will set you up for running the computer task*." The experimenter turned the computer monitor on, and the computer guided the participant through the rest of the study. The experimenter then left and the participant was alone in the experiment room with the door closed as the experimenter remained across the hall with the door open. The experimenter did not return to the room while the participant completed the DNMS task and questionnaires.

2.3.4.2 Behavioral Paradigm 2: Delayed Non-Matching to Sample task

The second behavioral paradigm tested stereotypic behavior while waiting for anticipated events during a cognitive task. This paradigm entailed a DNMS task completed on a computer. Both the DNMS and subsequent questionnaires were created in and administered to participants via the online-based Qualtrics® survey platform. In this study, there were two kinds of stimuli used that were designed to challenge the working memory of the participant (K. Cave, personal communication). The first kind of stimuli was a group of simple shapes that varied slightly in colors that defied simple labeling

(Fig. 2.2a), such as all being slightly different gradations of blue. The second kind of stimuli was a grid of nine squares in which each square was a different color (Fig. 2.2b).

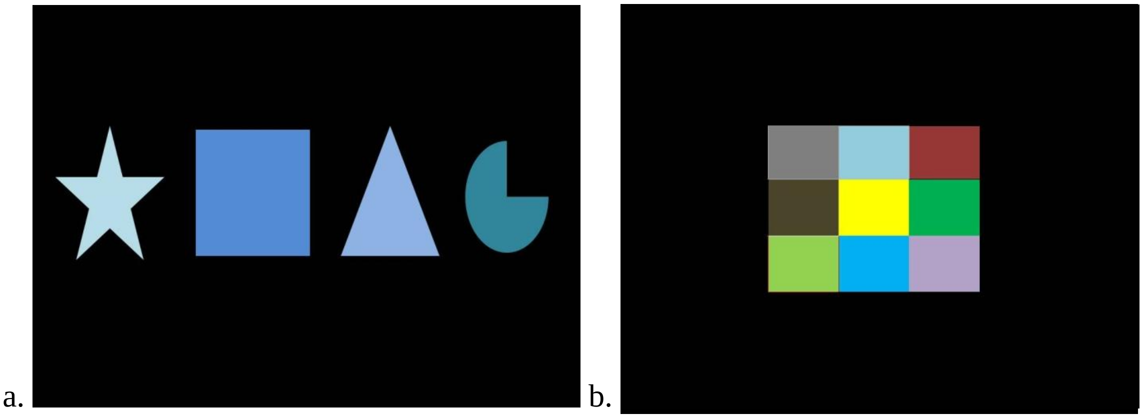

Figure 2.2 (a,b) Example stimuli used in the DNMS task. One type of stimulus consisted of four basic shapes with gradations of similar colors (a) and the other type was a grid of 9 squares of different colors (b). In the test condition, the participants had to identify the alternative stimulus.

The initial stimulus was presented for 5 seconds. Following the delay, two similar-looking stimuli were presented and the participant had to select the novel stimulus. There was no time limit placed on this decision, although time was tracked and all participants took less than 1 minute per answer choice. Following the participant's choice, he or she was presented with visual and auditory feedback. There were seven DNMS trials in each participant's session. For each participant, the seven trials were the same in terms of the order of stimulus presentation and the length of the delay. What was manipulated was the information conveyed to the participant about the delay.

In the predictable condition, following the initial stimulus presentation, the length of the delay was both indicated with the words "Delay of [time of delay]" and a countdown clock on the screen. In the unpredictable condition, the length of the delay was not clearly indicated, for the message on the screen read “Average delay of [time of delay]” and there was no countdown clock on the screen. The length of the delay varied between trials and ranged from one to four minutes. For the seven questions, each participant received delays in the following order: 1 minute and 30 seconds; 4 mins; 2 mins and 30 seconds; 2 mins; 1 min; 3 mins and 30 seconds; and 3 mins.

Because I did not want to call attention to the digital clock placed on the desk by turning it on or off in the middle of a participant’s session, the clock remained on or off from the Wait paradigm during the DNMS test. While the DNMS had 2 conditions within the computer task, there were actually four conditions as the digital clock may have affected the predictability of the DNMS paradigm. The four conditions of the DNMS paradigm are outlined in Table 2.3.

Table 2.3 The DNMS paradigm conditions and what was manipulated between conditions.

Wait Paradigm	Predictable		Unpredictable	
DNMS Paradigm	Predictable with Clock	Unpredictable with Clock	Predictable without Clock	Unpredictable without Clock
Digital Room Clock	ON	ON	OFF	OFF
Computer Countdown Clock	YES	NO	YES	NO
Text on Screen	“Delay of [1 minute, 30 seconds]”	“Average delay of [1 minute, 30 seconds]”	“Delay of [1 minute, 30 seconds]”	“Average delay of [1 minute, 30 seconds]”

Following the seven trials, directions on the computer indicated that the cognitive task was over and the questionnaire would appear on the next page. The participants then completed the questionnaires, described further in Section 2.3.6. After the questionnaire, instructions on the computer told the participants that they completed the study and directed them to the adjacent room to let the experimenter know that they were finished. Participants were then debriefed about the true purpose of the study.

2.3.4.3 Use of deception with participants

In order to capture naturalistic responses to my experimental conditions, deception was used in the study. I told participants that the study was a visual discrimination task and the camera on the desk was recording their visual search strategies as they engaged in the visual discrimination task on the computer. In actuality, performance on the cognitive task was not the purpose of the study but rather it was their behavioral response during the delays in the task. Secondly, waiting for another experimenter was a part of the study even though it was presented to the participant as an unexpected addition to the experimental session. Finally, each participant was told that for every correct answer on the visual discrimination task, he or she accumulated a raffle ticket for a prize. Instead, each participant earned an equivalent reward of a single raffle ticket for a prize for their participation.

I used the explanation of a visual discrimination test in order to prevent the participants from guessing the experimental conditions and purpose of the task as well as why they were being video recorded. If I told the participants that I was interested in their behavior during the task, then this may have caused the participants to monitor their behaviors and possibly inhibit stereotypic behavior. As for the deception with earning raffle tickets, I told the participants that they were earning raffle tickets in order to increase their motivation to fully participate in the task and earn an incremental reward for correct answers. However, because visual discrimination was not the true purpose of the study, each participant received the same raffle ticket reward no matter his or her visual discrimination ability.

All participants were debriefed at the end of the study and told the real purpose of the computer task and overall study (see Appendix A for form). With knowledge of the true purpose, each participant had the option to withdraw his or her data and still receive SONA participation credit and a raffle ticket for participating in the study. No participants elected to have their data withdrawn.

2.3.4.3.1 Participants' Understanding of the Purpose of the Study

In order to test whether the deception used in the study was effective, one of the final questions of the study was "*When you were completing the visual discrimination task with the colored stimuli, what did you think that the purpose of the study was*?" I phrased the question in this way to assess how the participants felt during the DNMS paradigm but before completing the questionnaires. Once the DNMS paradigm ended, participant behavior was no longer measured. Because I was no longer assessing their behavior, the participants answered questionnaires that had nothing to do with their visual discrimination ability which was the stated purpose of the study. While answering the questionnaires, it is likely that participants may have suspected that the study was not about their visual discrimination abilities. However, it cannot be assumed that participants followed my direction for answering this question and instead wrote what they currently thought the purpose of the study was after completing the questionnaires.

I qualitatively analyzed the participants' written responses to this question. Ultimately, four categories of responses were generated: Memory, Discrimination, Eye, Wait, and Other. I categorized responses that discussed concepts related to these terms. Participant responses could have included more than one of these categories. Through the categorizing of responses, 28 participants discussed Memory, 20 discussed Discrimination, 8 discussed Eye, and 6 discussed possible purposes that did not fall within these categories.

Twenty-two participants believed the purpose of the study was to investigate something related to waiting and the delays rather than the cognitive task. Importantly, even if this belief was held, no participant specifically discussed repetitive behaviors or predictability. As some participants received different conditions for the Wait and the DNMS paradigms, it is meaningful that no participants reported perceiving a difference between the paradigms related to the differences in condition. None of the participants also provided reports that could suggest that they discussed the study with previously debriefed participants. In addition, some participants discussed how their belief about the study purpose changed over the course of the study between the DNMS paradigm and the questionnaires. Ultimately, no participants were eliminated from subsequent analyses because they possibly understood the true purpose of the study. Even if participants believed that the purpose was a waiting, the repetitive behaviors that I observed during the Wait and DNMS paradigms still occurred despite any suspicions that I was filming them in order to measure these behaviors.

2.3.5 Video Coding of Participant Behavior

Participant behavior was scored via videotape from the two GoPro Hero 3 cameras. These cameras provided a ventral and back/full body view of the participant sitting at the desk. Videos were scored on computer at the frame-level (30 frames=1 second) using MPEG® Streamclip software. All observers (n=3) completed CITI training and the additional undergraduate human subjects Research Assistant workshop in the Department of Psychological and Brain Sciences before they viewed the participants' videos. All scorers achieved interobserver reliability above 90% and remain blinded to the experimental condition in the video.

I generated my own ethogram for stereotypic behaviors (see Appendix E). These behaviors addressed broad categories for stereotypic behavior such as stroke and tap. Any behavior that was both purposeless and repetitive was included. However, behaviors such as scratching and minute fingernail grooming were not included because they may possibly serve a purpose. Repetitive was defined as occurring in three consecutive cycles (Suomi, Harlow, & Kimball, 1971). I decided that the cycles must be within a half-second of each other in order to be considered in the same episode of behavior. Each repetitive behavior was categorized as well as where the behavior was targeted. For example, for tapping, it was indicated where the tap was directed (i.e. arm or desk) or what was being stroked (i.e. hair or arm). For leg and foot movements, it was indicated whether the entire leg was involved or just the foot and whether it was one or both legs or feet. Behaviors were scored using an all-occurrence sampling method (Altmann, 1974) with duration noted through frame numbers. These frames were ultimately converted to seconds.

The two paradigms that generated behavior were the Wait paradigm and the DNMS paradigm. For most analyses, the behaviors from these paradigms were analyzed separately. From the video coding, I calculated the total duration of stereotypic behaviors in seconds that these behaviors were performed either in the Wait or DNMS paradigms. I used this measure because I wanted to gain information from the perspective of a time budget for how long participants spent engaging in stereotypic behaviors during the two paradigms.

2.3.5.1 Types of Stereotypic Behaviors Performed

There were seven types of repetitive and purposeless behaviors observed and described in the ethogram: Tap, Stroke, Body Rock, Other, Object Manipulation, Bounce, and Swing. In order to assess the most common behaviors, I first calculated the total duration for the different types of behaviors performed across participants. I then accounted for how many participants performed these behaviors and divided the total duration by the number of participants who performed each category of behavior.

The most common behavior observed was tapping both with respect to many participants performing this behavior and spending the most amount of time performing this behavior (Figure 2.3). Notably, many participants also spent time in *other* stereotypic behaviors. Participants displayed many idiosyncratic behaviors that were repetitive and purposeless but could not be classified into one of the behavior categories. Examples include unusual hand motions or complicated sequences of behaviors that combine categories such as tapping and bouncing. These behaviors were counted as stereotypic behavior because they were repetitive, but were too different across participants in order to create specific and meaningful categories from them. Because tapping was by far the most common behavior, I combined these behavioral categories and used total duration scores to test my hypotheses.

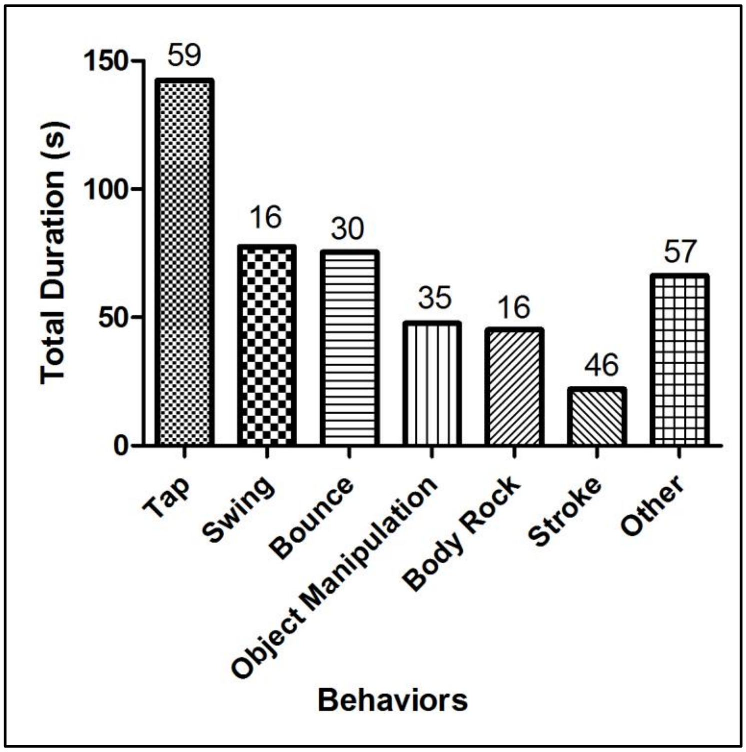

Figure 2.3 The prevalence (total duration divided by number of participants who performed the behaviors) of stereotypic behaviors observed both in the Wait and DNMS paradigms combined. The numbers above the bars represent the number of participants who engaged in these behaviors.

2.3.6 Questionnaire materials

2.3.6.1 Presentation to participants

After the two experiments, the participant was directed by instructions on the computer to proceed to a series of questionnaires. The order and purpose are presented in Table 2.4.

Table 2.4 The order and purpose of administered questions and scales to participants.

Prior to Experiment
Individual Differences
1. Ten-Item Personality Inventory (TIPI) (Gosling, Rentfrow, & Swann, 2003)
2. 5 questions from Shortened Intolerance of Uncertainty Scale (Carleton, Norton, & Asmundson, 2007)
Immediately following Wait & DNMS behavioral paradigms
Emotional State
1. Positive and Negative Affect Scale-Expanded (PANAS-X) (Watson & Clark, 1994)
2. Multidimensional State Boredom Scale (MSBS) (Fahlman et al., 2013)
Self-Report of Stereotypic Behavior tendencies
3. Fidgeting tendency scale (Mehrabian & Friedman, 1986) and follow-up questions about stereotypic behavior and severity
Characterization of Participants
4. Mental Health Check
Individual Differences
5. ADHD Current Symptoms Scale—Self -Report Form: Adults (Barkley & Murphy, 1998)
6. Behavior Inhibition/Behavioral Activation (BIS/BAS) Scale (Carver & White, 1994)
7. Shortened Intolerance of Uncertainty Scale (Carleton et al., 2007)
More Characterization of Participants
8. Demographic questions
Emotional State and Other Thoughts about Experiment
9. Free write: Perceived Purpose of study and Thoughts/Emotions

2.3.6.2 Generation of Test Variables from Questionnaires

In order to test how the participants' current emotional state and individual differences may have been related to stereotypic behavior performance, I used information from scales outlined in Table 2.4. Most of these scales had pre-established subscales. While exploratory factor analyses may reveal the true underlying latent variables in my population, my sample size (n=68) was most likely too small to adequately describe any such variables, as sample sizes of 100 or more are usually recommended for exploratory factor analyses (Budeav, 2010; MacCallum, Widaman, Zhang, & Hong, 1999). I thus worked with the pre-established sub-scales. The reliability of a potential factor was assessed with Cronbach's alpha, which measures the reliability with respect to the correlations between responses to questions that were designed the measure the same underlying latent variable. An alpha of 0.8 was considered a highly reliable factor.

2.3.6.3 Current Emotional State

2.3.6.3.1 Positive and Negative Affect Scale-Expanded (PANAS-X) (Watson and Clark, 1994)

The PANAS-X was designed to assess the participants' current general positive and negative affect as well as four negative emotions (fear, hostility, guilt, and sadness), three positive emotions (joviality, self-assurance, and attentiveness), and four complex affective states (shyness, fatigue, serenity, and surprise). Sixty emotion words were presented and the participants were asked to what extent they currently felt that emotion. Participants used a 5-point scale of *very slightly or not at all* through *extremely* to respond to each emotion. In addition to the 60 emotion words included in the scale, I added *bored*, *frustrated, pressed for time*, *anxious,* and *stressed* in order to specifically assess these emotions that may be associated with stereotypic behavior (α=.86). While the PANAS-X was presented first to all participants, I randomized the order of the emotion words in the PANAS-X when presented to each participant.

The first group of four factors described negative emotional states. The four factors were: General Negative Affect (α=.63), Fear (α=.40), Sadness (α =.77), and Guilt (α =.83). The next group of four factors for the PANAS-X described positive emotional states: General Positive Affect (α =.88), Joviality (α =.93), Self-assurance (α =.80) and Attentiveness (α =.85). The final group of pre-established factors for the PANAS-X did not describe necessarily positive or negative emotional states but instead complex emotional states. These factors were: Shyness (α =.69), Fatigue (α =.92), Serenity (α =.75), and Surprised (α =.75).

The five additional emotions of *Bored, Frustrated*, *Stressed*, *Pressed for Time*, and *Anxious* were added to the PANAS-X scale. I ran an exploratory factor analysis in order to see how an underlying latent variable may contribute to the response for these emotions. I ran these five items in a maximum likelihood factor analysis with Varimax rotation. In order to be included in a factor, each variable must have had an eigenvalue of at least 0.3. The resulting analysis of factor loadings created two factors: Bored/Frustrated (α =.52) and Pressed-for-Time/Anxious/Stressed (α =.57).

2.3.6.3.2 Multidimensional State Boredom Scale (MSBS) (Fahlman, et al. 2013)

While the PANAS-X assessed the current emotional state on a general basis, the MSBS more specifically probed how bored the participants felt during the test session.

The MSBS assessed aspects of arousal state, time perception, and how engaged participants felt with the current task. Participants used a 7-point Likert scale of *Strongly disagre*e through *Strongly agree* to indicate for 29 items how they currently felt while answering the questionnaire (α = .95).

The MSBS also had pre-established subscales, so I began my analysis by testing the reliability of these factors. Two of them had to do with arousal and were High Arousal (α =.82) and Low Arousal (α =.88). The three other factors were: Disengagement (α =.89), Time Perception (α =.93), and Inattention (α =.83).

2.3.6.3.3 Generation of Combined Boredom measure

Because both the PANAS-X and MSBS had factors that described Boredom as a potential emotional state, I tested whether the Bored/Frustrated measure from the PANAS-X correlated with the overall index from the MSBS scale. These measures were highly positively correlated ($r(68) = 0.459$, $p<0.0001$), thus suggesting that they may be capturing a similar emotional state. I created an additional factor for Boredom that combined the PANAS-X Bored/Frustrated and MSBS Index and named this factor *Boredom Combined* (α =.62).

2.3.6.4 Individual Differences

In order to test how differences between individuals with respect to tendencies or temperament may be related to the performance of stereotypic behavior, I used information from five scales. Two of the scales, the Ten-Item Personality Inventory (TIPI) (Gosling et al., 2003) (α= .68) and a modified Shortened Intolerance of Uncertainty Scale (Carleton et al., 2007) (α= .68), were from the Spring 2016 Prescreen questionnaire that the participants answered before being invited to participate in my study. I did not screen for certain answers form the questionnaires. Three of the scales were administered to participants during my study: the ADHD Current Symptoms Scale—Self-Report Form: Adults (Barkley & Murphy, 1998) (α= .88), Behavior Inhibition/Behavioral Activation (BIS/BAS) Scale (Carver & White, 1994) (α= .68), and the unmodified version of the Shortened Intolerance of Uncertainty Scale (Carleton et al., 2007) (α= .87).

2.3.6.4.1 Ten-Item Personality Inventory (TIPI) (Gosling, Rentfrow, & Swann, 2003)

The TIPI assessed personality traits of extraversion, agreeableness, conscientiousness, emotional stability, and openness to experiences. Participants read 10 phrases about these constructs and responded using a 7-point scale from *disagree strongly* through *agree strongly*. I included the TIPI in the Spring 2016 Prescreen Questionnaire in order to capture information about the participant without the experiment possibly affecting the response. In this light, personality should be a stable individual trait that can be assessed prior to the experiment and still be applicable when the participant is in the experiment.

For the TIPI scale, Pearson's correlations are reported because only two items went into each factor. The factors included Extraversion (Pearson's r=.46), Agreeableness (r= .39), Conscientious (r= .49), Emotional Stability (r= .62), and Openness to Experiences (r= .32).

2.3.6.4.2 Shortened Intolerance of Uncertainty Scale (Carleton, Norton, & Asmundson, 2007)

Five questions from the scale were included in the Spring 2016 Prescreen Questionnaire (see Appendix A for the included questions) and then the full length of the scale was administered during the experiment. Participants responded to statements about their perception of uncertain events and how they feel about uncertainty in general by responding to a 5-point scale of *Not at all characteristic of me* through *Entirely characteristic of me*.

For the Shortened Intolerance of Uncertainty Scale, I first created a factor from the 5 items that were included in the prescreen questionnaire (α= .72). For the full scale administered during the experiment, I created factors from the recommended subscales of Prospective Anxiety (α= .82) and Inhibitory Anxiety (α= .89). I then tested whether the prescreen responses were related to the responses to the same questions asked during my study. The responses were highly positively correlated ($r(67) = 0.569$, $p < 0.0001$) via a Pearson Correlation, which suggests that these responses are capturing a similar perception about uncertainty across time and contexts. I thus created a factor called Intolerance of Uncertainty Index that consisted of the overall index score from the Intolerance of Uncertainty Scale administered during the study and the 5 responses during the prescreen questionnaire (α= .89). I called this factor *Intolerance of Uncertainty Combined*.

2.3.6.4.3 ADHD Current Symptoms Scale—Self-Report Form: Adults (Barkley and Murphy, 1998)

I used the ADHD Current Symptoms Scale to assess participants' levels of hyperactivity and general difficulty on focusing on tasks. Even if symptoms were not severe enough to warrant an ADHD diagnosis, participants who tended to either be hyperactive or tend to have difficulty focusing on tasks may be more likely to perform stereotypic and associated fidgeting behaviors during my experiment. Participants answered 18 questions about distractibility, their ability to complete work, and ability to maintain attention and focus. They answered with a 4-point scale from *never or rarely* through *very often* (α= .88).

I kept the ADHD questionnaire as an overall Index (α= .89). As opposed to the other scales used in this study, the ADHD Current Symptoms Scale—Self-Report Form: Adults (Barkley and Murphy, 1998) is a questionnaire meant to represent an accumulation of symptoms associated with ADHD for diagnostic purposes. There were no recommended subscales, and when I included all of the items in a maximum likelihood factor analysis with Varimax rotation, no factors emerged that would more effectively describe a latent variable better than the overall Index.

2.3.6.4.4 Behavior Inhibition/Behavioral Activation (BIS/BAS) Scale (Carver and White, 1994)

The BIS/BAS scale measured sensitivity to impending reward. Participants who scored high on behavioral activation (BAS) were considered individuals who pursue movement towards goals and feel positive emotions such as happiness when presented with a signal of an upcoming reward. On the other hand, participants who scored high on behavioral inhibition (BIS) were considered individuals who avoid movement towards goals and feel negative emotions such as fear or anxiety in response to cues of upcoming punishment. These measures of BAS and BIS also are considered well-correlated with personality measures of neuroticism for BIS and measures of extroversion for BAS.

Participants responded to 24 statements about themselves using a 4-point scale of *Very true for me* through *Very false for me* (α=.68). I placed participants in an experiment in which they were presented with cues about the timing of an anticipated event, so the BIS/BAS scale was a relevant measure to assess how sensitive participants were to the signal of an upcoming event.

The BIS/BAS scale had a number of recommended subscales. The first one was Behavioral Activation: Drive (α= .73). This factor contained all of the recommended items with the exception of the statement *When I go after something I use a "no holds barred" approach* which was eliminated in order to improve the reliability of the factor from α= .73 to .77. Feedback from pilot participants suggested that not all participants may have understood the meaning of the phrase “no holds barred.” The other factors included Behavioral Activation: Fun-seeking (α= .57), Behavioral Activation: Reward Responsiveness (α= .70), and Behavioral Inhibition (α= .75).

2.3.6.5 Other Factors from Questionnaires

2.3.6.5.1 Mental Health check

Other variables could have potentially affected stereotypic behavior in my study. For example, amount of sleep the night before, caffeine intake, and mental health diagnoses could affect fidgeting and stereotypic behavior. I asked participants about how much sleep they received and caffeine they consumed within the last 24 hours. I also asked participants whether they have been diagnosed with attention deficit/hyperactivity disorder (ADHD), Tic disorder, ASD, and Stereotypic Movement Disorder with *yes* or *no* questions. These disorders are known to be associated with an increase in behaviors such as increased fidgeting (ADHD), tics (Tic disorder), and stereotypies (ASD, Stereotypic Movement Disorder) (DSM-V). I also asked whether participants have taken medicine typically prescribed for ADHD such as Adderall.

These questions were included primarily as a way to screen for participants who possibly should not be included in the data analysis because of certain clinical diagnoses or an extreme consumption of caffeine or extreme lack of sleep. Sensitivity analyses of potential participant issues are included in Appendix C.

2.3.6.5.2 Self-Report of Stereotypic Behavior

This section consisted of two parts: the first was the fidgeting tendency scale from Mehrabian and Friedman (1986) and the second part was questions about the participants' perceived performance of these behaviors during the experiment and perception of their severity. Because the fidgeting tendency scale was more aimed at fidgeting behavior, I included some behaviors that are regarded as stereotypic but were not already included in the scale. These behaviors included: *pull or twist my hair*, *touch my face*, *rock my body or torso back and forth or side-to-side when seated*, *pace a lot while waiting for something to happen*, *bite my nails*, *tap my fingers*, *crack my knuckles or fingers*, *click or grind my teeth* and *shake my leg*. Participants used a 9-point Likert scale of *Very strongly agree* through *Very strongly disagree* to indicate how often they performed each fidgeting behavior in the past six months ($\alpha = .89$). I randomized the order of the behaviors presented to each participant.

Following the fidgeting tendency scale (Mehrabian & Friedman, 1986), I also directly asked whether participants recalled performing any of the described behaviors during the experiment as well as what they considered their most common habit. Finally, I asked questions about the participant's perceived severity of the behavior in order to assess the level of disturbance participants feel about their own stereotypic behaviors. Two of these questions came from the survey of purposeless habits created by Hansen, et al. (1990).

The fidgeting tendency scale served as a way for participants to self-report how often they perform stereotypies and other fidgeting behaviors as well as what kind of behaviors they tend to perform the most. I used the fidgeting tendency scale from Mehrabian and Friedman (1986) in order to test how self-aware participants were about their performance of stereotypic behavior (analyses in Appendix B). There is evidence that fidgeting behavior occurs outside of a human's attentional awareness when the mind wanders (Carriere, Seli, & Smilek, 2013).

2.3.7 Statistical Analyses

2.3.7.1 The effects of predictability on stereotypic behavior

The first step in my analyses was to assess the effects of predictability on behavior prior to incorporating data from any of the scales. To test the effects of predictability in the Wait paradigm, I used an independent samples t-test for the two conditions of Predictable and Unpredictable. I also used these tests when analyzing the DNMS paradigm with respect to the two conditions of Predictable and Unpredictable. When I factored in the conditions of the Wait experiment on the DNMS paradigm, thus creating 4 conditions of Predictable with clock, Predictable without clock, Unpredictable with clock, and Unpredictable without clock, I used a one-way Analysis of Variance (ANOVA) test. I tested the assumption of ANOVAs of homogeneity of variance using a

Levene’s test. If the assumption was violated, then I log-transformed data in order to better meet the homogeneity of variance assumption.

2.3.7.2 The effects of emotional state and individual differences on stereotypic behavior

Once the emotion state and individual difference variables were created, I tested the relationship between these variables and Wait and DNMS behaviors with Pearson correlations. I used the total duration measure for both Wait and DNMS behaviors. It was evident from my preliminary analyses that I could collapse the stereotypic behavior across conditions.

2.3.7.3 Mediation analyses of the effects of emotional state and individual differences on stereotypic behavior

The next step in my analyses was to test what individual factors may contribute to the relationship between boredom and behavior. I used the PROCESS macro (Hayes, 2013) to test simple mediation models. It was clear at this point in my analyses that even though there was variability in stereotypic behavior observed in the paradigms, conditions in either of Wait or DNMS paradigms were not affecting stereotypic behavior. For the mediation analyses, I thus tested whether there were individual differences that accounted for differences in stereotypic behavior as mediated by reported current emotions.

2.4 Results

2.4.1 Characteristics of stereotypic behavior performance

Both the Wait and DNMS paradigms successfully elicited stereotypic behaviors from participants. In the Wait paradigm, 57/68 (83.85%) of the participants performed stereotypic behavior, and this number increased to 67/68 (98.5%) participants in the DNMS paradigm. The Wait paradigm was 5 minutes long and the DNMS paradigm was approximately 20 minutes long but the duration depended partially on participant response time. In the Wait paradigm, participants spent about 51 seconds, on average, performing these behaviors and on average 3.5 minutes in the DNMS paradigm (Table 2.5). To better compare time spent in stereotypies between paradigms, the total duration spent in stereotypies in the DNMS paradigm was divided by four to account for the 20-minute paradigm time. Corrected for time, the participants spent on average 53.34 (±48.8) seconds in the DNMS performing stereotypies. This is a similar time spent performing these behaviors during the Wait paradigm. Yet, it is also evident from the deviations and ranges presented that there is a large amount of individual variation in the performance of these behaviors.

Table 2.5 Mean and Range for Stereotypic behavior measures in the Wait and DNMS Paradigms

	Wait Paradigm	DNMS Paradigm
Mean total duration (±SD) (seconds)	51.21± 68	213.40±195.24
Total Duration Range: (with minimum above zero) (seconds)	3.3-331.96	1.9-755.80
Mean Average Duration (±SD) (seconds)	6.71±7.28	9.14±5.46
Average Duration Range: (with minimum above zero) (seconds)	2.18-41.45	1.9-29.36

2.4.2 Correlations between paradigms

Behaviors were scored during delays in two slightly different scenarios for the Wait and DNMS paradigms. The delay in the Wait paradigm consisted of the participants sitting at a desk in front of an off computer screen for 5 uninterrupted minutes while waiting for the experimenter to return. In the DNMS paradigm, the participants were engaged in a computer task for approximately 20 minutes. Within these 20 minutes, there were 7 delays ranging from 1-4 minutes during the paradigm that were interrupted by participant actions on the computer.

Despite that each delay within the DNMS paradigm was shorter than the 5-minute Wait paradigm, there was individual consistency in stereotypic behavior between paradigms. Behaviors in the two paradigms were highly positively correlation between the Wait and DNMS paradigms (r(68)=0.582, *p*<0.0001). Participants who performed more behaviors in the Wait paradigm also performed more behaviors in the DNMS paradigm.

2.4.3 The effects of predictability on Stereotypic Behavior

In the Wait paradigm, there were no significant differences observed between the Predictable and Unpredictable conditions (t(66)=0.7644, *p*= 0.447) (Figure 2.4). In the DNMS paradigm, there were no significant differences observed when analyzed as two conditions, Predictable and Unpredictable conditions (t(66)= (0.462, *p*= 0.645) (Figure 2.5a). There were also no significant differences observed when the clock from the Wait condition was accounted for, thus creating four DNMS conditions: Predictable with clock, Predictable without clock, Unpredictable with clock, and Unpredictable without clock (F(3)= 0.169, *p*= 0.917) (Figure 2.5b).

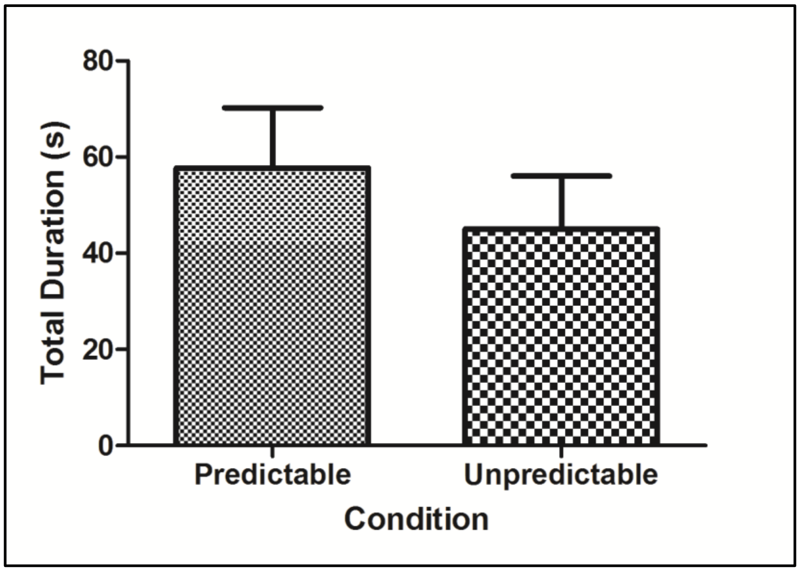

Figure 2.4 Total duration of stereotypic behavior observed in the Wait paradigm, by condition. Bars represent ±1 SEM.

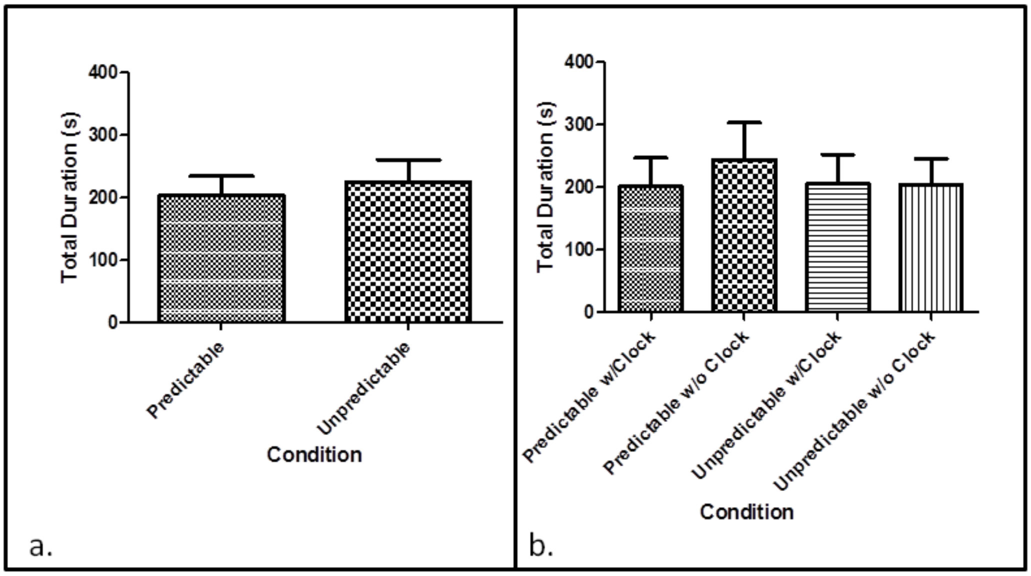

Figure 2.5 Total duration of stereotypic behavior observed in the DNMS paradigm, by the two DNMS conditions (a), and with the conditions of the Wait paradigm included, thus creating four conditions (b). Bars represent ±1 SEM.

2.4.4 Correlations between stereotypic behavior, emotional state, and individual differences

Preliminary analyses indicated that my composite measure of boredom (Boredom Combined; (α =.62) and individual difference of Behavioral Inhibition (α =.75), the ADHD Index (α=.88), Intolerance of Uncertainty combined (both prescreen and experiment questionnaires) (α=0.89) and the Emotional Stability measure (α=.62) from the TIPI were most important to examine as predictors of stereotypic behavior.

Table 2.6 demonstrates that there were significant correlations between stereotypic behavior during the Wait paradigm and behaviors displayed in the DNMS paradigm, as discussed previously. In addition, there was a significant positive relationship between the stereotypies performed in the Wait paradigm and the measure of boredom (Table 2.6).

Table 2.6 Pearson correlation coefficient (r) for Stereotypic behavior, Emotion, and Individual Difference measures. Bolded values represent $p<0.05$. *$p<0.05$, **$p<0.01$

	Wait (s)	DNMS (s)	Bored Combined	IUS Combined	BIS Index	ADHD Index	TIPI EmotStab
Wait (s)	1						
DNMS (s)	**.582****	1					
Bored Combined	**.271***	.062	1				
IUS Combined	.074	.063	**.453****	1			
BIS Index	-.093	-.066	**.437****	**.343****	1		
ADHD Index	.097	.086	**.394****	**.328****	**.245***	1	
TIPI EmotStab	-.004	-.027	**-.383****	**-.479****	**-.622****	**-.331****	1

As for the individual differences, the boredom measure had a significant relationship with the Intolerance of Uncertainty combined, the BIS Index, the ADHD Index, and the TIPI Emotional Stability measure (Table 2.6). These correlations were positive with the exception of the TIPI Emotional Stability measure which had a significant negative relationship with boredom and the other individual difference measures.

2.4.5 Mediation analyses on individual differences and emotion on stereotypic behavior

2.4.5.1 The effect of Behavioral Inhibition and Boredom on stereotypic behavior

Differences in self-report as reflected in the Behavioral Inhibition Scale factor were found to have significant effects on stereotypic behavior in the Wait paradigm through increasing boredom (*b*=0.80, 95% CI [0.39, 1.20]) and increasing stereotypic behavior (*b*=29.09, 95% CI [9.67, 48.52]). The indirect effects of boredom through the BIS index on stereotypic behavior was significant with the *Boredom Combined* factor acting as a significant mediator (*b*=23.13, 95% CI [8.04, 46.30]). After controlling for these indirect effects, the direct effect of Behavioral Inhibition on stereotypic behavior was also significant (*b*=-35.92, 95% CI [-71.28, -0.56]) and had a negative relationship with stereotypic behavior whereas the indirect effects had a positive relationship with stereotypic behavior.

The direct and indirect effects on stereotypic behavior in the DNMS paradigm were not as robust. Behavioral Inhibition had a significant effect on increasing boredom (*b*=0.80, 95% CI [0.39, 1.20]), but there was no significant effects of boredom on stereotypic behavior (*b*=24.28, 95% CI [-34.95, 83.51]) and boredom was thus not a significant mediator (*b*=19.30, 95% CI [-25.33, 66.71]). Additionally, there was no significant direct effects of Behavioral Inhibition on stereotypic behavior in the DNMS paradigm (*b*=-45.46, 95% CI [-153.28, 62.37]).

2.4.5.2 The effects of Intolerance of Uncertainty and Boredom on stereotypic behavior

I next analyzed intolerance of uncertainty using my combined factor from the pre-experiment Prescreen and the full Shortened Intolerance of Uncertainty scale administered during the experimental session. There were significant effects on stereotypic behavior in the Wait paradigm through increasing boredom via the Boredom Combined factor (*b*=0.62, 95% CI [0.32, 0.93]) and increasing stereotypic behavior (*b*=22.56, 95% CI [2.38, 42.74]). The indirect effects of boredom through the Intolerance of Uncertainty measure on stereotypic behavior was significant with the *Boredom Combined* factor acting as a significant mediator (*b*=14.10, 95% CI [2.49, 31.03]). After controlling for these indirect effects, there was not a significant direct effect of Intolerance of Uncertainty on stereotypic behavior (*b*=-6.36, 95% CI -34.20, 21.48]).

As with the Behavioral Inhibition mediation, the direct and indirect effects of Intolerance of Uncertainty and Boredom on stereotypic behavior in the DNMS paradigm were not as strong as observed for the Wait paradigm. Intolerance of Uncertainty had a significant effect on increasing boredom (b=0.62, 95% CI [0.32, 0.93]), but there was no significant effects of boredom on stereotypic behavior (b=9.11, 95% CI [-50.94, 69.15]) and boredom was also not a significant mediator (b=5.69, 95% CI [-69.84, 95.85]). There was also no significant direct effects of Intolerance of Uncertainty on stereotypic behavior in the DNMS paradigm (b=13.01, 95% CI [-69.84, 95.85]).

2.5 Discussion

2.5.1 The effects of predictability on stereotypic behavior

The Wait and DNMS paradigms successfully elicited stereotypic behavior. However, there were no differences in behavior observed between predictable and unpredictable conditions in either paradigm, thus failing to support my predictions that there would be more behavior observed in the Unpredictable condition (Figures 2.4, 2.5). Previous studies have demonstrated that when people have to wait to for an anticipated event, they seek out information about waiting time (Pamies et al., 2016). Customers also prefer waits with duration information (Hui & Zhou, 1996) or their progressing position in a line (Munichor & Rafeli, 2007), and are more satisfied with public transportation experiences that provide wait information (Brakewood et al., 2015; Watkins et al., 2011). These preferred conditions would be akin to my Predictable condition in which participants knew when the experimenter would return in the Wait paradigm and countdown clocks indicated when the delay would end in the DNMS paradigm.

I propose two reasons why my results did not align with these previous studies. First of all, these prior studies did not assess behavior during waiting scenarios. It is possible that while people prefer predictable conditions, this preference does not alter behavior when presented with either more or less predictable conditions. This idea is not entirely supported by the few studies on human stereotypic behavior, for if the environment induces boredom or anxiety, stereotypic behaviors have been different depending on condition (Barash, 1974; Woods & Miltenberger, 1996; Lang et al., 2015). However, these studies did not test predictability. It is possible that while people prefer predictable conditions, predictability is not enough of a salient factor in the environment to shift people's behavioral response to the environment.

Secondly, it is possible that stereotypic behavior and predictability are related but my study scenario was too artificial to demonstrate this relationship. I was ethically obligated to inform the participants prior to the commencement of the study that the study would transpire for no more than 1 hour of their time. Because participants knew that ultimately they would be in the study for 1 hour or less, this knowledge may have abolished any response to the predictability of the wait or task delay scenarios.

Additionally, the artificiality may not have elicited the relationship between stereotypic behavior and predictability because the anticipated events in the Wait and DNMS paradigms were not overly positive or negative in nature. Participants received raffle tickets as a reward in the study. However, they did not physically receive any reward until after the study ended. During the study, they received no positive reinforcement except for the visual and auditory feedback for correct answers.

On the other hand, no event was particularly anxiety-inducing as the participants waited to either start the study in the Wait paradigm or answer the next question in the DNMS task. Previous studies that elicited stereotypic behaviors in challenging conditions used the scenario that participants had to prepare a presentation to present in front of other people on a topic that they were just provided (Woods & Miltenberger, 1996; Lang et al., 2015). This scenario borrows from but is not the full protocol for a procedure called the Trier Social Stress Test (Kirschbaum, Pirke, & Hellhammer, 1993). The Trier Social Stress Test is a laboratory scenario in which participants provide a 5-minute speech in front of managers on why they should be hired for a job and then are spontaneously told during the public presentation to perform a complex mathematical operation as fast and as accurately as possible. Saliva samples collected during the experiment consistently demonstrate that this protocol effectively shifts biological indicators of acute psychological stress such as cortisol associated with hypothalamus-pituitary-adrenal (HPA) axis activity (Kirschbaum et al., 1993; Foley & Kirschbaum, 2010).

In the artificial laboratory scenario, manipulations in social stress of a certain magnitude might be needed in order to overcome the artificial nature of the environment and to shift physiology and behavior. In my study, the only social components were the waiting for the arrival of the experimenter as well as the cameras in the room. Without sufficient social stress, my laboratory scenario in the Unpredictable conditions of both the Wait and DNMS paradigms may not have been challenging enough in order to shift behavior between participants.

Despite the finding that the environmental manipulations failed to shift behavior, participants who reported higher levels of boredom exhibited more stereotypic behaviors. These results align with the hypotheses that stereotypic behavior may be elicited to serve a psychological function in the current environment (Mason & Latham, 2004). While the usual perception of boredom would suggest that the environment was under-stimulating, boredom as an emotional state is poorly understood. Eastwood, Frischen, Fenske, and Smilek, (2012) proposed that while boredom is usually associated with low arousal, from a perspective of engaging attention, both low and high stimulation can negatively affect attention and create a state of boredom. This high stimulation bored state might contain restlessness and irritability (Eastwood et al., 2012). While I cannot conclude what levels of stimulation may have contributed to boredom in my participants, stereotypic behavior may have served a psychological function to compensate for a misalignment between homeostatic and environmental levels of stimulation.

Most studies of boredom to date have focused on how to reduce boredom in students and workers in order to increase learning and productivity (Loukidou, Loan-Clarke, and Daniels, 2009; Vogel-Walcutt, Fiorella, Carper, and Schatz, 2012). Recently, psychologists have focused on studying it as a psychological phenomenon because proneness to boredom has been linked to mental health issues such as depression, anxiety, and recovery from traumatic brain injuries as well as impulse control deficits related to gambling and drug addictions (Eastwood et al., 2012). Yet, these studies rarely focus on behavioral components of boredom.

To my knowledge, no previous study has established a link on the individual level between boredom and the performance of stereotypic behavior. In Woods and Miltenberger's (1996) study, participants in the *bored* condition performed more object-related behaviors than when placed in the *neutral* or *anxious* conditions. However, participants were only asked "*How bored were you during this situation*?" and it is not reported how an individual's answers related to his or her behavior other than that participants overall felt bored in the *bored* condition.

The cause of boredom may not only be related to the environment as discussed above, but people may have different propensities for getting bored. Analyses suggest that both of these components, environmental and person-based, contribute to the tendency to feel bored and can interact with each other (Mercer-Lynn, Bar, and

Eastwood, 2014). The results of my mediation analyses can shed light on how individual propensities for feeling bored may contribute to the performance of stereotypic behavior.

As for proneness to boredom in my study, there was an indirect effect of Behavioral Inhibition (BIS), and Intolerance of Uncertainty on stereotypic behaviors in the Wait paradigm through feelings of boredom. These measures both pertain to inhibitory responses to the environment. Participants who tend to have these inhibitory responses to unfavorable conditions may have found the Wait paradigm an aversive experience because the stimulation in the room was either under or over the homeostatic levels of the participant as suggested by the report of feeling bored. These results suggest that these may have used stereotypic behavior to serve a psychological function of regulating their stimulation during the Wait paradigm.

The BIS/BAS scale (Carver & White, 1994) is based on the Reinforcement Sensitivity Theory of personality (Corr, 2004) in which the scale measures sensitivity to impending reward. Specifically, the Behavioral Activation Scales (BAS) were intended to capture approach motivation and the Behavioral Inhibition Scale (BIS) was designed to assess avoidance motivation (Carver & White, 1994). My results demonstrated that participants who rated high on the BIS were more inclined to feel bored and perform stereotypies. Yet, the direct effect between the BIS factor and stereotypic behavior was negative when boredom was controlled for.

The effect of BIS levels predicting boredom via the MSBS was found by Mercer-Lynn, Flora, Fahlman, and Eastwood (2011) and then replicated by Mercer-Lynn et al., (2014). In both these studies and my results, boredom and the BIS had a positive relationship. Yet, in Mercer-Lynn et al. (2014), the BIS was not predictive of MSBS scores when participants were in non-boring conditions. This may account for the negative relationship in my meditation analysis between BIS and stereotypic behaviors when boredom is controlled for.

The BIS pertains to individual tendencies to withdraw from aversive situations. Results suggest that people with these tendencies were more prone to feeling bored in a boring situation. Eastwood et al., (2012) hypothesized that tendencies reflected in the BIS/BAS represent a chronic hyper (BIS) or hypo (BAS) sensitivity to stimulation and may be a psychological cause of boredom. Additionally, people placed in boring situations report attempts to cope with boredom either through behavioral or cognitive means (Nett, Goetz, and Daniels, 2010). In my study, it is possible that people who scored high on the BIS scale were more sensitive to feelings of boredom and sought to avoid these feelings through increased expression of stereotypies.

The Intolerance of Uncertainty scale is associated with state anxiety and other anxiety disorders (Carleton et al., 2007). In the Wait paradigm, there was a positive relationship between intolerance of uncertainty and stereotypic behaviors through a mediation with feelings of boredom. With respect to how an individual’s intolerance of uncertainty affects waiting periods, Sweeny and Andrews (2014) measured responses to a long-term waiting of receiving test scores 4 months later. They found that people who scored high on Intolerance of Uncertainty engaged in more emotional regulation during the waiting period (Sweeny & Andrews, 2014). A similar effect may have occurred during my study. Participants who were less tolerant of uncertainty might have found the wait to be a more challenging condition than those who were more tolerant of uncertainty and may have used stereotypic behaviors to regulate themselves while waiting.

2.5.2 Characterizations of Human Stereotypic Behavior

Beyond the analysis of other factors, I also characterized behaviors across time and different contexts. For the participants in my study, stereotypies observed across the Wait and DNMS paradigms were highly positively correlated (Table 2.3). The Wait paradigm consisted of an uninterrupted waiting period and the DNMS paradigm consisted of engaging in a task. While the context was largely the same as the participants encountered both the Wait and the DNMS paradigms within the same room in the same hour, these results suggest that stereotypic behavior does not differ between slight changes in context.

In this light, the Wait and DNMS paradigms both successfully produced stereotypic behavior in participants. These behaviors were similar across the two paradigms and while the DNMS paradigm had multiple but shorter delays than in the Wait paradigm, there were more behaviors displayed in the longer DNMS paradigm than in the short Wait paradigm. It is possible, depending on the question being asked, that the Wait paradigm would be sufficient to understand how the environment and how individual differences in emotion and temperament affect stereotypic behavior.

Finally, a limitation of my analyses is that my participant pool represented a narrow sample of the human population. I could not answer any questions about age differences in behavior, and sex differences were not present (see Appendix C) but males were not well-represented. It is possible that while the Wait and DNMS paradigms elicited stereotypic behavior in my participants, these same paradigms may produce different results in other swaths of the general population.

2.5.3 Conclusion

There has been little research on the topic of stereotypic behavior in nonclinical adult humans. My results suggest that this is not because people do not perform them, as both of my experimental paradigms elicited a great amount and variety of stereotypic behavior. However, with respect to my hypotheses and predictions, these behaviors were not affected by the predictability of the delays in either the Wait or DNMS paradigms.

Yet, while not elicited by experimental manipulations of predictability, participants who reported feeling the emotional state of boredom performed more stereotypic behaviors. The feelings of boredom were predicted by individual temperament measures that may have been affected by the environment although not by variables of predictability. The relationship between boredom and stereotypies may align with my overall hypotheses that stereotypies may serve as a compensatory response in certain environmental conditions.

CHAPTER 3

THE EFFECTS OF PREDICTABILITY ON BEHAVIOR IN ADULT RHESUS MACAQUES

3.1 Overview

The objective of Experiments 2 and 3 with rhesus macaques was to assess the behavioral response to predictability in anticipated events. In contrast to other studies that have manipulated predictability in captive animals, my studies did not manipulate husbandry routines such as feeding time. Because these monkeys are routinely tested in cognitive paradigms, their expectations of such events were manipulated. These experiments were designed not only to replicate other scenarios that the macaques encounter, but also to mirror the design of Experiment 1 in humans as closely as possible. As in the human study, the response of interest was stereotypic behavior as compared with the stereotypic behaviors generated by humans in Experiment 1.

Experiments 2 and 3 with rhesus macaques consisted of two similar paradigms as conducted in Experiment 1 with humans. However, whereas each human participant received both the Wait and DNMS paradigms within a single testing session, the two paradigms were conducted with macaques in separate sessions on separate days. The experiment of the modified DNMS task, called a delayed response task (DRT), was carried out first. When all of the sessions for the DRT were completed, I then used the macaques’ general anticipation for participating in the delayed response task and ran the Wait Experiment. Following reports of Experiment 2 and 3, I compared the macaque responses between their two experiments and then compared responses between macaques and the human participants from Experiment 1.

3.2 Introduction

3.2.1 Rhesus Macaque Natural History

Rhesus macaques are an Old World monkey and one of the most common species of primates used in biomedical research because of their close evolutionary relationship with humans, resulting in similar anatomy, physiology, and behavior. They shared their last common ancestor with humans around 25 million years ago (Kumar & Hedges, 1998). Rhesus macaques are found throughout southern Asia and even thrive in human-modified environments from agricultural regions to cities and have thus been called weed macaques for their ability to persist in a variety of different habitats (Richard, Goldstein, & Dewar, 1989). Concordantly, they are considered a species of Least Concern by the International Union for Conservation of Nature and Natural Resources (IUCN: Timmins, Richardson, Chhangani, & Yongcheng, 2008).

Rhesus macaques are generally diurnal, omnivorous, and reside both on the ground and in trees or man-made structures. They live in large social groups consisting of multiple males and females. The backbones of the groups are multigenerational matrilines, for females remain in the social group in which they were born, whereas around puberty, males emigrate and attempt to integrate into a new social group (Melnick, Pearl, & Richard, 1984). Within a social group, macaques use dominance status and rank in order to regulate access to resources. Matrilines within a social group are ranked from high to low, and within matrilines, daughters rank below their mothers in reverse age order, so a younger daughter outranks an older one (Silk, 2009). Affiliative behaviors such as grooming are used to reinforce social bonds whereas aggression ranging from facial threats through physical contact is used to reaffirm rank differences (Southwick, 1967).

Depending on the housing arrangements in captivity, many of these social relationships and behaviors are observed in captive macaques. When housed in large social groups, macaques establish matrilines. These matrilines compete for resources and can even engage in group violence when the social group becomes unstable (Dettmer, Woodward, & Suomi, 2015). Most macaques that are used in biomedical research are housed indoors in pairs or without direct physical contact in cases of incompatibility. While matriline formation may not be possible, macaques appear to form ranks and relationships with both the macaques they share a room with as well as the human caretakers they interact with (Asakura, 1958). These macaques not only use similar social signals as wild macaques in terms of affiliative or aggressive behavior, but can also eavesdrop and interpret the ranks of human experimenters who specifically use macaque-like facial expressions while interacting with each other in an experimental paradigm (Hamel, unpublished data).

3.2.2 Types of Stereotypic Behavior

It is largely unstudied whether stereotypic behaviors occur in wild animals, both in general and for rhesus macaques specifically. For example, Mason and Latham's (2004) previously reported estimation of the prevalence of stereotypies across species only included animals in captivity. The subsequent discussion about stereotypic behaviors will thus pertain to captive rhesus macaques.

There is no standardized categorization of stereotypic behaviors. Yet, there are consistencies across ethograms utilized for different populations in different periods of time (see ethograms in Lutz, et al., 2003 and Gottlieb, et al. 2013a). First of all, stereotypies in rhesus macaques can be categorized either as whole-body or motor stereotypies such as pacing, rocking, and swinging, or self-directed, such as hair-pulling and digit-sucking (Lutz, et al. 2003). Secondly, motor stereotypies are usually required to be repetitive actions in order to be considered a stereotypy with three iterations usually considered the minimum number of required repetitions (Lutz, et al. 2003; Gottlieb, et al., 2013a). Self-directed behaviors do not have to be repetitive in order to be considered a stereotypy. It is currently unknown whether there is a different function for motor or self-directed stereotypies or whether different scenarios elicit these types of stereotypies.

In this dissertation, stereotypies will refer to both motor and self-directed stereotypies unless I distinguish further. However, the studies described below may discuss stereotypies in general but not actually include both self-directed and motor behaviors in their observations and analyses. As there is little information on any distinctions between the cause of and function of these two types of stereotypies, I will consider that findings related to one kind of stereotypic behavior can possibly apply to both types of stereotypies.

3.2.3 Prevalence of Stereotypic Behavior

For rhesus macaques housed in laboratories, the prevalence of stereotypies in the population can range from 18.4% to 78% (Lutz, Well, & Novak, 2003; Lutz, Coleman, Maier, & McCowan, 2011) depending on the sex, age, and developmental history of the monkeys as well as the management practices at the facility. These percentages reflect the presence of stereotypic behavior and represent animals at all points in the severity range. It is estimated that fewer than 10% of the rhesus macaques housed in laboratories exhibiting stereotypic behavior display it at levels that interfere with basic biological processes (M. Novak, personal communication), thus demonstrating the importance of understanding the potential function of performing mild levels of stereotypic behavior.

Without a standardized severity scale, it is difficult to systematically assess the severity of stereotypies. However, two metrics can be used to identify monkeys with severe stereotypic behavior: the time spent in stereotypic behavior across the day (base rate) and the level of voluntary participation in cognitive studies, not involving any food deprivation. Monkeys classified as severe score high on the first dimension and low on the second dimension. The macaques studied in my dissertation display mild stereotypic behavior inasmuch their base rates were low and their participation in cognitive tasks was high.

3.2.4 Hypothesized Functions of Stereotypic Behavior

A key feature of stereotypies observed in rhesus macaques is that they do not appear to serve a function such as acquiring resources or moving towards or away from a clear stimulus. These behaviors are thus hypothesized to serve a psychological purpose. As discussed for both human and nonhuman primates in the introduction, there are four leading hypotheses for why animals perform these behaviors (Mason & Latham, 2004). The first two are considered divergent responses to the current environment. One of these hypotheses is that an animal may perform stereotypies in order to increase stimulation in response to an under-stimulating environment, whereas another hypothesis is that an animal may perform stereotypies in order to cope with a stressful or otherwise challenging environment (Mason & Latham, 2004).

However, two additional hypotheses do not pertain to the current environment and may possibly affect behavior in my experiments. The third hypothesis is that stereotypies reflect a previously developed habit rather than a response to the current environment, although it may reflect previous exposure to an under or over-stimulating environment (Mason & Latham, 2004). The current environment would thus not affect expression of stereotypic behavior if it is a habit rather than a response to the environment. Finally, stereotypies may have an underlying physiological, rather than environmental, cause. Psychostimulants such as cocaine and methamphetamines and genetic disorders such as Fragile X syndrome are known to increase stereotypic behavior. The monkeys in this study and in the studies reviewed below have not had exposure to the known drugs that cause stereotypic behavior and are not known to have genetic disorders linked to these behaviors. However, the possibility remains that there are underlying physiological differences, currently not known, that affect expression of their stereotypic behavior.

3.2.5 Factors that Affect Stereotypic Behavior

3.2.5.1 Individual Factors

Whether certain traits of individuals confer differential risk for developing stereotypies has been studied in rhesus macaques, although sometimes with unclear or conflicting results. First of all, male macaques were more likely to display stereotypies than females (Lutz et al., 2003; Vandeleest et al., 2011; Gottlieb et al., 2013a). This is a stable result; however, no consistent explanation has been put forth for why this sex difference occurs when studied.

Researchers also assessed whether there were certain temperaments and tendencies in rhesus macaques that were associated with increased risk of developing stereotypic behavior. Infants involved in the BioBehavioral Assessment (BBA) program at the California National Primate Research Center were rated for temperament after a 25-hour test session of behavioral observation and challenges such as a novel object and an unfamiliar human intruder (described further in Golub, Hogrefe, Capitanio, & Widaman, 2009). The possible temperaments that characterized infants based on factor analyses were *vigilant*, *gentle*, *confident*, and *nervous*.

When monkeys face environmental stressors, these temperaments may differentially affect the development of stereotypic behaviors. Two studies using the BBA provided partial confirmation of this idea, but it depended on environment. A relationship between temperament characteristics and stereotypic behavior was present only for indoor housed mother-infant pairs but not for infants housed with their mothers in large, species-typical outdoor groups (Vandeleest et al., 2011; Gottlieb et al., 2013a).

However, for monkeys housed indoors, relationships between temperament and stereotypic behavior were found but were mixed. Vandeleest et al. (2011) noted that nervous and gentle macaques were more at risk for developing stereotypic behaviors, yet Gottlieb et al. (2013a) found, in the same facility as Vandeleest et al.'s (2011) study, that only macaques that were *not* gentle were more at risk for developing stereotypic behaviors. Gottlieb et al. (2013a) did not acknowledge or explain the discrepancy in these results. While none are offered by the authors, there are some possible explanations for these contradictory results. First, there may be different genetic predispositions in the two populations studied that led to these contrasting results. Second, there were age differences between the populations studied, for the age range in Vandeleest et al.'s (2011) study was 0.5-4.7 years and in Gottlieb et al.'s study (2013a) was 1-10 years. It is possible that differences in age contributed to different results of how infant temperament predicts future stereotypic behavior. Finally, it may be that the *gentle* temperament or the temperaments in general as generated from the BBA may not be reliable predictors of future performance of stereotypic behavior.

Provided with inconsistent results in macaques, an assessment of stereotypic behavior in humans can provide some important synthesis to how individual differences in temperament may relate to stereotypic behavior. Temperament can be assessed both with behavior and established scales. These scales can more directly address latent variables that underlie behavior than observing the behaviors themselves. Assessments of human temperament such as those used in this dissertation may reveal relationships that can lead to hypotheses and comparisons in rhesus macaque behavior.

In addition to temperament, there are other metrics of behavioral tendencies that may reveal risk factors for developing stereotypic behavior in rhesus macaques. Two other tests used in the BBA, response to a novel object and to an unfamiliar human, have identified potential risk factors for developing stereotypic behavior. Infants that scored high on *activity* during the Human Intruder Test (as opposed to *emotionality*, *aggression*, and *displacement*) displayed more stereotypic behavior when stereotypies were assessed later in life (Gottlieb et al., 2013a).

A novel object test has been used in the BBA with infants as well as with adult macaques at another facility. Individuals have displayed variations in responses to novel objects from individuals that will explore the novel object, or quickly inspect it, to those that display inhibition and do not approach the object (Coleman, Tully, & McMillan, 2005). It was found in two studies that monkeys that had more contact with the novel object (Gottlieb et al., 2013a) or approached the novel object (Gottlieb, Maier, & Coleman, 2015) displayed more stereotypic behavior than monkeys that had less object contact or did not approach it at all. One of the proposed functions of stereotypic behavior is that it serves to decrease stimulation in a challenging environment. The results that macaques that perform stereotypies were also more apt to contact novel objects appear at odds with the notion that these macaques seek to decrease stimulation from their environment. These results suggest that macaques that perform stereotypic behaviors and also contact novel objects may be using these behaviors to increase stimulation from their environment.

3.2.5.2 Environment

3.2.5.2.1 Long-term risk factors

Researchers have identified environmental factors that contribute to the tendency to perform stereotypic behavior in primates. These factors, described below, suggest that stereotypies may indicate current or previous exposure to stressful situations. Rhesus macaques in laboratory settings are reared in different ways, from remaining with their mother in a large social group, remaining with their mother but not in a large social group, or being reared in a nursery without their mother but with varying access to similarly-aged playmates.

The development of stereotypies has been associated with macaques that were reared in a nursery either with a cloth surrogate and intermittent access to playmates

(Lutz et al., 2003; Rommeck, Gottlieb, Strand, & McCowan, 2009; Gottlieb et al., 2013a) or with continuous access to playmates (Champoux, Metz, & Suomi, 1991; Bauer & Baker, 2016). The social environment can also exert effects on stereotypic behavior later in life. Macaques that were reared and housed with conspecifics but then placed in housing without physical contact displayed more stereotypies than monkeys housed with other monkeys (Bayne, Dexter, & Suomi, 1992; Lutz et al., 2003; Vandeleest et al., 2011, Gottlieb et al., 2013a).

In addition to the social environment, there are other long-term housing and management conditions that may contribute to the development of stereotypies. Monkeys housed indoors have been found to be more likely to develop stereotypies than monkeys housed outdoors (Vandeleest et al., 2011; Gottlieb et al., 2013a; Gottlieb et al., 2015). There are also more subtle aspects of the environment that can be risk factors for stereotypic behavior. For example, macaques that were housed in the bottom row, closest to the ground, of multi-row cages were more vulnerable to developing stereotypies (Gottlieb, et al. 2013a) as well as those that were closest to the room entrance (Gottlieb et al., 2013a) or that did not have a foraging device affixed to their cage (Gottlieb et al., 2015). Finally, macaques that were involved in more research projects or blood sampling were more likely to develop stereotypies (Lutz et al., 2003; Vandeleest et al., 2011; Gottlieb et al., 2013a).

These environmental factors suggest that consistent encounters with stressful situations may contribute to the incorporation of stereotypic behaviors into the behavioral repertoire. However, these are retrospective analyses that analyzed why animals that already incorporated stereotypies into their repertoire may have developed them. It is unclear from these studies whether the stressful situations are immediate triggers for the performance of stereotypic behavior or contribute to stereotypies through indirect relationships.

3.2.5.2.2 Short-term risk factors

Researchers have performed short-term experimental manipulations in the environment to test the relationship between environmental factors and stereotypies. Usually these experiments are carried out in order to assess possible ways to decrease stereotypic behavior in captive animals. For laboratory-housed macaques, manipulable objects placed in the cage decreased stereotypic behavior (Novak, Kinsey, Jorgensen, & Hazen, 1998; Kessel & Brent, 1998; Cannon, Heistermann, Hankison, Hockings, & McLennan, 2016). However, interest in these objects has been shown to quickly wane (Pruetz & Bloomsmith, 1992). In order to maintain interest in enrichment items, objects are typically rotated in and out of monkey cages (Lutz & Novak, 2005).

Aside from the assessment of how environmental enrichment relates to stereotypic behavior, other short-term environmental factors for stereotypic behavior have not been investigated for laboratory-housed macaques with the exception of predictability which is discussed in more detail below. Provided with the prevalence of stereotypies in captivity, understanding what external factors may directly contribute to an animal’s motivation to perform stereotypic behaviors has the potential to inform captive animal management and positively contribute to animal welfare.

3.2.5.2.3 Predictability

Much of animal care in captivity is predictable because animals are fed, cages are cleaned, and animals interact with caretakers at fixed times of day. Although there are substantial benefits to living in predictable environments, the relationship between predictability and stereotypic behavior in captive primates remains unclear. Two experimental manipulations in which previously temporally predictable feeding regimes were made unpredictable demonstrated that stereotypies were more frequent in predictable conditions than unpredictable ones (chimpanzees: Bloomsmith & Lambeth, 1995; stump-tailed macaques: Waitt & Buchanan-Smith, 2001). However, in two other studies, the opposite conclusion was reached: stereotypies were more frequent when monkeys were fed on an unpredictable schedule than a predictable one (rhesus macaques: Gottlieb, Coleman, & McCowan, 2013b; capuchins: Ulyan et al., 2006).

The contradictory results for captive primates suggest that the relationship between stereotypies and predictability for captive animals may depend on other factors such as aspects of the current captive environment. Most of the studies in primates were conducted in social groups, with the exception of Gottlieb et al.'s (2013b) work with rhesus macaques housed in non-physical contact. When food is provisioned to social groups, primates experience feeding competition as individuals either scramble for or engage in a contest for food resources against other members of the group. In this light, the predictability of when or how much food will be acquired may vary greatly between individuals in a social group and in differently sized social groups. In primate social groups, these manipulations of the feeding schedule in order to assess the relationship between predictability and stereotypies may thus be unreliable for understanding behavioral responses to predictability.

On the other hand, food is more closely regulated for rhesus macaques housed in non-physical contact and they do not experience feeding competition. As the relationship between when food is provisioned and then attained by the animal is more straight-forward for these rhesus macaques, the predictability of the feeding event can be more effectively manipulated. As feeding competition can complicate feeding predictability in social groups, it is more compelling that predictability was more precisely manipulated with rhesus macaques housed in non-physical contact than in previous studies of primate social groups.

3.3 Experiment 2: The effects of predictability of delays in a delayed response task on behavior in adult rhesus macaques

As with Experiment 1 in humans, I manipulated predictability in the delays of a response task in macaques. However, in contrast to the DNMS task in humans, macaques are likely to cease participating in the task if they are incorrect and do not receive a reward. The task thus was not a DNMS task with correct and wrong choices but rather a delayed response task (DRT) in which there was only one choice that was always correct. In addition, as opposed to Experiment 1 in humans, the DRT experiment was conducted first before the Wait experiment. This was done in order to utilize the monkeys' expectation for the DRT experiment when conducting the Wait Experiment. The monkeys were first trained on the predictable delay length of 15 seconds. During data collection, what was manipulated was whether the delays in the DRT were the predictable length (15s) or other unpredictable lengths.

3.3.1 Hypothesis and Predictions

Previous literature on how captive primates respond to predictability is mixed (Bloomsmith & Lambeth, 1995; Waitt & Buchanan-Smith, 2001; Ulyan et al., 2006; Gottlieb et al., 2013b). However, the most relevant study in terms of using the same species in a similar environment is Gottlieb et al.'s (2013b) work with adult rhesus macaques housed in non-physical contact with other macaques. The researchers found that stereotypic behavior decreased when the macaques experienced temporal predictability in the animal care routine via feeding, enrichment distribution, and cleaning (Gottlieb et al., 2013b). This result aligns with the human literature that found that humans prefer certain or predictable conditions to unpredictable or uncertain ones (Munichor & Rafeli, 2007; Brakewood et al., 2015; Watkins et al., 2011). As with Experiment 1 in humans, I hypothesized that the frequency of stereotypic behavior is related to the predictability of the event. More specifically, I predicted that macaques would perceive the unpredictable condition as a more challenging environment and perform more stereotypic behavior in an unpredictable condition than in a predictable one during the Wait Experiment.

3.3.2 Methods

3.3.2.1 Subjects

The rhesus macaques that participated in this dissertation (n=14) were housed at the UMass Amherst Primate Laboratory. All macaques were adults, ranging in age from 12-24 years old, with the mean age of 15. The majority were male (female=5). Twelve monkeys originated from the National Institutes of Health Animal Center in Poolesville, MD and arrived at UMass either as juveniles (n=8) or as adults (n=4). Two monkeys were born at UMass as offspring to two different males and females housed at UMass (see Appendix F for demographic information). Four monkeys were pair-housed with a same-sex conspecific and another four had protected contact with an adjacent monkey; the rest were housed without physical contact due to pair incompatibility but had close proximity to other monkeys in their colony room. Since their arrival at UMass, the macaques have continuously participated in behavioral and cognitive studies conducted by UMass students. However, there have been no previous studies that assessed macaque response to predictability.

3.3.2.2 Housing and Experimental Setting

The monkeys were housed indoors in two suites. In each suite, there was a common hallway and the monkey rooms had opaque doors that opened into this hallway. All monkeys were housed in a room with at least one other monkey and no more than three other monkeys. There were two forms of housing conditions for the monkeys, a pen (n=8) or an Allentown® cage (n=6). The pens were fenced-in cubes with an open bottom that was supplemented with wood shavings. There were shelves, perches, and hammocks so that the monkeys could access both the ground level and a higher level. Allentown® cages are typically the most common housing used for laboratory-housed macaques in a wide variety of facilities (e.g. Gottlieb et al., 2013b). These consist of a large metal cage on wheels with four quadrants, two on top and two on the bottom that can be opened or closed. All of the macaques housed in Allentowns® in the UMass facility had full access to all 4 quadrants.

As for the general husbandry procedures for the macaques, beginning at approximately 8 am, the monkeys received a treat (fruits, grains, or peanuts) during a morning health and wellbeing check. They were then fed Purina monkey chow in a fixed amount customized for each monkey based on body mass. All monkeys had *ad libitum* access to water. Between 9-10 am, the animals' behaviors were recorded in 5-min samples; followed by cognitive testing between 10-11:30 am. The daily enrichment program was implemented around 12 pm and, depending on the day, consisted of ice cube treats, presentation of videotapes, exposure to music, and rolled-up paper bags containing treats. At approximately 2 pm, the monkeys received their afternoon ration of Purina monkey chow. After that, additional behavioral data were collected. The light cycle in the colony rooms was 13:11 (0700-2000) and the rooms were maintained at 23ºC between 35-50% humidity.

3.3.2.3 Apparatus

Monkeys in the UMass Primate Lab all approach tangible objects arranged on a board. The monkeys reach through their cage to interact with an apparatus placed within their reach but outside of the cage. I constructed a new apparatus for the DRT (Figure 3.1) that attached to a camera tripod via a camera mount. As the home cages for each monkey were variable in terms of where they can sit and reach to interact with the apparatus, the tripod height could be adjusted so that every monkey could reach the apparatus. A large black rubber rectangle was attached to the bottom of the tripod as a counterweight in order to prevent the monkeys from pushing the apparatus over.

The stimulus used for the DRT was a 4cm yellow-colored Brio® wooden high-gloss block intended for children. This block was affixed to a platform and slid in place on the apparatus. The macaque pushed the block back in order to reveal a food well below that had a treat (Figure 3.1). The reward in the DRT was either half of a raisin or a quarter of a peanut, as these are similar in size but some monkeys prefer peanuts to raisins and vice versa.

Figure 3.1 The apparatus used in the Delayed Response Task (DRT) for Experiments 2 and 3. The photos display the front (left) view, back (center), and side (right) views of the apparatus. The center image is Friday (N01) using the apparatus in an Allentown cage and the right image is Coby (V43) using the apparatus in a pen.

3.3.2.4 DRT Procedure

The general logistics of a DRT trial (depicted in Figure 3.2) was that the apparatus was placed in the typical home cage testing position but out of reach of the test monkey. I first announced the trial number for later information when scoring via video. In clear sight of the test monkey, I then held up the treat and baited the apparatus by placing the treat in the food well and pushed the yellow block over the baited food well to cover the treat. Then I placed an opaque white board occluder in front of the apparatus. The occluder was in place both to obscure the apparatus during the delay and also to increase the precision of the delay. The delay ended when the apparatus was placed in front of the monkey cage and the occluder was removed. In order to maintain a consistent overall session time between sessions, the monkeys had 30 seconds to engage with the apparatus and obtain the treat following the delay. If the monkey did not respond within 30 seconds, then the apparatus was pulled back, and the trial was repeated. Otherwise, as soon as the monkey obtained the treat, I pulled the apparatus back out of reach and began the next trial.

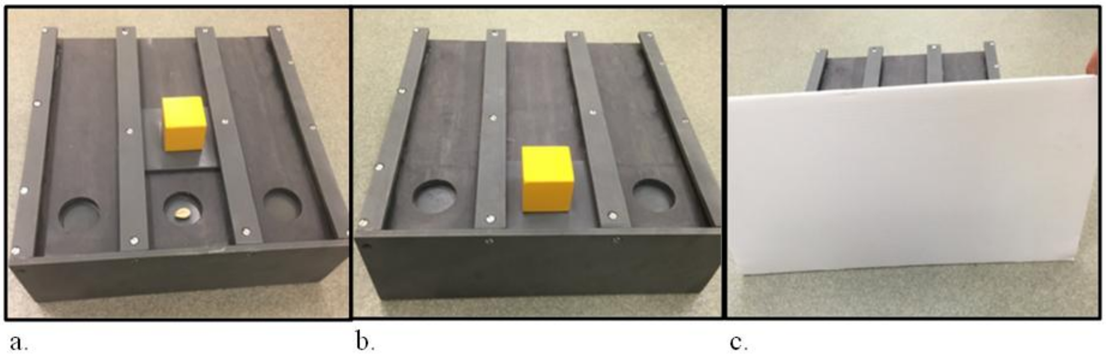

Figure 3.2 (a,b) The logistics of the DRT visualized on the apparatus. The treat (a peanut, pictured) was baited into the center treat well (a). The yellow block was then slid over the treat (b). The white occluder (made of polypropylene plastic sheet) was then placed in front of the apparatus for the duration of the delay (c). When the delay ended, the occluder was removed and the monkey slid the yellow block back and obtained the treat.

3.3.2.5 Training

While the human participants read instructions on how to complete the DNMS task, the rhesus macaques had to be incrementally trained on how to complete the DRT. The monkeys were first familiarized to the apparatus in June 2016, in which they approached the apparatus and successfully obtained a treat by pushing the block back in order to expose the food well underneath. This familiarization was conducted on all monkeys with 2 trials per day. When a monkey was able to successfully push back the block and obtain the treat on both trials in a day, then the monkey was considered successfully familiarized.

All monkeys became familiarized except for one, N02 (Lily). N02 approached the apparatus and took a treat when placed in front of the block, but she would not approach the apparatus if she did not already see the treat. Her behavior toward the apparatus suggested that this was not due to a lack of understanding of the treat location but rather an aversion to directly touching the apparatus instead of the treat. N02 participated in Experiment 2 with the modification that the treat was placed in front of the block in lieu of the treat well underneath the block.

Following familiarization with the apparatus, I began training the monkeys on the DRT procedure. During the experiment, the predictable condition consisted of 15-second delays and the unpredictable condition consisted of varying delays between 0 (no delay) and 30 seconds. The training period thus consisted of running multiple sessions with 15-second delays in order to create the predictability for the predictable condition during the future experiment. The monkeys participated in the training period for 4 weeks in August 2016 in which the monkeys participated in a maximum of 9-10 trials per day totaling 148 training trials by the end of the training sessions. The training trials were counted only if the monkey responded and took the treat. If the monkey did not engage with the apparatus after 30 seconds, then I repeated the trial. This occurred rarely, seven times in total involving three different monkeys. If the monkey did not engage for 2 trials in a row, then training was stopped for the session. This occurred rarely, three times in total for two of the three above monkeys. These trials were run later in the day or on the next day until 148 trials were ultimately reached.

3.3.2.6 Data Collection Procedure

As this experiment tested the effects of predictability on macaque behavior, it was of the utmost importance to test monkeys on days when they received their normal husbandry routine described earlier. On the few occasions of physical plant repairs to the facility, or unexpected loud noise or unfamiliar human visitors entering the suite, the monkeys were not run in the test paradigm on those days. On each testing day, I tested seven monkeys in a suite and then on the next day, the other suite of seven monkeys was tested. Within each suite, one monkey in a room participated in the study at a time and then I switched to another room. This is a standard testing procedure in the UMass Primate Laboratory in order to minimize any potential learning or behavioral effects from one monkey in a room observing the trial of another monkey. The order of monkeys tested was block randomized with some restrictions in order to ensure that rooms with more monkeys were still run in an alternating order.

At the beginning of a test session, I set up a video camcorder (Canon VIXIA HF R700) in the room in order to capture the movements of the monkey throughout its home cage during the test session. Data were ultimately collected via scoring behaviors from this video recording. Because of our extensive cognitive testing protocols, the UMass macaques are habituated to video cameras. Following camera set-up, I then removed the water bottles and thus water access for the monkey. Because test sessions lasted for on average 5 minutes and 54 seconds (±44s) this was not considered a water restriction that may affect behavior or compromise animal welfare. Yet, this was done to maintain focus on the task at hand. The monkeys were not deprived of food or water prior to the study and food was available *ad libitum* during the study. Finally, for pair-housed monkeys, pairs were separated in order to prevent disruption from the task or competition for treats for the test monkey. Monkeys that were pair-housed were trained to go to either the left or right side of their home cage for separation during routine husbandry events and for all cognitive testing.

Each testing session consisted of 9 trials. Predictable test sessions had a standard delay of 15 seconds across sessions and trials, whereas Unpredictable sessions had delays of different lengths between trials and a different order of delays between sessions. In the Unpredictable condition, the delay lengths varied from 0 seconds through 30 seconds with possible increments of 0, 5, 10, 15 (delay associated with the predictable condition), 20, 25, or 30 seconds. The 15-second delay was included in Unpredictable trials in order to probe monkey response to predictable trials in the midst of other unpredictable trials. In the 9 trials of an unpredictable session, the first (1), middle (5), and last (9) trials had 15-second delays. The other trials of an unpredictable session were block randomized for delay lengths of 0, 5, 10, 20, 25, or 30 seconds. Ultimately, both the predictable sessions (9 trials with 15s delays) and the unpredictable sessions (9 trials of variable delays) lasted approximately 135 seconds (2.25 minutes) in total depending on monkey response time after the delay within the 30-second timeframe.

As opposed to Experiment 1 with humans, each monkey received more than one DRT session. Each monkey received 8 test sessions in the order of Predictable (P), Unpredictable (U): P, U, P, U, P, U, P, U. Each monkey in a room received the same condition on a test day. In addition, the unpredictable condition with respect to the order of delays was the same across monkeys. Within an unpredictable session, all monkeys received the same order of trial delays. Across unpredictable sessions, the delay order varied.

3.3.2.7 Video Coding of Macaque Behavior

There were four categories of behavior investigated based on a subset of existing categories (see Appendix G): stereotypies, anxious behavior (yawn and scratch), aggressive behavior (cage shake and threat behaviors), and tactile and oral exploration (manipulation of objects or features in the environment with hands or mouth). These behavioral categories can be described as four different possible ways to respond to a delay. For stereotypies, I combined all types of stereotypies (pacing or self-directed ones such as eye-poke and self-stroke) into a stereotypic behavior category. Stereotypies have been demonstrated to occur as animals wait for an anticipated event; however, it is unclear why, as outlined in the Introduction. Yawn and scratch behaviors may represent an anxious response as these behaviors increase with anxiogenic drugs and decrease with anxiolytic drugs (Schino, Perretta, Taglioni, Monaco, & Triosi, 1996). Cage shake and threat are considered aggressive social signals and may represent a heightened emotional response to the delays. Finally, tactile and oral exploration behaviors may represent a response of filling the waiting period time with other stimulating activities. Self-injurious behaviors such as self-bite were also noted if they occurred, although these behaviors typically occurred too rarely to be analyzed.

Macaque response during the delays was scored from video. Videos were scored on a computer at the frame-level (30 frames=1 second) using MPEG Streamclip software. All observers (n=2) achieved interobserver reliability above 90% and were blinded to the experimental condition in the video.

3.3.2.8 Statistical Analyses

Because the delays between Predictable and Unpredictable conditions had different durations, I did not use total duration measures as I used in Experiments 1 and 3. Instead, I calculated both frequency counts as well as an average duration score. I calculated the average amount of time spent in each behavior in each session. I then calculated an average duration score across the four sessions for the Predictable and Unpredictable conditions. Preliminary analyses demonstrated that the behavioral data were not normally distributed via Shapiro-Wilk tests (see Appendix H). For both the frequency and duration measures of behavior, I used the nonparametric Wilcoxon signed-ranks test to assess potential differences in behavior between the Predictable and Unpredictable conditions.

3.3.2.8.1 Assessment of potential sex and housing effects on behavior

There were other potential intervening variables that could account for differences in response to the experimental conditions. These include sex differences (5 females; 9 males) and housing (6 in Allentown cages, 8 in pens). I tested whether these variables accounted for differences in response to the DRT experiment with two mixed design ANOVAs, the first one used sex as the between-subjects variable and the experimental conditions as the within subjects variables and the second one used housing as the between subjects variable and the experimental conditions as within subjects variables. I used the Levene’s test to assess for equality of the variances. If the variances were significantly not equal, then I log-transformed the data prior to running the ANOVAs.

3.4 Results

3.4.1 The effects of Predictable and Unpredictable conditions on behavior

The Delayed Response Task paradigm elicited stereotypic behavior from the monkeys. However, it was the least frequent behavior of the categories that were measured (Figure 3.3a). Yawning and scratching behaviors were the most frequent behaviors monkeys performed during the experiment, followed by tactile-oral exploration and then cage-shaking and threat behaviors. For the average duration of behaviors during the DRT, yawn and scratch behaviors were also the longest, but the monkeys spent more time engaging in stereotypic behaviors than tactile-oral exploration or cage shake and threat (Figure 3.3b). Given that the test session averaged 6 minutes per monkey and given that a small percentage of time was devoted to retrieving the treat, nonetheless, the total time spent in all of these behaviors comprised no more than 5-10% of the time available to express them.

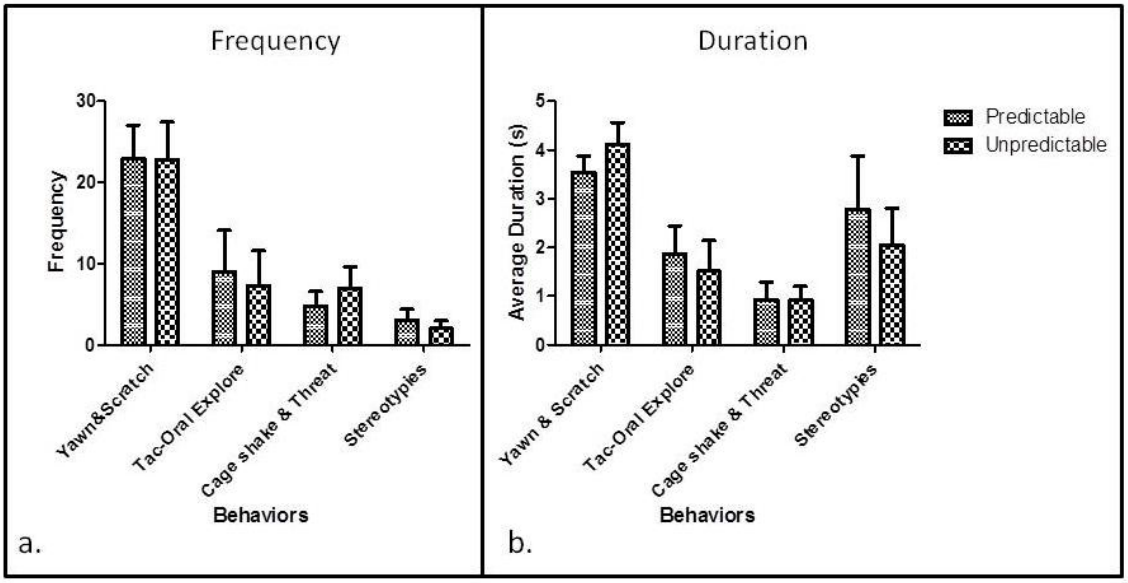

Figure 3.3 (a,b) Average frequencies of behaviors observed in Experiment 2, by condition (a). Average behavior duration observed in Experiment 2, by condition (b). Bars represent ±1 SEM.

There was no statistically significant difference in the frequency or duration of stereotypic behavior in the Predictable and Unpredictable conditions using the Wilcoxon signed-rank tests (see Table 3.1). However, frequencies of cage shake and threats were significantly higher in the Unpredictable condition than in the Predictable condition (*p*=0.035), although there was no difference between durations of cage shake and threat behaviors (see Table 3.1). The converse was true for yawn and scratch with no difference in frequency but a trend for increased duration in the Unpredictable condition. There was also a trend for an increase in the average duration of yawn and scratch behaviors in the Unpredictable condition although there was no difference in the frequencies of yawn and scratch behaviors. Finally, there was no difference between conditions for frequency of tactile-oral exploration or average duration (Table 3.1).

Table 3.1 Wilcoxon Signed-Ranks Tests for behaviors during DRT. Bolded values represent significant. *$p<0.05$; +$p<0.06$.

	Wilcoxon-Signed Rank Tests	
Behaviors	Frequency	Duration
Stereotypies	Z=1.197, p=0.231	Z=0.652, p=0.515
Cage Shake & Threat	**Z=2.111, p=0.035***	Z=1.07, p=0.285
Yawn & Scratch	Z=0.0, p=1.00	**Z=1.92, p=0.056**$^{+}$
Tactile & Oral Exploration	Z=1.064, p=0.287	Z=0.874, p=0.382

3.4.2 Individual differences in DRT response

I assessed how the behaviors differed between individual monkeys. I collapsed across conditions and used the average of the Predictable and Unpredictable conditions (Figure 3.4). All four behavioral categories were observed in 9/14 monkeys. Every monkey performed yawn and scratch behaviors, and all but one (Linus) performed tactile and oral exploration. For stereotypies and cage shake and threat behaviors, 9/14 monkeys performed these behaviors. While these behaviors were prevalent, there was variation with no observable consistency between individuals in the time spent performing the four categories of behaviors.

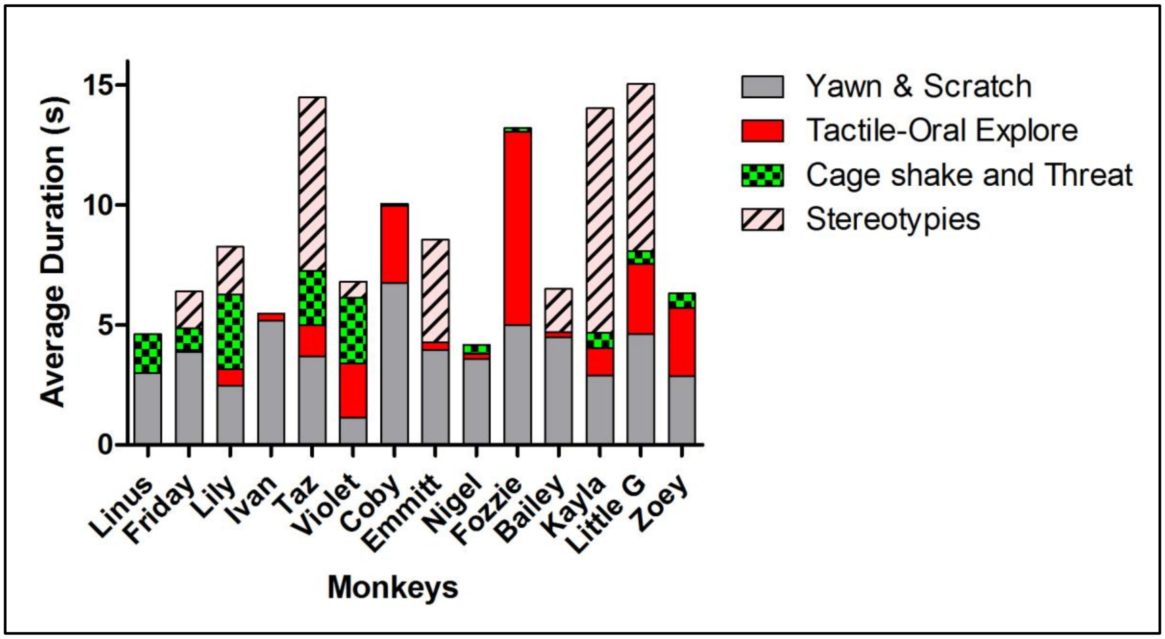

Figure 3.4 Average duration of behaviors performed, by monkey.

3.4.3 The effects of sex and housing factors on behavior in the DRT Experiment

There were no effects or interactions of sex and housing on stereotypic behavior in the DRT Experiment (see Appendix I). However, there were significant main effects of sex for both yawn and scratch and cage shake and threat behaviors. For yawn and scratch, both the frequency and duration, males yawned and scratched more than females, (trend detected, *p*=0.051), (see Figure 3.5a) and duration: (*p*=0.007) (see Figure 3.5b). The significant main effect of sex for cage shake and threat behaviors was the converse. Females performed longer of cage shaking and threat behaviors than males (*p*=0.008) (Figure 3.6). There were no effects of housing on yawn and scratch and cage shake and threat behaviors. Finally, there were no effects of sex or housing on tactile oral exploration.

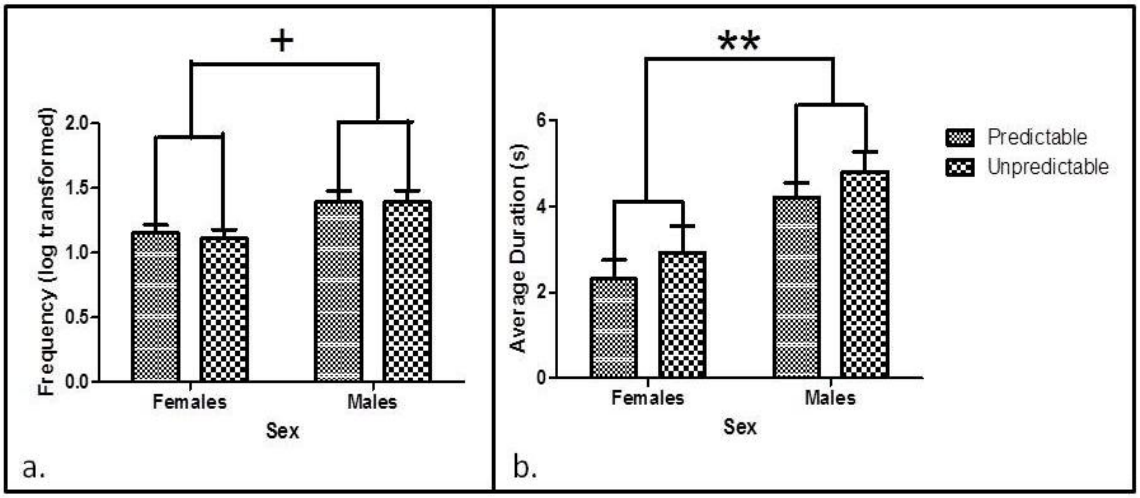

Figure 3.5 (a,b) Frequency of yawn and scratch behaviors, by sex (a). Average bout duration of yawn and scratch behaviors, by sex (b). += 0.06>*p*>0.05, **=*p*<0.01. Error bars represent ±1 SEM.

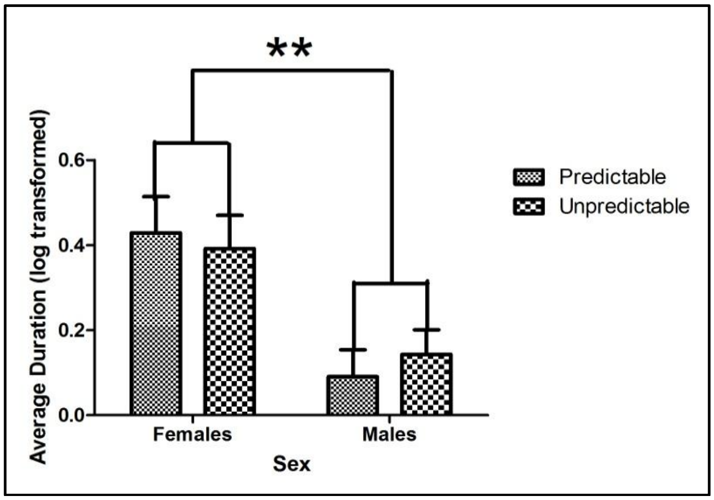

Figure 3.6 Main effect of sex in duration of cage shake and threat behaviors.

3.4.4 Rare but Notable Behaviors

Self-injurious behavior such as self-bite is present as a mild form in some of the monkeys. In Experiment 2, there were 7 observed instances of self-bite by 4 different monkeys. All occurred during Unpredictable sessions. Secondly, two of the monkeys who performed self-bite did not have a history of self-biting behavior. The qualitative difference and elicitation of self-bite in monkeys that typically do not display this behavior is potentially an important response to predictability in macaques.

Secondly, social behaviors other than threat were rare but did occur. There were 3 rump presents by one monkey. Two of these behaviors occurred in two separate Unpredictable sessions and during the one 30-second delay of the session. While threats and cage shake were the most common social behaviors observed, the rump present behavior may represent an alternative response of communicating with the experimenter in order to end the delay during the longest wait period experienced in the study.

3.5 Discussion

Stereotypic behavior did not occur frequently in the experiment nor did it increase in the Unpredictable condition, thus failing to support my prediction. Instead, the monkeys responded with anxious and aggressive behaviors via yawning, scratching, cage shaking, and threats, although they spent less than 10% of the available time engaging in these behaviors. Additionally, there was a significant increase in the frequency of cage shake and threat behaviors and a strong trend for an increase in yawn and scratch length in the Unpredictable condition. These results suggest that the macaques were sensitive to the difference in conditions and the Unpredictable condition was the more challenging condition.

The macaques responded to the DRT with what can be interpreted as emotional responses of anxious and aggressive behaviors. As opposed to the humans in Experiment 1, these macaque behaviors elicited do not suggest that the macaques were bored or under stimulated but rather in a higher arousal state. For whether monkeys responded with anxious or aggressive behaviors, there may be an influence of sex in that females displayed longer durations of cage shake and threat behaviors than males. Males, on the other hand, yawned and scratched more, but this finding was not universal. For example, two monkeys rarely displayed aggressive behaviors but instead frequently displayed anxious behaviors (i.e., Ivan and Coby).

Because the experimenter was in the room with the macaques and controlling both the treat and the length of the delay, the macaques may have used social signals in order to influence ending the delay or obtaining the treat from the experimenter. However, it is also possible that the social signals were expressed because of displaced irritation, or frustration without the additional intention of attempting to influence the experimenter.

Two previously discussed studies found that agonistic behavior, of which aggression is one type of agonistic behavior, increased in the delayed or unpredictable conditions. In stump-tailed macaques, agonism generally increased before feeding and then decreased when fed. However, if feeding was delayed, then agonistic behaviors increased until they were fed (Waitt and Buchanan-Smith, 2001). Additionally, Bloomsmith and Lambeth (1995) measured agonistic behavior in chimpanzees and while not explicitly discussed in the paper, the presented data demonstrated that agonistic behavior was higher than abnormal or stereotypic behavior in the unpredictable condition. Two other studies that pertained to predictability in captive primates either did not assess social behavior beyond vocalization (Gottlieb et al., 2013b) or did not distinguish between agonistic or affiliative social behavior (Ulyan et al., 2006). While not as widely discussed as stereotypic behavior, it is possible that there is a stable relationship between predictability and agonistic behavior in captive primates.

The previous studies assessed the relationship between stereotypic behaviors and predictability by manipulating the predictability of husbandry routines (Bloomsmith & Lambeth, 1995; Waitt & Buchanan-Smith, 2001, Ulyan et al., 2006; Gottlieb, et al., 2013b). All of these studies manipulated the timing of feeding although Gottlieb et al. (2013) added other components of the husbandry routine of cleaning and enrichment distribution. To my knowledge, my study is the first to assess behavioral response to predictability within a task paradigm in captive primates. It is possible that the macaques in my study would respond to a manipulation of predictability of their feeding routine in a similar way as captive primates in the previous published studies. However, my results demonstrate that when the environmental conditions elicit either aggressive or anxious behaviors, these emotional states do not align with those that elicit stereotypic behavior. Secondly, it is possible that stereotypic behavior is unconnected to these immediate variations in emotional state and requires more long-term environmental factors as seen in changes in husbandry routines in order to elicit stereotypic behavior.

3.5.1 Conclusions

My predictions were not supported as the macaques did not perform stereotypic behaviors differently across the conditions. However, emotional responses via anxious and aggressive behaviors were elicited and the monkeys differentiated their behavior between conditions by performing more cage shake and threat behaviors and longer yawn and scratch behaviors in the Unpredictable condition. The paradigm of an experimenter manipulating the delay before the treat distribution may have shifted macaque behavior to emotional responses rather than stereotypic behavior. Rather than display stereotypic behavior, the macaques either displayed behaviors indicative of increased anxiety or demonstrated displaced irritation or frustration, or used social signals to influence the experimenter.

3.6 Experiment 3: The effects of predictability of a waiting period on behavior in adult rhesus macaques

After the macaques had experience with the apparatus and DRT, I utilized this anticipation and ran the Wait experiment. The DRT apparatus was placed in front of the test monkey, thus signaling to the monkey that it was about to participate in the task. However, the experimenter then left the room for 5 minutes. This was the same duration as the waiting time as in the Experiment 1: Experimenter Wait paradigm with humans. What was manipulated in this paradigm was amount of information the macaque received about when the experimenter will return to the room based on whether the room door was open (Predictable) or closed (Unpredictable).

It is currently unknown whether the UMass monkeys perceive a difference in the timing of human entry when the door is opened or closed. Yet, it tends to be the practice for the lab personnel and caretakers to leave a room door open when they are working in that room. On the other hand, when personnel and caretakers are no longer working with monkeys in a particular room, then the door is shut and remains closed.

3.6.1 Hypothesis and Predictions

I hypothesized that the frequency of stereotypic behavior was related to the predictability of the event. More specifically, I predicted that macaques would perceive the unpredictable condition as a more challenging environment and perform more stereotypic behavior in an unpredictable condition than in a predictable one during the 5-minute delay in the Wait Experiment.

3.6.2 Methods

3.6.2.1 Subjects

The same 14 rhesus macaques housed at UMass were used in Experiment 3. As with Experiment 2, the monkeys were tested in their home cage, with the cage and pairing arrangements unchanged from Experiment 2.

3.6.2.2 Training

This experiment sought to explore macaque behavior in a situation that they commonly encounter in their interactions with human caretakers and researchers. Because I used a naturalistic scenario, I did not train the macaques on the procedure. On the day prior to the first day of data collection, there was a run-through of the procedure in each monkey room, in the predictable condition. This was done not for monkey training but rather to finalize camera positioning in the room as well as other experiment logistics. Furthermore, the apparatus used in Experiment 3 was the same as Experiment 2, so the monkeys were already trained on how to approach and obtain a treat from the apparatus.

3.6.2.3 Wait Experiment Procedure

Experiment 3 started three days after the cessation of the Experiment 2 in order to fully utilize the macaques' expectation to participate in DRT of Experiment 2. I conducted Experiment 3 at the same time of day (10am-12pm) as Experiment 2. All 14 monkeys participated in a test session in one day. All rooms were tested in one session in one day. Because the opening or closing of the room door was a large stimulus apparent to all monkeys in the room, all monkeys in a room were tested and video recorded in one session simultaneously. The order for the 5 testing rooms for one day was block-randomized across sessions with the restriction that I alternated between the 2 suites between sessions. Each room received the same condition on a testing day.

To begin a test session, a video camera (Canon VIXIA HF R700) and the DRT apparatus that the monkeys had prior experience working with was placed in the center of the room. Whereas Experiment 2 involved interaction between the experimenter and monkeys, in Experiment 3, I sought to create a scenario more similar to husbandry procedures in which humans may be in the room or setting things up but not yet directly interacting with the monkeys. This meant that as opposed to when the monkeys participated in Experiment 3, I did not separate pairs or remove water bottles. The monkeys were not deprived of food prior to the study and food was available *ad libitum* during the study. Once the apparatus was rolled into the room, I then exited the room for 5 minutes, as was done with the human participants. What was manipulated was the information provided to the monkey about when I would return based on whether the room door was open (Predictable) or closed (Unpredictable). For both conditions, the hallway outside the door remained clear of human activity. Following the 5-minute wait period, I then returned to the room and commenced the DRT with the apparatus already placed in the room.

When I returned to the room following the 5-minute experiment, I ran each monkey through 1 trial of the DRT and the monkey ultimately approached the apparatus to obtain a treat. This ended the session for that monkey. The order for monkeys in a room to participate in the DRT was randomized so that every monkey in the room was tested in every possible order depending on the number of monkeys in the room. In contrast to the human experiments, each monkey received more than one Wait session. Each monkey received 8 test sessions in the order of Predictable (P), Unpredictable (U): P, U, P, U, P, U, P, U.

3.6.2.4 Baseline Morning Data Collection

Because Experiment 3 entailed assessing macaque behavior for 5 uninterrupted minutes, I compared the behaviors observed in Experiment 3 to another scenario in which macaque behavior is measured in a 5-minute session. Every weekday morning at 9 am, undergraduate research assistants, lab technicians, and graduate students collect behavioral data on every monkey. For each room of monkeys, the order for which monkeys are observed first through last are randomly determined prior to data collection.

For this morning data collection, an observer walked into a monkey room, sat down and observed each monkey in the room for a separate 5-minute session. A focal animal sampling procedure was used, and social signals were scored both with respect to initiation of the focal monkey and with respect to receipt from other animals in the room. Monkeys were habituated to all observers as one must spend a significant amount of time with the monkeys prior to being a reliable observer. Each monkey was then observed in real-time (sans video recording) with minimal interaction between the focal monkey and observer. Observers were instructed not to interact with the monkey aside from passive observation, but the monkey may have sent social signals to the observer. When finished with data collection for all monkeys in the room, the observer then left the room.

3.6.2.5 Sampling and Coding of Macaque Behavior

As with Experiment 2, I scored the Experiment 3 videos for: stereotypies, anxious behavior (yawn and scratch), aggressive behavior (cage shake and threat behaviors), and tactile and oral exploration (manipulation of objects or features in the environment with hands or mouth). I also noted whether any self-injurious behaviors occurred, although it is usually a rare behavior.

3.6.2.5.1 Assessment of Behavior during the Wait Experiment

I scored behaviors for Experiment 3 through two different sampling methods. The first method was an all-occurrence sampling method (Altmann, 1974) in which I scored every instance of my four behavioral categories. For the 14 monkeys, I generated a sum total of number of seconds each monkey was engaged in each of the behaviors for each experimental session. I then averaged the number of seconds for each behavior across the four Predictable and four Unpredictable sessions, so each monkey had one mean behavior score for each behavior in the Predictable and Unpredictable conditions, respectively.

I then re-sampled the videos using a modified frequency sampling procedure. I added this additional modified frequency data to my analyses because it allowed me to compare my Experiment 3 data to the AM Data collection that may serve as a potential baseline measure of macaque behavior. A form prepared for modified frequency data collection consisted of a grid of 20 columns representing 15-second intervals and a row for each potential type of behavior observed. After the 5-minute observation, the observer counted the total number of intervals that each behavior occurred in. This is a *modified* frequency count because this count represented the number of intervals that each behavior occurred in, whether once or multiple times, instead of a complete frequency count. With 20 intervals, the range of possible behavior scores was 0-20.

3.6.2.5.2 Comparison with AM Baseline Data

The modified frequency sampling procedure was performed both in real-time for the 9am morning data collection as well as through Experiment 3 videos. In this procedure, every behavior the monkey performs was accounted for, so there were additional behaviors not included in the all-occurrence sampling described previously. An ethogram for all of the behaviors measured is outlined in Appendix G.

In my analysis of the macaque behavior during AM data collection, I used all data collected between Tuesday September 6, 2016 and Tuesday November 1, 2016, resulting in 38 sessions. This range of dates included both the days that Experiment 3 was conducted as well as weeks prior and after data collection. There were no major disruptions or changes to the monkeys’ routines, feeding, or housing during this time. On November 2, a group of monkeys transferred rooms, so I did not use any AM data collected after the relocation.

3.6.2.6 Statistical Analyses

3.6.2.6.1 Assessment of Predictability of Waiting period on Behavior

Duration measures of the four behavior categories were not normally distributed via Shapiro-Wilk tests (Appendix H). In order to analyze the effect of predictability on macaque behavior, I analyzed the behavioral data with Wilcoxon Signed-Ranks tests as a nonparametric alternative to paired samples.

3.6.2.6.2 The effects of habitation on responses to predictability

I tested whether there were possible effects on habitation to the predictable or unpredictable conditions over the course of the eight sessions. To do this, I conducted paired samples t-tests with the four categories of behaviors for sessions 1 and 2 and sessions 7 and 8. These test the differences between responses in the first predictable trial and the first unpredictable trial and then separately test the responses in the last predictable trial with the last unpredictable trial. Results are presented in Appendix K.

3.6.2.6.3 Comparison of Wait Experiment Conditions and baseline AM Data

I tested for differences between the two experimental conditions and AM data condition by first placing the behavior in the three conditions of Predictable, Unpredictable, and AM data collection in an omnibus repeated measures one-way ANOVA. I used Mauchly's test of Sphericity to test whether the variances between all three measures were equal. If they were not equal, then I used a Greenhouse-Geisser correction for my significance values. If the ANOVA showed a significant difference within the three conditions, then I viewed the pairwise comparisons. I ran a paired-samples t-test on the two conditions that the pairwise comparison indicated were different.

3.6.2.6.4 Assessment of potential sex and housing effects on behavior

To assess possible contributions of sex (5 females; 9 males) and housing (6 in Allentown cages, 8 in pens) effects on behavior, I ran mixed-design ANOVAs using sex and housing as between-subjects variables and conditions as the within subjects variables. I performed these analyses both with the two Predictable and Unpredictable conditions represented in the all-occurrence duration data and with the three conditions of AM Data, Predictable and Unpredictable represented in the modified frequency data.

3.7 Results

3.7.1 The effects of Predictable and Unpredictable conditions on behavior

3.7.1.1 The effects of Predictability on behavior: duration measures

The Wait experiment elicited the behaviors of interest in terms of stereotypies, yawning and scratching, tactile-oral exploration, and cage shake and threat behaviors. Of these, stereotypic behavior was the category of behavior that the monkeys spent the most time performing (Figure 3.7). They spent the least amount of time engaging in cage shake and threat behaviors (Figure 3.7).

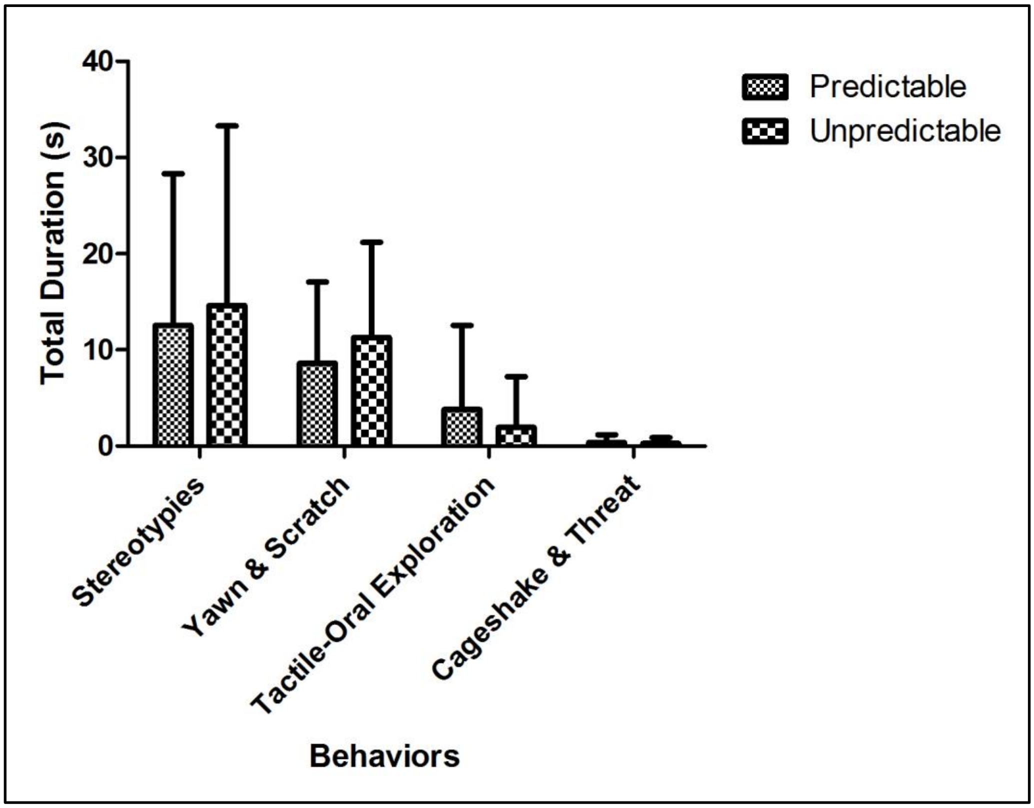

Figure 3.7 The total duration of behaviors observed, averaged across sessions, by condition. Error bars represent ±1 SEM.

There were no statistically significant differences between conditions for all four behavior categories of stereotypies (Z=0.178, p=0.859), yawn and scratch (Z=1.475, *p*=0.140), tactile and oral exploration (Z= 1.214, *p*=0.225), and cage shake and threat behaviors (Z=0.944, *p*=0.345).

3.7.1.2 The effects of Predictability on behavior: modified frequency measures

When measured through the modified frequency sampling method, there were also no statistically significant differences between conditions for all four behavior categories of stereotypies (Z=1.29, p=0.197), yawn and scratch (Z=0.945, *p*=0.345), tactile and oral exploration (Z=0.962, *p*=0.336), and cage shake and threat behaviors (Z=0.0, *p*=1.00) (Figure 3.8).

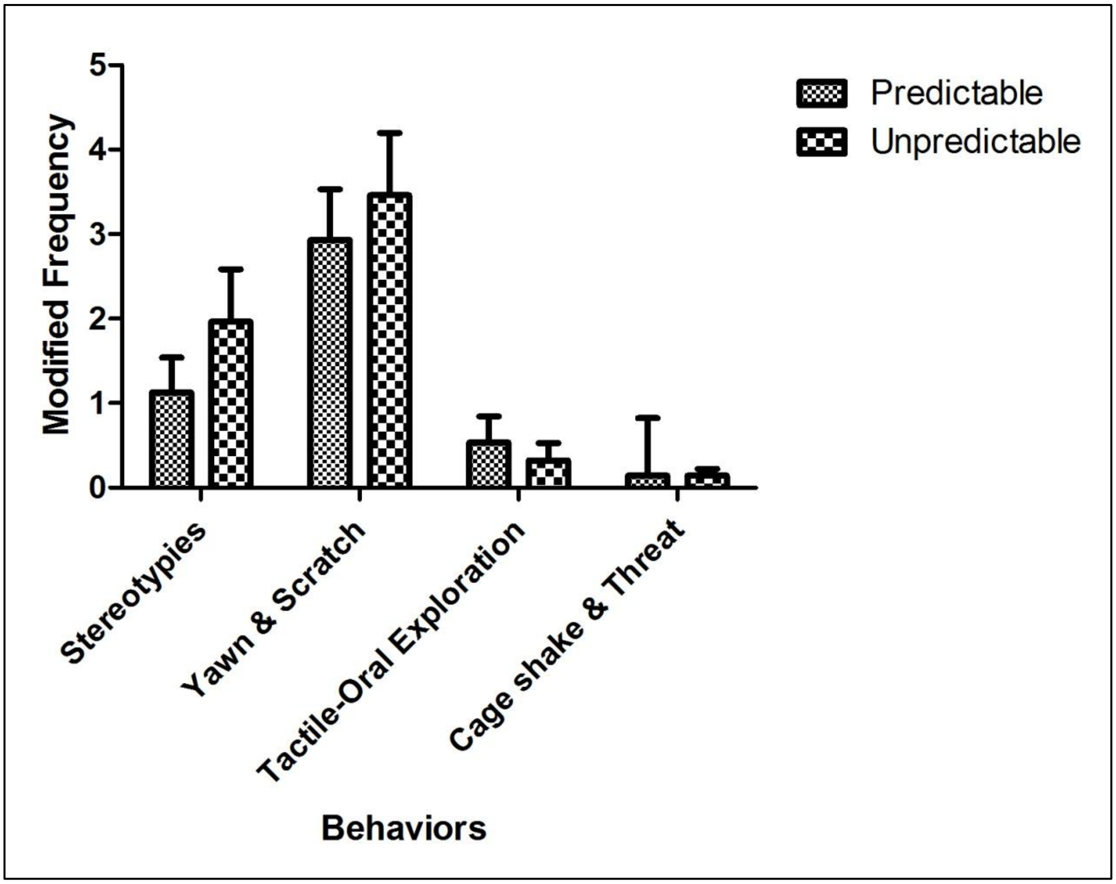

Figure 3.8 The modified frequency count of behaviors observed, by condition. Error bars represent ±1 SEM.

3.7.2 The effects of Wait Experiment conditions on AM behavior

With no statistically significant differences between the Predictable and Unpredictable conditions, I tested whether there was a response to the overall experimental manipulation when compared with AM Data, conducted daily at 9am, which was 1 hour prior to the time period in which I ran the Wait Experiment. There were no statistically significant differences between stereotypic behavior expressed during the two Wait conditions and AM Data F(2,26=2.34, *p*=0.116) (Figure 3.9). In addition, there was still no statistically significant differences between cage shake and threat behaviors (F(1.279, 16.623=2.086, *p*=0.165) and tactile and oral exploration (F(1.03,13.36)=3.389, *p*=0.087) (Figure 3.9).

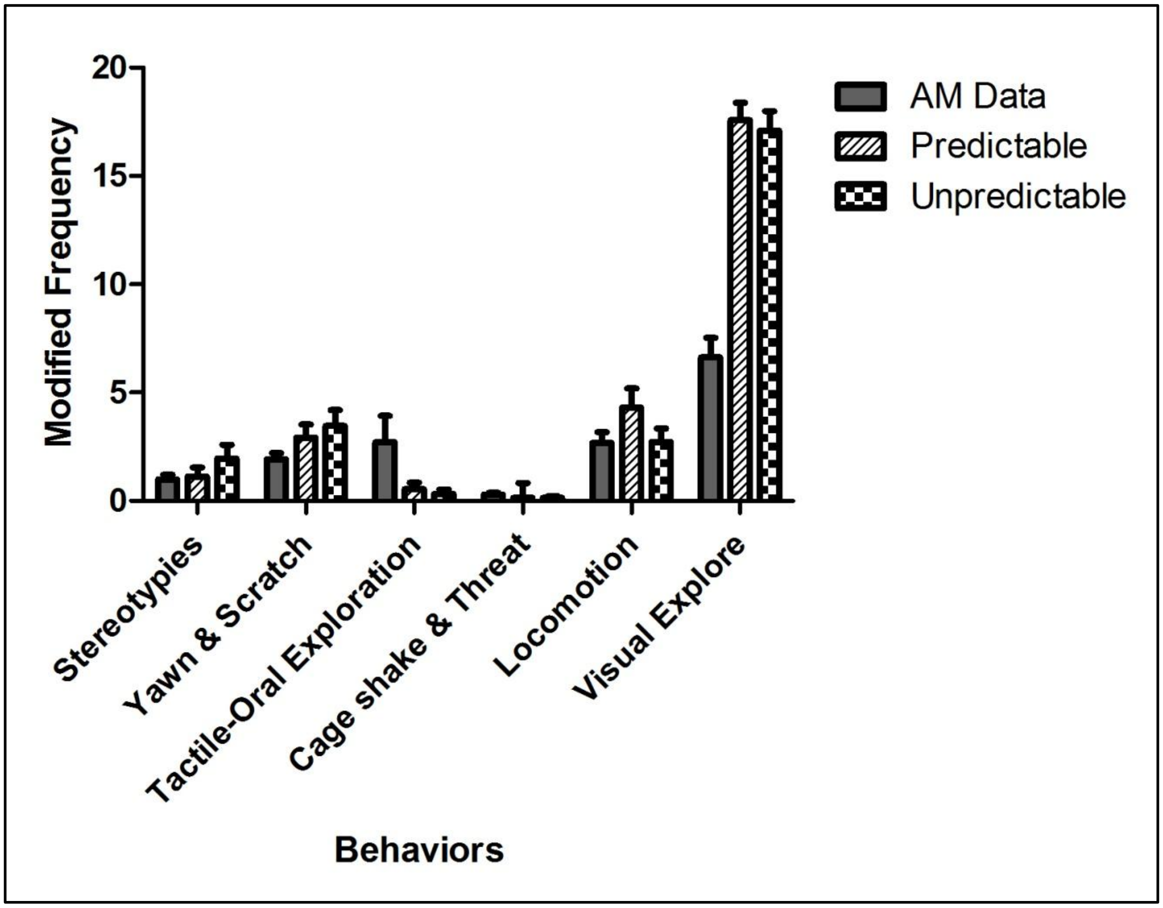

Figure 3.9 The modified frequency count of behaviors observed, by the two Wait conditions and AM Data. Error bars represent ±1 SEM.

However, Comparisons with AM Data revealed differences in response between the Wait Experiment and AM Data. These differences were made more apparent with the addition of two broader categories of behavior, called *visual explore* and *locomotion.* First of all, there was a significant difference in visual explore in the AM Data, Predictable, and Unpredictable conditions (F(2,26)=23.870, *p*<0.001). Post-hoc paired samples t-tests with a Bonferroni correction (p<0.01) demonstrated that the AM condition had significantly less visual explore behaviors than the Predictable (t(13)=5.412, *p*<0.0001) and the Unpredictable (t(13)=5.652, *p*<0.0001) conditions (Figure 3.9).

Secondly, there was a significant difference in locomotion in the AM Data, Predictable, and Unpredictable conditions (F(2,26)=5.388, *p*=0.011). Post-hoc paired samples t-tests with a Bonferroni correction (*p*<0.01) demonstrated that the locomotion in the Predictable condition was significantly higher than the Unpredictable condition (t(13)=3.350, *p*=0.005) and there was a strong trend for being higher than the AM Data (t(13)=2.48, *p*=0.028) (Figure 3.9). Finally, there was a marginally significant difference in the AM Data, Predictable, and Unpredictable conditions for yawn and scratch behaviors (F(2.26)=3.349, *p*=0.051). Post-hoc paired samples t-tests with a Bonferroni correction (p<0.01) did not demonstrate any significant differences within these conditions, although there was a trend for an increase in yawn and scratch behaviors from the AM Data to the Unpredictable condition (t(13)=2.237, *p*=0.037) (Figure 3.11).

I also analyzed two characteristics of the behavioral response in the three conditions: behavior rate, or how many behaviors the macaque iterated through, and behavior range, or how many different categories of behaviors the macaque iterated through. There were no differences in behavior rate between the AM Data, Predictable, and Unpredictable conditions (F(1.325, 17.220)= 2.958, *p*=0.095) (Figure 3.10). However, there was a difference between the behavior range in the AM Data, Predictable, and Unpredictable conditions (F(2,26)=4.841, *p*=0.016) (Figure 3.12). A post-hoc paired t-test determined that this difference was driven by a significant decrease in the Predictable behavior range from the AM Data (t(13)=3.01, *p*=0.010).

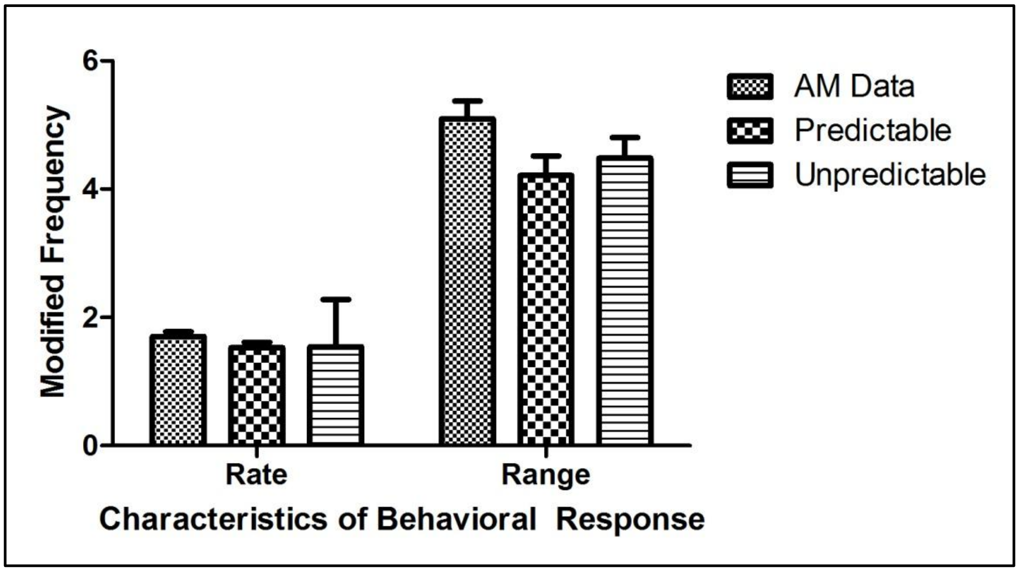

Figure 3.10 The rate and range of behaviors observed, by the two Wait conditions and AM Data. Error bars represent ±1 SEM.

3.7.3 The effects of sex and housing factors on behavior in the Wait Experiment and AM Data

3.7.3.1 Duration Measures and effects on the two Wait conditions

Mixed-design ANOVAs with sex or housing as the between subjects variable and behaviors measured as durations in the Predictable and Unpredictable conditions demonstrated that there were no effects or interactions of sex and housing on stereotypic, and tactile and oral exploration behaviors (see Appendix I). For yawn and scratch behaviors, there was a strong trend for a main effect of sex in which males yawned and scratched more than females (F(1,12)=4.64, p=0.052) (Figure 3.11a). For cage shake and threat behaviors, there was a significant interaction between sex and condition (F(1,12)=8.89 , *p*=0.011) (Figure 3.11b). None of the post-hoc t-tests indicated a direction for interaction as none of the analyses were statistically significant.

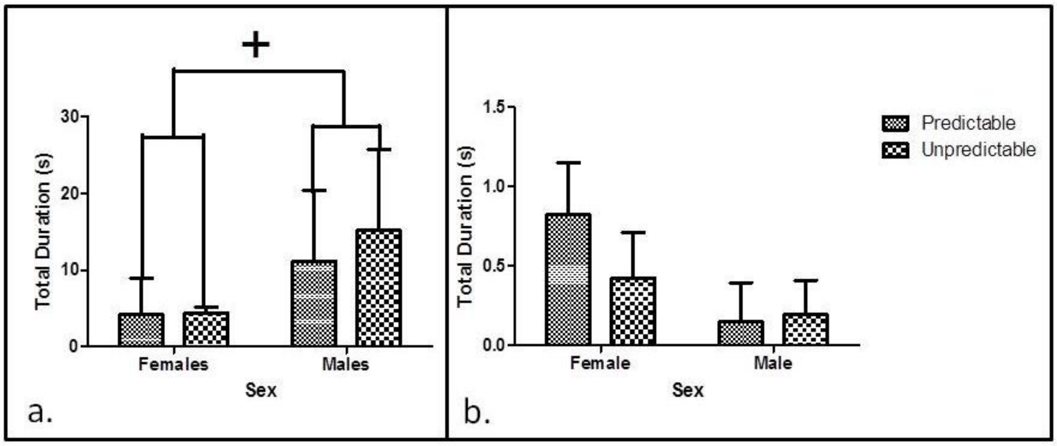

Figure 3.11 (a,b) The main effect of sex on yawn and scratch behaviors (a). The interaction between sex and cage shake and threat behaviors. The post-hoc t-tests with a Bonferroni correction applied did not reveal a significant direction for the interaction (b). Error bars represent ±1 SEM.

3.7.3.2 Modified Frequency measures and effects on the two Wait conditions and AM Data

Mixed-design ANOVAs demonstrated that there were no effects or interactions of sex and housing on stereotypic behavior (see Appendix I for results). However, there was a significant main effect of sex difference on yawn and scratch behaviors in which males yawned and scratched more than females (F(1,12)=5.66, *p*=0.035) (Figure 3.12). There was also a trend for a main effect of sex in visual explore behavior as males performed more visual exploration behavior than females (F(1,12)=4.419, *p*=0.057).

Figure 3.12 Main effect of sex on yawn and scratch behaviors. * p<0.05. Error bars represent ±1 SEM.

There were significant main effects of housing on tactile and oral exploration behaviors as well as cage shake and threat behaviors. The housing effects were that monkeys housed in Allentowns performed more cage shake and threat behaviors than monkeys housed in pens (F(1,12)=4.95, *p*=0.046) (Figure 3.13a), yet, monkeys housed in pens performed more tactile and oral exploration behaviors than monkeys housed in Allentowns (F(1,12)=4.864, *p*=0.048) (Figure 3.13b). In addition, there was a significant interaction between housing and condition for visual exploration behaviors (F(2,24)=5.704, *p*=0.009) (Figure 3.16). However, post-hoc paired samples t-tests with a Bonferroni correction applied (*p*<0.01) did not reveal a significant direction of the interaction.

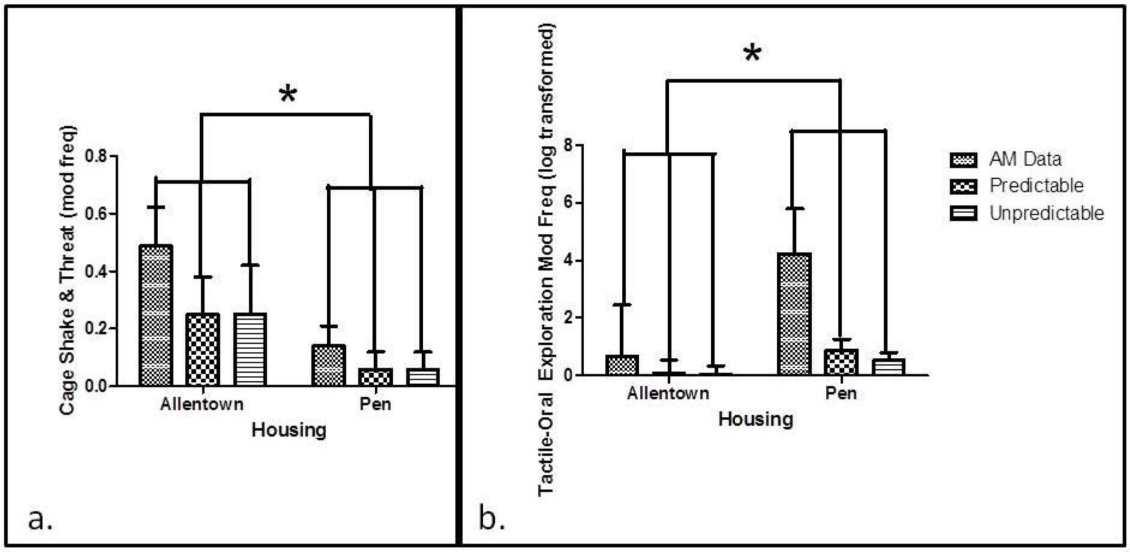

Figure 3.13 (a,b) Main effect of housing on cage shake and threat behaviors (a) and on tactile and oral exploration behaviors (b). * $p<0.05$. Error bars represent ±1 SEM.

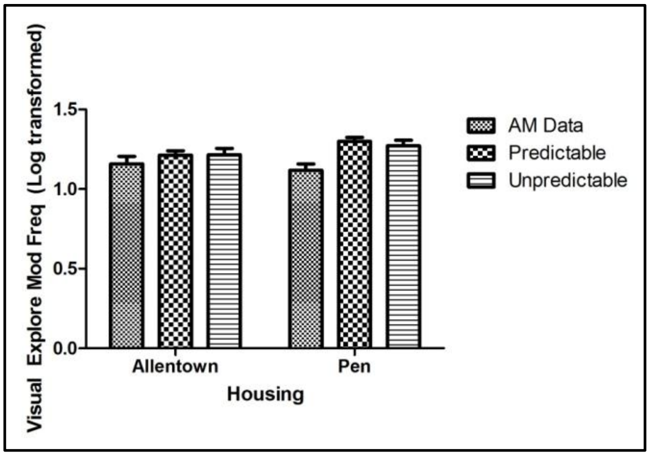

Figure 3.14 The interaction between housing and condition for visual explore behaviors. The post-hoc t-tests with a Bonferroni correction applied did not reveal a significant direction for the interaction. Error bars represent ±1 SEM.

3.8 Comparison of Stereotypic Behavior across Experiments

3.8.1 Stereotypic Behavior: Comparison of macaque stereotypic behavior in Experiments 2 & 3

Both the Wait and DRT experiments elicited stereotypic behavior. As with Experiment 1 in humans, I compared the stereotypies across these different contexts. For the monkey experiments, the Wait paradigm consisted of 5 uninterrupted minutes of waiting for the arrival of the experimenter and the DRT paradigm lasted on average 5 minutes and 54 seconds (±44s) with 9 interruptions for obtaining the treat. I corrected the total duration of stereotypic behavior for the time differences between the two experiments. Only 2/14 (14.3%) macaques (Nigel and Zoey) never displayed stereotypic behaviors across the two experiments. In the Wait Experiment, 11/14 (78.6%) monkeys displayed stereotypies and 9/14 (64.3%) monkeys displayed stereotypies in the DRT experiment.

The macaques performed significantly more stereotypic behaviors in the Wait experiment than the DRT experiment in terms of total duration (Z=2.20, p=0.028) (Figure 3.15). This result differs from Experiment 1 human results in which humans performed more behaviors in the longer paradigm, the DNMS paradigm, than in the 5-minute Wait paradigm. Furthermore, Table 3.1 presents that the monkeys had a shorter duration range and average length in the DRT paradigm as well as a smaller frequency of stereotypic behaviors performed across the group.

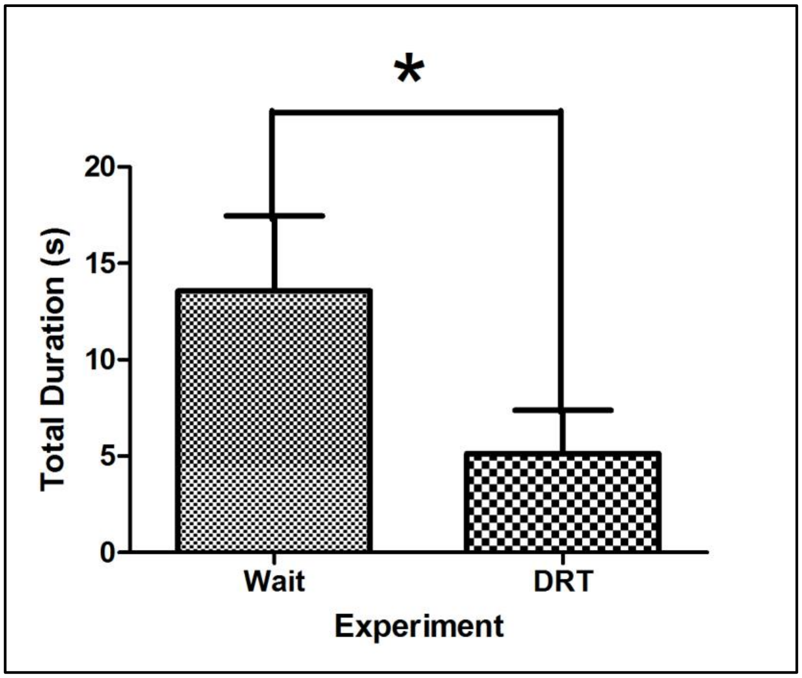

Figure 3.15 Total duration of stereotypic behavior observed in macaques in Experiments 2 (Wait) and 3 (DRT). The total duration time in the DRT was multiplied by (5/6, 0.833) in order to correct for the time difference between experiments. *p<0.05. Error bars represent ±1 SEM.

Table 3.2 Characteristics of stereotypic behavior observed in macaques in Experiments 2 and 3.

	Wait Experiment	DRT Experiment
Number of Monkeys that performed stereotypies	11	9
Average Duration Length (±1SE)	5.59 (±1.46) seconds	2.52 (±0.88) seconds
Total Duration (±1SE)	13.58 (±3.88)	6.16 (±2.70)
Frequency	92	69
Range (with minimum above zero)	0.7-60.4 seconds	0.53-32.3 seconds

The average length of stereotypies in the DRT was less than that observed in the Wait paradigm (Table 3.2). As most delays in the DRT were 15 seconds long, these results suggest that the monkeys may have shortened the length of their stereotypies because of the end of the delays. The interruptions of a delay ending and the monkey subsequently obtaining a treat affected the expression of stereotypic behavior. Importantly, this also demonstrates that the stereotypies were overall not preservative or severe enough to continue despite the end of a delay and opportunity to obtain a treat.

There were also different stereotypic behaviors expressed between experiments. Pacing was the predominant behavior observed in both experiments (Table 3.3), however, the other behaviors differed. A contrast exists between self-stroke appearing in the Wait Experiment but not in the DRT, and the converse for oral stereotypies that appeared in the DRT but not the Wait Experiment. As the monkeys consumed treats during the DRT experiment, it is possible that the treats elicited oral stereotypies not observed during the Wait Experiment. There were idiosyncratic behaviors that did not fit a behavioral category, as is also observed in human stereotypies (Table 3.3).

Table 3.3 Types of stereotypic behaviors and frequencies of these behaviors observed in macaques in Experiments 2 and 3.

Behaviors	Wait Experiment	DRT Experiment
Pace	49	36
Eye poke	25	3
Oral Stereotypies-Mouth	0	26
Self-Stroke	14	0
Oral Stereotypies-Licking	0	3
Other: Rubbing hands together	1	1
Hair Pull	1	0
Other: Tapping cage	1	0

With respect to individual consistency, I also tested whether stereotypies were correlated between experiments. I measured this for three different measures of behavior:

frequency, total duration, and average duration. I used a Spearman's rank correlation. Table 3.4 displays that each of these behavior measures were correlated with other measures from the same experiment. However, there were no significant correlations between Wait and DRT behavior measures. These results suggest that the levels of stereotypic behavior expressed by each monkey were not consistent between experiments.

Table 3.4 Spearman Rank correlation rho values for the Wait and DRT experiments. * $p<0.05$, ** $p<0.01$. Bolded values represent $p<0.05$.

	Wait Total Freq	Wait Total Duration	Wait Average Duration	DRT Total Freq	DRT Total Duration	DRT Average Duration
Wait Total Freq	1.000					
Wait Total Duration	**.950****	1.000				
Wait Average Duration	**.888****	**.925****	1.000			
DRT Total Freq	.461	.279	.150	1.000		
DRT Total Duration	.453	.291	.142	**.985****	1.000	
DRT Average Duration	.480	.323	.160	**.939****	**.959****	1.000

3.8.2 Human and Monkey Stereotypic Behavior

Experiment 3 with the rhesus macaques and the Wait paradigm of Experiment 1 with human participants both assessed how primates respond to an uninterrupted 5-minute waiting period prior to an anticipated event. I compared stereotypic behavior performance between species during this 5-minute time period using my all-occurrence sampling data from the human participants and rhesus macaques. However, the macaque stereotypic behavior included both self-directed and motor or repetitive stereotypies whereas the human stereotypic behavior was only repetitive stereotypies. In this comparison, the human behaviors without self-directed non-repetitive behavior may be an underestimate in comparison to the rhesus macaques.

Because both species did not display differences in stereotypic behavior in response to the predictable and unpredictable conditions, I combined behavior data from the within-subjects rhesus macaques. This also allowed me to statistically analyze data as I had one wait behavior score for each human and monkey participant. I log-transformed the data in order to meet the equality of variances assumption as tested with a Levene's Test. An independent samples t-test determined that the rhesus macaques performed significantly less stereotypic behavior in a 5-min sample than the humans ($t(80)=2.22$, $p=0.029$) (Figure 3.16).

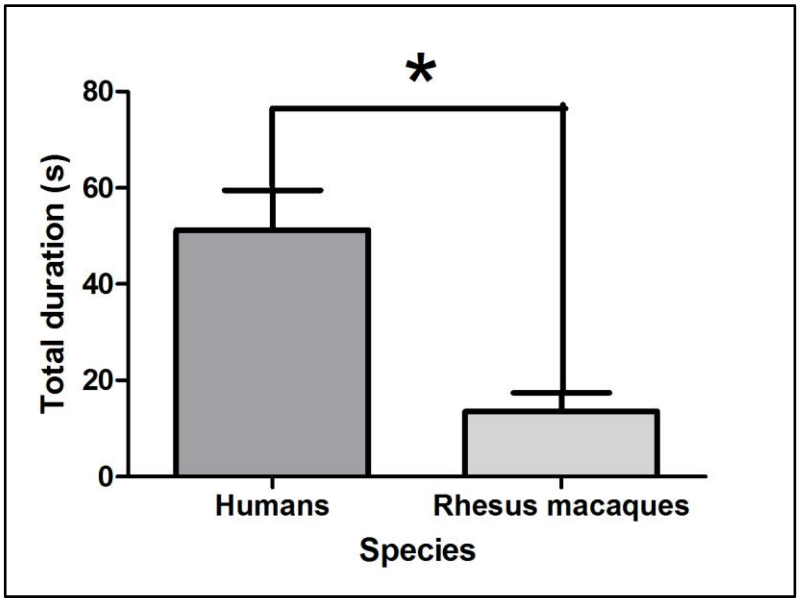

Figure 3.16 Stereotypic behavior observed in the Wait paradigm of Experiment 1 in humans and Wait experiment of Experiment 3 in rhesus macaques. Data were log transformed for analyses but are presented not-transformed in seconds.

The behaviors observed between rhesus macaques and humans were different, for while humans perform many episodes of tapping, monkeys never performed observable tapping behaviors. Yet, both species engage in behaviors utilizing objects around them while waiting as measured via object use in humans and tactile-oral exploration in rhesus macaques. However, the monkeys explored their home cage environment and humans were in a completely new room. This may be a promising new direction for future research. The discussion of these results is in Chapter 4: General Discussion.

3.9 Discussion

3.9.1 The effect of predictability on stereotypic behavior in the Wait Experiment

There were no significant differences found between stereotypic behavior performed in the predictable and unpredictable conditions, so the prediction that more stereotypic behavior would be observed in the unpredictable condition was not supported.

The other behaviors of interest of yawning, scratching, cage shake, threat, and tactile-oral exploration also did not differ by condition. I then compared behavioral responses in the Wait experiment with AM Data, a possible baseline scenario in which animals are observed with no other event following the observation to be anticipated by the monkeys. This comparison demonstrated that visual exploration behaviors significantly increased from the AM data condition to both the Predictable and Unpredictable conditions. The monkeys most likely responded to the Wait Experiment by suppressing their usual range of behaviors and instead remained vigilant as they anticipated the experimenter returning to the room to run the DRT experiment.

The behavioral responses indicate that the monkeys did not find either condition more challenging than the other. Yet, comparisons with the AM Data also revealed some differences between conditions that were likely too slight to be detected through tests of just the two conditions alone. First of all, yawn and scratch behaviors were significantly increased form the AM Data to the Unpredictable condition. As yawn and scratch behaviors can be indicative of anxiety (Schino et al., 1996), these results suggest that the monkeys may have found the Unpredictable condition to be slightly more challenging than the Predictable condition.

Secondly, locomotion significantly increased from the AM Data to the Predictable condition. Compared to the other two conditions, the Predictable condition was the one scenario in which the monkeys’ room door was open. It is possible that the macaques moved around more in this condition because they sought information from the open door. One must move around in order to view all possible angles out of an open door.

The monkeys may have moved around to exploit as much information as they could gather through the open door about where the experimenter went during the delay.

The results also demonstrated that sex and housing differences between subjects may have influenced their behavior. First of all, male macaques yawned and scratched more than female macaques (Figure 3.7). This is most likely a result primarily from yawning and this sex difference is a stable result found both in the UMass rhesus macaques and others. One of the sexually dimorphic features of male macaques is their large canine teeth. It has been hypothesized that when males are in situations that warrant displacement behaviors, males may utilize yawning more than females because it is a way to display their canines to others around them.

To my knowledge, this was the first study to measure behavior in anticipation of performing a task in captive primates. Other studies that assessed the relationship between the predictability of anticipated events and stereotypies in captive primates used husbandry events, mainly the distribution of food (Bloomsmith & Lambeth, 1995; Waitt & Buchanan-Smith, 2001, Ulyan et al., 2006; Gottlieb, et al., 2013b) although Gottlieb et al. (2013b) added additional husbandry events of cleaning and enrichment distribution. While working with a human experimenter and performing a task with an apparatus is a naturalistic scenario for many laboratory-housed rhesus macaques, the nature of the task-related scenario is different than husbandry-related ones.

It is possible that the macaques in this experiment responded differently to the anticipation of an event that offers optional participation in a task than they would respond to the anticipation of husbandry events. Alternatively, it is possible that they had a similar behavioral response to the anticipation of participating in a task as they do with husbandry events, but the macaques were not sensitive to the predictability of the delay. This may be a key difference between the optional task that is associated with a nominal treat and husbandry-related events in which the monkey anticipates an important component of its well-being such as receiving a large portion of its food, if not all of its food, for the day.

3.9.2 Conclusion

The macaques in the wait experiment largely did not respond differently to the experimental conditions. Comparisons with the AM Data demonstrated that the monkeys suppressed much of their usual behavioral repertoire and remained vigilant for the return of the human experimenter. Yet, an increase in yawning and scratching behavior between the AM Data and the Unpredictable condition suggest that the Unpredictable condition may have been a slightly more challenging experience than the Predictable. However, the macaques overall found the Predictable and Unpredictable conditions equally challenging and responded to both conditions with similar behaviors.

CHAPTER 4

GENERAL DISCUSSION

My dissertation was the first comparison of stereotypic behavior in nonclinical adult humans and captive animals, to my knowledge. Although the presence of stereotypic behavior nonclinical adult humans has been discussed by captive animal researchers (Ridley & Baker, 1982; Mason & Latham, 2004), it has never been quantified and directly compared with captive nonhuman primates using a similar experimental paradigm.

In my novel experimental manipulations of the predictability of anticipated events, I found no differences in stereotypic behavior performance in either humans or rhesus macaques. This could suggest that both species were not sensitive to changes in predictability. However, based on the results in my study, I suggest instead that predictability may be contextually difficult to manipulate. Responses other than stereotypic behavior in my experiments, such as through questionnaires in humans and other behaviors in macaques, suggest that both species responded to the experimental scenario overall rather than shifting their behavior between predictable and unpredictable conditions. Furthermore, the rhesus macaques and humans potentially had different emotional reactions to the experiments.

Humans reported feelings of boredom across both conditions. On the other hand, the rhesus macaques seemed to find both conditions equally non-boring and challenging as demonstrated through aroused and emotional responses such as yawning, scratching, cage shaking, and threat behaviors. These divergent responses to the experimental paradigms occurred despite predictions in both species that the Unpredictable conditions would be more challenging. My results suggest that the effect of predictability in the experiments were indistinguishable to both humans and monkeys and they responded instead to the overall paradigms. Future studies on predictability in humans and rhesus macaques should increase the differences between conditions in order to make the difference in conditions discernible enough in order to possibly shift behavior.

Despite the lack of response to predictability, I found that when purposeless and repetitive behaviors were quantified, nonclinical adult humans spent significantly more time performing stereotypic behavior than captive adult rhesus macaques. With much more scientific research devoted to the understanding and mitigation of stereotypic behaviors in captive animals than humans, the greater amount of stereotypic behaviors observed in adult humans was an unexpected result. As nonclinical adult humans rarely seek treatment or intervention for their levels of stereotypic behavior, the prevalence of these behaviors in people suggest that captive animal managers, based on the amount of scientific research and discussion, may overestimate the need to prevent or mitigate the performance of these behaviors in captive animals.

There are multiple possibilities for why there was an observed species difference in stereotypic behavior performance. Both the human and macaque paradigms measured stereotypic behavior in an uninterrupted 5-minute sample while the primates waited for an anticipated event. Yet, there were differences between the human and monkey experiments that may account for some of the observed differences in behavior. First of all, the monkeys were tested in their familiar home cage with their roommates present and the humans were tested alone in a novel room. The humans might have experienced an overall more challenging situation than the macaques. However, I do not propose to account for this difference by testing the macaques in a novel room by themselves, for macaques would perceive this as a highly challenging situation. My assumption is that because the humans received verbal and written information about the study they were about to participate in, the humans had more information about the overall experimental situation than the macaques would have when going into a novel room by themselves. The difference in the novelty of the room thus may have had a modest contribution to the differences in behavior between species.

Another difference between the species is in regards to possible different levels of stimulation humans and macaques usually receive from their environment. The human participants were separated from their mobile phones for the entire duration of the experiment, including the Wait paradigm. As discussed in the Methods, I had to tweak my procedure so that the participants were in a different room from their phones. When my initial group of participants was in the same room as their phones during the Wait paradigm, two out of seven of these participants went to the other side of the room to retrieve their phones to engage with while they waited for the return of the experimenter. These responses suggest that one behavioral response to a waiting scenario is to engage with one’s mobile phone in order to pass the time.

When separated from their phones, participants may have perceived less stimulation from the environment than monkeys would when in a similar situation. They may have found the Wait paradigm even less stimulating than the monkeys. On the other hand, studies demonstrated that there was an increase in anxiety when participants were separated from iphones as evidenced by both self-reported levels of anxiety (Cheever, Rosen, Carrier, & Chavez, 2014; Clayton, Leshner, & Almond, 2015) and physiological indicators of anxiety such as heart rate and blood pressure (Clayton, Leshner, & Almond, 2015). In addition, Cheever et al. (2014) analyzed these effects with respect to level of phone usage and found that anxiety increased for participants with heavy or moderate phone usage but not low phone usage. The authors of both studies proposed that there is a social component of smart phones that may drive the increase in anxiety, in which participants were anxious about being out of touch with events happening in their social circle, a psychological concept called Fear of Missing Out or *FoMO* (Przybylski, Murayama, DeHaan, & Gladwell, 2013).

If the separation from their mobile phones influenced the participants' behavior in my experiment, then this is a highly concerning effect given that the practice for most psychology studies is to have the participants not engage with their phone while in the middle of an experiment. Especially if this effect differs with levels of phone usage, which may not be held constant between experimental conditions in a typical psychology study. Future studies of human behavior, physiology, and cognition should implement such scales as the FoMo scale (Przybylski et al., 2013) or query participants on the level of usage or dependence on their smart phones in order to account for the effect of phone separation in behavior, physiology, or cognitive performance during experiments. While I discuss factors that were different between the monkey and human experiments, there was no easy way to adjust or account for these factors and make the experimental scenarios more equivalent between the species in my study. The results of my comparative study may provide information for what factors to focus on and how to adjust for these factors in future studies.

One hypothesis about stereotypic behavior is that these behaviors serve a psychological function in response to the current environment (Mason & Latham, 2004). My results suggest that stereotypic behavior may serve a psychological function in response to the current environment and this function is retained between two species of primates. The experimental paradigms were designed to be as similar as possible between the human and macaque experiments. Furthermore, a similar amount of stereotypic behaviors were observed. While I cannot conclude that the humans and macaques experienced similar emotions during the experiment, it seems plausible that both species used stereotypic behavior as part of their response to the current environment.

Additionally, the ability to administer questionnaires and established scales to human participants provided valuable information on how emotional states and individual traits may relate to the performance of stereotypic behavior. Participants who reported being bored performed more stereotypies. As for who felt more bored, it was participants who found a waiting scenario more challenging, possibly because of a general intolerance of uncertain conditions.

As a response to the current environment, stereotypic behaviors may serve a psychological function as a compensatory response to certain environmental conditions. Provided that these behaviors may be found in both under or over-stimulating environments, stereotypic behavior may serve a self-regulatory function. In this respect, an individual has a lower baseline frequency of these behaviors when the stimulation in the environment is at a level that allows the individual to maintain homeostasis. When stimulation in the environment rises above or drops below this level, stereotypic behaviors increase in order to modulate stimulation that the individual receives and then returns the individual to homeostasis. The idea of a self-regulatory function of stereotypic behavior has been discussed before (Mason & Latham, 2004). However, the link to the emotional state of boredom in humans as found in my dissertation is one of the strongest links to date of this potential function of stereotypic behavior.

From a psychological perspective, the underlying mental processes that occur during boredom are not well understood. There are different approaches to explaining boredom, from existential ones, arousal theories, and cognitive perspectives. Overall, researchers conclude that boredom is an aversive state: when someone feels bored, they want to not feel bored anymore (Eastwood et al., 2012). One study even found that people may be more prosocial and sacrifice resources in order to end boredom, for people who were in a *high boredom* condition felt more willing to give to charity than those in a less bored condition (van Tilburg & Igou, 2017).

It is possible to extend this sentiment of sacrificing resources to end boredom as a similar one found when people purchase items that elicit fidgeting and stereotypic behavior. People buy objects such as *Fidget Cubes*® (Figure 4.1) (McLachlan & McLachlan, 2016). The Fidget Cube is currently the 10th most funded project on a crowd sourcing website called Kickstarter.com (Kickstarter.com: Most Funded, 2017) with $6,465,690 in profits when the project only aimed to generate $15,000 for production costs. These data suggest that people were willing to sacrifice monetary resources in order to perform stereotypic behavior. This is a similar theme as van Tilburg and Igou’s (2017) findings that people may be willing to sacrifice $20 in order to alleviate boredom.

Figure 4.1 Image of Fidget Cube from its website.

Future studies can further develop on the result of boredom and stereotypic behavior in a number of ways. First of all, boredom may not be necessary for the development of stereotypic behavior. Perhaps animals have a general trait-level amount of stereotypic behavior that they perform across multiple contexts. In this light, I did find a consistency in stereotypic behavior performance between the Wait and DNMS paradigm in humans, although the results were less similar for monkeys. Different contexts, such as more stressful situations, should be tested. Secondly, the relationship between boredom and stereotypic behaviors should be further explored. It is unclear whether boredom causes stereotypic behavior, or if stereotypic behavior and boredom co-occur but have no effect on each other, or importantly whether stereotypies are specifically implemented to reduce boredom.

Both my experiment and the Fidget Cube demonstrate a possible universality in stereotypic behavior propensities in humans. However, these findings are limited to humans in industrialized and Western cultures. While not specifically studied, to my knowledge, these behaviors have been recorded as occurring in other cultures through anthropological accounts. Table 4.1 outlines when fidgeting, pacing, or other nervous habits were mentioned as occurring in adults in other cultures that span both the globe and different points in history. The use of labels such as *fidgeting*, however, does not constitute agreement across writers as to what behaviors constitute fidgeting. Secondly, these entries do not imply that the behavior of fidgeting was recognized as such by members of non-Western cultures. It is possible that fidgeting is a Western concept. While the behaviors appear to occur in humans across cultures, future studies can further explore how other cultures view stereotypic behaviors and what factors may elicit these behaviors.

Table 4.1 Examples of anthropological accounts of stereotypic behaviors across cultures.

Cultural Group	Location	Behavior Mentioned	Reference
Kwoma	New Guinea	Fidgeting	Whiting, 1970
Buddhist Monks	Thailand	Fidgeting	Terwiel, 1975
Hopi	Arizona	Fidgeting	Titiev, 1944
Akan	Ghana	Fidgeting	Field, 1970
Iroquois	New York	Nervous Habits	Fenton, 1953
Canela	Eastern South America	Fidgeting and pacing	Nimuendaiu & Lowie, 1946

In addition, the perception of and calibration of time is highly cultural. Event time, clock time, and expectations of promptness may not be as tightly regulated in other cultures. The 5-minute Wait time for the arrival of another experimenter, as seen in my study, could be an innocuous or expected event in other cultures. Time as a psychological and environmental concept depends on social, economic, environmental, and cultural factors (Levine, 1997). Ultimately, while fidgeting may be observed across cultures, other approaches beyond the experiment utilized in my dissertation should be considered in order to test these phenomenon in participants from other cultures.

Potential differences in perception also arise when I compare my results between human participants and rhesus macaques. In the DRT, rhesus macaques received an immediate small food reward on every trial whereas the humans in the DNMS paradigm received immediate visual and auditory feedback but their tangible reward was not provided to them until the cessation of the experiment. The differences in the rewards and the timing of reward distribution may have generated the divergent responses to the experiments between humans and monkeys. In my study, humans may have reported feeling less bored if they were waiting for food rewards. Comparative studies like mine attempt to create equally meaningful rewards in both species. However, differences inherent in the species and logistical considerations may make it difficult to provide similar rewards across species and thus affect our abilities to compare behaviors even in similar experimental scenarios.

4.1 Complexities of Stereotypic Behaviors in Captive Animals

The traditional perception of captive animal management was that behaviors that were performed more often in captivity than in the theoretical wild were abnormal pathological aberrations. Furthermore, these behaviors indicated that an animal was placed in a suboptimal environment (Mason & Latham, 2004). Environmental enrichment, or the human implementation of physical or social complexity in a captive animal’s environment, has one of its primary goals as promoting *species-typical* behaviors and reducing stereotypies and other *abnormal* behaviors (Lutz & Novak, 2005). In this respect, the long-held idea for promoting animal welfare was to, as nearly as possible, replicate naturalistic conditions in the captive environment so that captive animals can perform behaviors observed in their wild counterparts instead of behaviors observed more often in captive counterparts.

When quantitatively measured, the relationship between abnormal behaviors and welfare has been unclear. For example, Mason and Latham (2004) conducted a literature review on the relationship between stereotypies and welfare and found that 153 studies linked stereotypic behavior with poor welfare, yet 133 studies did not find this link, and some were even associated with good welfare. Secondly, coprophagy in chimpanzees, usually considered an abnormal and undesirable behavior, loaded onto a factor with positive social behaviors rather than the abnormal factor in a principal components analysis of 60 zoo or sanctuary-housed chimpanzees (Hopper, Freeman, & Ross, 2016).

Additionally, emerging evidence suggests that there may be subtypes of stereotypic behavior. For example, 22 horses that had a history of performing an oral stereotypy called crib-biting were exposed to an ACTH challenge test (Freymond et al., 2015). When challenged, 15 of these horses displayed crib-biting behavior and 7 did not. These crib-biting horses and a control group of horses did not have a significant difference in baseline cortisol prior to the challenge. Yet, the horses that had a history of crib-biting but did not display crib-biting during the test had a significantly higher increase in cortisol during the challenge than the control group. Phrased another way, the horses that had a history of crib-biting and expressed this behavior during the test had statistically equivalent levels of cortisol as control horses (Freymond et al., 2015). Similar results were found in laboratory-housed rhesus macaques. Adult monkeys that had a history of pacing but suppressed these behaviors during a challenging scenario of an unknown human intruder, had significantly higher hair cortisol levels than monkeys that had a history of pacing and continued to pace in response to a stranger (Novak et al., 2015).

Finally, in a study of mink housed at a fur farm, minks reared in enriched environments had overall lower levels of stereotypic behavior than minks reared in nonenriched environments (Díez-León et al., 2016). The differences were reflected in scrabbling (resembling scratching or digging the cage floor) but locomotor behavior such as pacing were not different between enriched and non-enriched minks. These studies (Freymond et al., 2015; Novak et al., 2015; Díez-León et al., 2016) suggest that there may be different functions of different behaviors that we currently categorize together as stereotypies. Furthermore, a similar behavior such as pacing and crib-biting may serve different functions in different animals. My dissertation may provide support for a self-regulatory function or at least a response to the current environment, but this does not imply that there are other possible functions depending on the type of behavior or that there are individual differences between animals.

4.2 Conclusion

In my dissertation, I found that stereotypic behaviors are performed in similar frequencies in nonclinical adult humans and a group of captive rhesus macaques. There is an extensive literature on understanding captive animal stereotypies and comparably less scientific research on understanding these behaviors in nonclinical adult humans. The potential for a biological continuity of these behaviors between humans and captive primates can create opportunities for a cross-translational model. Stereotypies can be studied in humans using techniques such as scales that are available in humans and not in captive animals in order to investigate potential functions of stereotypies in novel and revealing ways.

As quantitative analyses of these behaviors suggest that we question the assumed link between stereotypies and poor welfare, captive animal researchers are beginning to shift away from the idea that behaviors observed more in captive animals than wild counterparts are pathological behaviors that should be mitigated. Results from my dissertation can contribute to this evolution in thinking in captive animal management. A culmination of this thinking is outlined in a position statement by the American Society of Primatologists, Association of Primate Veterinarians, and the American College of Laboratory Animal Medicine that policy for animal welfare as reflected in the Animal Welfare Act should be directed towards “functionally appropriate nonhuman primate environments” instead of “ethologically appropriate environments” (Bloomsmith, Hasenau, & Bohm, 2017).

Instead of developing a pathology, captive animals may use stereotypic behaviors as a constructive and compensatory response to the captive environment. In this light, the behavioral repertoire of a captive animal may look different than that of a wild counterpart, but captive animals are indeed living in different environments and are responding to the environmental challenges they face. The captive environment, in turn, should be functional rather than attempt to be naturalistic (Bloomsmith, et al., 2017). A functional captive environment and management program may include tracking the occurrence of but not necessarily seeking to eliminate stereotypic behaviors.

APPENDIX A

MATERIALS FOR RECRUITMENT, CONSENT, DEBRIEFING FORM, AND QUESTIONNAIRES FOR PARTICIPANTS IN EXPERIMENT 1

A.1 Email to qualified prescreen participants

Greetings!
Thank you for completing the long version of the Psychology SONA prescreening questionnaire. Based on your data, you are eligible to participate in a study called Search Strategies in a Visual Discrimination Task. This is an hour-long study in which you can earn 2 credits and also raffle tickets for a giftcard! There is no advance preparation needed. You will complete a visual discrimination task while being videotaped, then, you will be asked to complete a few questionnaires about yourself.

You are required to have 20/20 or corrected-to-normal vision and not be color blind in order to participate in this study.

We will be holding a limited number of sessions this week. If you are interested participating in this study, please sign into SONA at https://umasspsych.sona-systems.com and select “Search Strategies in a Visual Discrimination Task.”

This study requires an invitation code in order to participate. Your special code when signing up on SONA is TOBINROCKS29.

I thank you for your time and hope that you decide to participate in this study. Please let me know if you have any questions! You can reply directly to this email address or contact me directly at amyr@cns.umass.edu.

Best regards,
Amy M. Ryan, M.A.
Department of Psychology
University of Massachusetts, Amherst
135 Hicks Way, Tobin Hall 512
Amherst, MA 01003
amyr@cns.umass.edu

A.2 Consent Form

CONSENT FORM

Study name: Search Strategies in a Visual Discrimination Task (Study# 2016-2927)

Investigator: Amy Ryan, Doctoral Student, 908-347-7857

By signing this consent form you, ______________________________________ indicate that you willingly agree to participate in this project. The essence of this project is as follows:

Purpose of the Research and Procedures

The purpose of this research is to examine how visual search strategies are used in a visual discrimination task. We will ask you to complete a visual discrimination task and we will videotape you while you are engaged in the task. Following completion of the visual discrimination task, we will ask you to complete a few questionnaires about yourself. Each session will take place in Tobin 652 or 656, will last 1 hour, and will be scheduled at a time that is convenient for you. During the session, you will be asked to participate in a couple separate tasks. First, you will be seated at a computer in order to complete the visual discrimination test. The test is a delayed non-matching to sample task, which means that you will view one image on the screen, the image then disappears, and after a delay, two test images will appear—one is the previously presented image and one is a novel image. The test is whether you can recall the previously presented image and identify and select the novel image. The computer will record your selection and provide feedback. We will be using the video recording to assess how long it takes you to select an image as well as how you scan the images while making your decision. After this, you will complete a series of questionnaires about your attitudes and behavior. Next, we will ask you to complete a demographic questionnaire where we will ask you questions about your age, year in college, etc. Finally, we will ask you a few questions about the study and then you will be debriefed.

Please note: Videotaping is a required part of the study; if you do not wish to be videotaped then you are not eligible to participate in this study. You will have several options as to how your videotapes can be used. These options will be fully explained in a separate videotaping consent form, which you will be presented with next.

Benefits

You may not directly benefit from this research; however, we hope that your participation in this study may offer some insight into your visual discrimination abilities. The experimenter will also be willing to discuss the study with you at the end of the session.

Risks and Discomfort

There is a small risk that you might feel some discomfort during the computer task or when answering the questions, but you may stop participating at any time or you may refuse to answer any question. You may choose to participate or not. You may answer only the questions you feel comfortable answering, and you may stop at any time. Although we hope that you will fully participate in this study, please understand that your participation is entirely voluntary and that you have the right to withdraw consent or discontinue participation at any time without penalty.

Academic Credit and Compensation

You will receive one credit for every 30 minutes of participation rounded to the nearest half hour (2 credits for an hour long study). If you should decide to discontinue your participation you will be credited for the time you have participated. The credits can be applied toward any psychology class that accepts human subjects extra credits. If you are earning experimental credits through your participation, please understand that participating in this study is not the only way to earn credit. You may contact your instructor who will offer you an appropriate alternative activity.

In addition, for every correct answer in the visual discrimination task, you will earn 1 raffle ticket for a $50 Amazon gift card. The raffle will be drawn at the conclusion of the study.

The University of Massachusetts does not have a program for compensating subjects for injury or complications related to human subjects research but the study personnel will assist you in getting treatment.

Length of the Study

The experimental session will take 1 hour, for which you will receive 2 SONA credits.

Confidentiality

Your identity as a participant in this research project will be kept confidential to the fullest extent possible. Although we ask for your name on this informed consent, any information that you give us is confidential. Your name will not be associated with your data. At the beginning of the study, you will be assigned a number code that will be written on all your response materials. By identifying your responses with a number code, your name will never be associated with your responses. The link between your name and the number code will be kept in a separate locked location, and none of the study data will have your name on it. Only the research team will have access to the dataset in most cases. If these data are ever shared with researchers beyond our team, no identifying information will ever be provided.

Given that the experiment will be videotaped, however, you will be asked to complete a separate consent form (Consent Form for Videotaping of the Experimental Session) for the videotaping, in which you will be able to set restrictions on how your videotape data are used.

Request for Additional Information

You may ask more questions about this research at any time. If you have questions about this project or if you have a research-related problem, you may contact the researcher, Amy Ryan (908-347-7857; amyr@cns.umass.edu), or Professor Brian Lickel (413-577-0493; blickel@psych.umass.edu) to answer your questions and concerns now and after your participation in this research if you agree to do so. You may also contact either Amy Ryan or Dr. Brian Lickel by postal mail (Department of Psychology, Tobin Hall, University of Massachusetts, Amherst, MA 01003). If you would like to speak with someone not directly involved in the research study, you may contact Dr. Harold Grotevant, Chair of Psychology (hgroteva@psych.umass.edu, 413-577-0837) or the Human Research Protection Office (HRPO) at the University of Massachusetts via e-mail (humansubjects@ora.umass.edu); telephone (413-545-3428); or mail (Human Research Protection Office, Research Administration Building, University of Massachusetts Amherst, 70 Butterfield Terrace, Amherst, MA 01003-9242).

Participant’s statement: When signing this form I am agreeing to voluntarily enter this study. I have had a chance to read this consent form, and it was explained to me in a language which I use and understand. I have had the opportunity to ask questions and have received satisfactory answers. A copy of this signed Consent Form has been given to me.

If you have any questions about anything, please do not hesitate to ask the experimenter. Also, if you have any questions concerning what you will be asked to do during this experiment, please ask the experimenter **before** signing this consent form.

If you are willing to participate in this study, please sign your name below.

Signed Name: ____________________________________

Printed Name: ____________________________________

Date: _____________________

Student I.D. Number: _______________________________

STUDY REPRESENTATIVE STATEMENT:

I have explained the purpose of the research, the study procedures, the possible risks and discomforts, the possible benefits, and have answered any questions to the best of my ability.

____________________________________ _____________________

Amy Ryan, Principal Investigator Date

A.2.1 Video Consent Form

CONSENT FORM FOR VIDEO RECORDING OF THE EXPERIMENTAL SESSION
Study name: Strategies in a Visual Discrimination Task (Study# 2016-2927)

Investigator: Amy Ryan, Doctoral Student, 908-347-7857

This consent form is to obtain your permission to videotape your behavior during a computer task and to use the videotape for research purpose specified below. If you do not wish to be videotaped, you are not eligible to participate in this study.

Consent for the Experimental Session to be Videotaped

For the purposes of understanding what visual search strategies are used in a visual discrimination task, you will be videotaped. Your behavior will be coded and analyzed by trained members of the research team. Please initial next to one of the options below to indicate your preferences:

_______ I consent for my experimental session to be videotaped for research purposes.

_______ I <u>do not</u> consent for my experimental session to be videotaped for research purposes. (If you do not wish to be videotaped, you are not eligible for this study. Thank you for your time.)

Confidentiality

We will only use your videotaped material for the purposes you have consented to; otherwise it will be deleted. Other than your image and information provided in your responses, we will keep all other information about you entirely confidential. Your name and other personal information such as your questionnaire answers will not be given to anybody looking at the videotapes. When the results of this study are presented or published, no identifying information about you will be revealed. Your data will be combined with other participants and presented as an anonymous group.

Storage of the Videotaped Materials

Your videotaped material will be stored electronically in password-protected files on the researcher’s computer, a laboratory computer, and on a back-up hard drive. During data collection, videos on memory cards will be stored in a locked filing cabinet only accessible by the research team. Once data collection is complete, only electronic copies of the videotapes will be retained. Regarding the archiving and retention of your videotape, please indicate whether you give permission to archive and retain your video indefinitely, or if you prefer for your videotape to be destroyed after 7 years post-publication. Please initial next to your preferred videotape archiving and retention option:

_______ My videotaped materials may be archived and retained <u>indefinitely</u> for the purposes I have agreed to above and for secondary analyses. If secondary analyses are performed, they will be approved by the Institutional Review Board to ensure they are consistent with the original purpose of the study I consented to participate in.

_______ My videotaped materials may be retained for <u>up to 7 years</u> post-publication for the purposes I have agreed to above and for secondary analyses. If secondary analyses are performed, they will be approved by the Institutional Review Board to ensure they are consistent with the original purpose of the study I consented to participate in. After 7 years, all copies of the videotapes will be destroyed.

Request for Additional Information

You may ask more questions about this research at any time. If you have questions about this project or if you have a research-related problem, you may contact the researcher, Amy Ryan (908-347-7857; amyr@cns.umass.edu), or Professor Brian Lickel (413-577-0493; blickel@psych.umass.edu) to answer your questions and concerns now and after your participation in this research if you agree to do so. You may also contact either Amy Ryan or Dr. Brian Lickel by postal mail (Department of Psychology, Tobin Hall, University of Massachusetts, Amherst, MA 01003). If you would like to speak with someone not directly involved in the research study, you may contact Dr. Harold Grotevant, Chair of Psychology (hgroteva@psych.umass.edu, 413-577-0837) or the Human Research Protection Office (HRPO) at the University of Massachusetts via e-mail (humansubjects@ora.umass.edu); telephone (413-545-3428); or mail (Human Research Protection Office, Research Administration Building, University of Massachusetts Amherst, 70 Butterfield Terrace, Amherst, MA 01003-9242).

Participant's statement: When signing this form I am agreeing to allow the researchers of this project to use my videotaped experimental session for experimental purposes. I understand that I can change my decision about allowing the use of this video at any time and the researchers agree not to continue using the video. I have had a chance to read this consent form, and it was explained to me in a language that I understand. I have had the opportunity to ask questions and have received satisfactory answers. A copy of this signed Consent Form has been given to me.

If you have any questions about anything, please do not hesitate to ask the experimenter. Also, if you have any questions concerning what you will be asked to do during this experiment, please ask the experimenter **before** signing this consent form.

Please sign your name below to affirm that you consent to the videotaping of your experimental session for the purposes you have indicated above:

Signed Name: ______________________________________

Printed Name: ______________________________________

Date: ______________________

Student I.D. Number: _________________________________

STUDY REPRESENTATIVE STATEMENT:

I have explained the purpose of the research, the study procedures, the possible risks and discomforts, the possible benefits, and have answered any questions to the best of my ability.

______________________________________ ______________________

Amy Ryan, Principal Investigator Date

A.3 Debriefing Form

DEBRIEFING FORM

Study name: Search Strategies in a Visual Discrimination Task (Study# 2016-2927)

Thank you for participating in this research. Your participation is greatly appreciated.

Purpose of the study:

Earlier in our consent form we told you that the study purpose was to examine how visual search strategies are used in a visual discrimination task. **In actuality, the purpose of our study is to examine whether events that are predictable or unpredictable differentially induce stereotypic behaviors. People engage in behaviors called stereotypies, which are repetitive motions that serve no apparent purpose, typically considered either fidgeting or nervous habits. Examples include hair and face manipulation with hands, putting parts of body or objects in mouth, flexion–extension of legs, tapping of limbs against a surface or one's own body, repetitive object manipulation, and rocking.**

While we originally told you that another experimenter will entering the room, we will actually use your video recording to measure your frequencies of stereotypic behavior while you waited for the anticipated event. Additionally, while we originally told you that we were measuring your performance on the delayed non-matching to sample task and visual search strategies, we will actually use your video recording to measure your frequencies of stereotypic behavior during the delay between the original image and test images. For both of these manipulations, you received one of two possible conditions: predictable or unpredictable wait time or delay. In the predictable condition, the time of the wait for the next experimenter was indicated, whereas in the unpredictable condition, the time of the wait was not specifically indicated. In the computer task, in the predictable condition, the time of the delay was indicated, whereas in the unpredictable condition, the time of delay was not specifically indicated. Finally, you will have one raffle ticket for the $50 Amazon gift card as opposed to a number based on your performance.

In order to properly test our hypothesis, we could not provide you with all of these details prior to your participation. This ensures that your reactions in this study were spontaneous and not influenced by prior knowledge about the purpose of the study. We had you participate in a task in which we could manipulate the predictability of an anticipated event. There was no second experimenter involved in this study, and the visual discrimination test was in this light a fabricated research activity. If we had told you the actual purpose of our study, your stereotypic behavior could have been affected once you became aware of our measurements of this typically innocuous activity. We regret the deception but we hope you understand the reason for it.

Confidentiality:

Please note that although the purpose of this study has changed from the originally stated purpose, everything else on the consent form is correct. This includes the ways in which we will keep your data confidential. Your name will not be associated with your data and your videotape recording will be kept in a locked location.

Do you allow us to use your data now that you understand the true purpose of this research? [*If no, we will withdraw and delete their data*.] If you select, 'no', we will withdraw your data. If you choose to withdraw your data, we will delete your data from the master data file by the end of the business day. Once you leave the lab, your data will no longer be associated with your identity, so we will be unable to withdraw your data after you leave. You will still be credited for your participation if you choose to withdraw your data, but we do hope that you will permit us to use your data in this research study.

[ ] **Yes, you may use my data**

[ ] No, you may **NOT** use my data

Please do not share the true purpose of this study with anyone else as data collection is ongoing.

Final Report:

If you would like to receive a copy of the final report of this study (or a summary of the findings) when it is completed, please feel free to contact us.

*** If you find that you are distressed by any part of study, you may contact the Center for Counseling and Psychological Health (CCPH) at UMass Amherst's University Health Services at UMass (545-2337) or the Psychological Service Center at UMass (545-0041) for counseling.

You may ask more questions about this research at any time. If you have questions about this project or if you have a research-related problem, you may contact the researcher, Amy Ryan (908-347-7857; amyr@cns.umass.edu), or Professor Brian Lickel (413-577-0493; blickel@psych.umass.edu) to answer your questions and concerns now and after your participation in this research if you agree to do so. You may also contact either Amy Ryan or Dr. Brian Lickel by postal mail (Department of Psychology, Tobin Hall, University of Massachusetts, Amherst, MA 01003). If you would like to speak with someone not directly involved in the research study, you may contact Dr. Harold Grotevant, Chair of Psychology (hgroteva@psych.umass.edu, 413-577-0837) or the Human Research Protection Office (HRPO) at the University of Massachusetts via e-mail (humansubjects@ora.umass.edu); telephone (413-545-3428); or mail (Human Research Protection Office, Research Administration Building, University of Massachusetts Amherst, 70 Butterfield Terrace, Amherst, MA 01003-9242).

Signed Name: ____________________________________

Printed Name: ____________________________________

Date: _____________________

Student I.D. Number: _______________________________

A.4 Questionnaires and Scales

The questionnaires were presented in Qualtrics® software. Below is the text for the questionnaires, but the formatting was consistent throughout the questionnaire when presented to participants.

A.4.1 Spring 2016 Prescreen Questionnaires
A.4.1.1 Ten-Item Personality Inventory (TIPI) (Gosling, Rentfrow, & Swann, 2003).

Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

Disagree strongly	Disagree moderately	Disagree a little	Neither agree or disagree	Agree a little	Agree moderately	Agree strongly
1	2	3	4	5	6	7

I see myself as:
1. _____ Extraverted, enthusiastic.
2. _____ Critical, quarrelsome.
3. _____ Dependable, self-disciplined.
4. _____ Anxious, easily upset.
5. _____ Open to new experiences, complex.
6. _____ Reserved, quiet.
7. _____ Sympathetic, warm.
8. _____ Disorganized, careless.
9. _____ Calm, emotionally stable.
10. ____ Conventional, creative.

A.4.1.2 5 questions from Shortened Intolerance of Uncertainty Scale (Carleton, Norton, & Asmundson, 2007)

Intolerance of Uncertainty Scale

Not at all characteristic of me	Somewhat uncharacteristic of me	Neither characteristic nor characteristic of me	Somewhat characteristic of me	Entirely characteristic of me
1	2	3	4	5

1. Unforeseen events upset me greatly.
2. It frustrates me not having all the information I need.
3. I always want to know what the future has in store for me.
4. When I am uncertain I can't function very well.
5. I must get away from all uncertain situations.

A.4.2 Questionnaires Immediately following Wait and DNMS Behavioral Paradigms

A.4.2.1 PANAS-X (Watson and Clark, 1994)

This scale consists of a number of words and phrases that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you currently feel this way. Use the following scale to record your answers:

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely

______ cheerful ______ sad ______ active ______ angry at self

______ disgusted ______ calm ______ guilty ______ enthusiastic

______ attentive ______ afraid ______ joyful ______ downhearted

______ bashful ______ tired ______ nervous ______ sheepish

______ sluggish ______ amazed ______ lonely ______ distressed

______ daring ______ shaky ______ sleepy ______ blameworthy

______ surprised ______ happy ______ excited ______ determined

______ strong ______ timid ______ hostile ______ frightened

______ scornful ______ alone ______ proud ______ astonished

______ relaxed ______ alert ______ jittery ______ interested

______ irritable ______ upset ______ lively ______ loathing

______ delighted ______ angry ______ ashamed ______ confident

______ inspired ______ bold ______ at ease ______ energetic

______ fearless ______ blue ______ scared ______ concentrating

______ disgusted with self ______ shy ______ drowsy ______ dissatisfied with self

ADDED: Bored; frustrated; pressed for time; anxious; stressed.

A.4.2.2 Multidimensional State Boredom Scale (MSBS): Fahlman, et al. 2013

Instructions. Please respond to each question indicating how you feel right now about yourself and your life, even if it is different from how you usually feel.
Use the following choices:
1 = Strongly disagree;
2 = Disagree;
3 = Somewhat disagree;
4 = Neutral;
5 = Somewhat agree;
6 =Agree; and
7 = Strongly agree.

1. Time is passing by slower than usual.
2. I am stuck in a situation that I feel is irrelevant.
3. I am easily distracted.
4. I am lonely.
5. Everything seems to be irritating me right now.
6. I wish time would go by faster.
7. Everything seems repetitive and routine to me.
8. I feel down.
9. I seem to be forced to do things that have no value to me.
10. I feel bored.
11. Time is dragging on.
12. I am more moody than usual.
13. I am indecisive or unsure of what to do next.
14. I feel agitated.
15. I feel empty.

16. It is difficult to focus my attention.
17. I want to do something fun, but nothing appeals to me.
18. Time is moving very slowly.
19. I wish I was doing something more exciting.
20. My attention span is shorter than usual.
21. I am impatient right now.
22. I am wasting time that would be better spent on something else.
23. My mind is wandering.
24. I want something to happen but I'm not sure what.
25. I feel cut off from the rest of the world.
26. Right now it seems like time is passing slowly.
27. I am annoyed with the people around me.
28. I feel like I'm sitting around waiting for something to happen.
29. It seems like there's no one around for me to talk to.

A.4.2.3 Stereotypic Behavior and severity (Mehrabian and Friedman, 1986)

9-Point Scale:

1	2	3	4	5	6	7	8	9
Very strong Agreement	Strong Agreement	Moderate Agreement	Slight Agreement	Neither Agree nor Disagree	Slight Disagreement	Moderate Disagreement	Strong Disagreement	Very Strong Disagreement

1. I frequently rub my neck
2. I hardly ever pinch my cheeks
3. I usually have something in my hands to play with
4. I hardly ever close my eyes tight and then open them)
5. I never make clucking or smacking noises with my mouth
6. I don't scrunch my shoulders
7. I often bend paper cups or aluminum cans after I drink their contents
8. I hardly ever blow up or puff out my cheeks
9. I hardly ever press my hands or fingers against each other
10. I scratch myself a lot
11. I usually jiggle my pen when I am holding it but not writing with it
12. When seated, I don't move around restlessly in my seat
13. I often stretch out my arms
14. I often bite my lip
15. I rarely rub my legs
16. I frequently lace my fingers together
17. I hardly ever suck on my tongue

18. When sitting someplace where my feet don’t reach the floor, I often swing my legs back and forth
19. I hardly ever rub my scalp
20. I don’t put nonedible objects in my mouth
21. I often rip up things such as napkins, wrappers, etc. into little pieces.
22. I don’t move my torso around when seated
23. I frequently rub my forehead and the areas around my eyes
24. When I have a hangnail or healing cut, I often play with it and make it worse
25. I usually bend or play with a straw when drinking through it
26. I don’t suck in my lips or cheeks
27. I tap my foot a lot
28. I don’t fondle or play with my clothes
29. I often click my teeth
30. I don’t rub my own arms or shoulders
31. I hardly ever move my fingers around just for stimulation
32. I have a lot of restless movements
33. I don’t unbend paperclips
34. I frequently roll my tongue around in my mouth
35. I don’t play with my watch once I put it on
36. I don’t scratch my head
37. When standing, I often shift my weight from one leg to another
38. I frequently bite the inside of my cheek
39. I don’t tap or drum on things
40. I rub my fingers and/or hands together a lot

Added:
I frequently pull or twist my hair
I hardly ever touch my face
I hardly ever rock my body or torso back and forth or side-to-side when seated
I don’t pace a lot while waiting for something to happen
I often bite my nails
I frequently tap my fingers
I hardly ever crack my knuckles or fingers
I don’t click or grind my teeth

A.4.2.3.1 Follow-up questions about stereotypic behavior not from Mehrabian and Friedman’s Scale

1. a. Can you recall performing any of these behaviors during the experiment?
A. Yes B. I probably did, but can’t recall C. I most likely didn’t, but can’t recall D. No- E. Unsure

b. If Yes, please describe what behaviors you performed.

2. Which of these habits do you consider your overall most frequent habit? Please indicate your top 3 habits.

3. For your most frequent habit, please use the following scale to rate how much of a problem this habit has been for you in the past 30 days?

None—Mild—Moderate—Severe—Very Severe

4. Have you attempted to stop your most frequent habit? YES/NO
If Yes, Degree of success:
0 (not successful) 1 2 (moderately successful) 3 4 (highly successful)

A.4.2.4 Mental Health Check

1. Approximately how many hours of sleep did you receive last night?
0-2 3-5 6 7 8 9 10-12

2. Approximately how much caffeine have you consumed within the last 24 hours?
None 1 cup of coffee 2-3 cups of coffee 1 energy drink (Red Bull, 5-hour energy) ≥4 cups of coffee

3. Do you have a tic (for example: eye blinking, grunting, shoulder shrugging, throat clearing, sniffing) that you are aware of?

Yes/No Diagnosis Questions:
4. Have you been diagnosed with Aspergers or Autism Spectrum Disorder?
5. Have you been diagnosed with ADHD (Attention-deficit/hyperactivity disorder) or ADD (attention deficit disorder)?
6. Have you taken medicine typically prescribed for ADHD or ADD in the past 6 months [for example, Ritalin, Adderall, Concerta, Focalin]?
6. Have you been diagnosed with a tic disorder or Tourette's disorder?
7. Have you been diagnosed with Stereotypic Movement Disorder?

A.4.2.5 Barkley and Murphy, 1998: ADHD Current Symptoms Scale—Self-Report Form: Adults

Please circle the number next to each item that best describes your behavior during the past 6 months.

0	1	2	3
Never or Rarely	Sometimes	Often	Very Often

1. Fail to give close attention to details or make careless mistakes in my work
2. Fidget with hands or feet or squirm in seat
3. Have difficulty sustaining my attention in tasks or fun activities
4. Leave my seat in situations in which seating is expected

5. Don’t listen when spoken to directly
6. Feel restless
7. Don’t follow through on instructions and fail to finish work
8. Have difficulty engaging in leisure activities or doing fun things quietly
9. Have difficulty organizing tasks and activities
10. Feel “on the go” or “driven by a motor”
11. Avoid, dislike, or am reluctant to engage in work that requires sustained mental effort
12. Talk excessively
13. Lose things necessary for tasks and activities
14. Blurt out answers before questions have been completed
15. Am easily distracted
16. Have difficulty awaiting turn
17. Am forgetful in daily activities
18. Interrupt or intrude on others

A.4.2.6 BIS/BAS Scale (Carver & White, 1994)

Each item of this questionnaire is a statement that a person may either agree with or disagree with. For each item, indicate how much you agree or disagree with what the item says. Please respond to all the items; do not leave any blank. Choose only one response to each statement. Please be as accurate and honest as you can be. Respond to each item as if it were the only item. That is, don't worry about being "consistent" in your responses. Choose from the following four response options:

1 = very true for me
2 = somewhat true for me
3 = somewhat false for me
4 = very false for me

1. A person's family is the most important thing in life.
2. Even if something bad is about to happen to me, I rarely experience fear or nervousness.
3. I go out of my way to get things I want.
4. When I'm doing well at something I love to keep at it.
5. I'm always willing to try something new if I think it will be fun.
6. How I dress is important to me.
7. When I get something I want, I feel excited and energized.
8. Criticism or scolding hurts me quite a bit.
9. When I want something I usually go all-out to get it.
10. I will often do things for no other reason than that they might be fun.
11. It's hard for me to find the time to do things such as get a haircut.
12. If I see a chance to get something I want I move on it right away.
13. I feel pretty worried or upset when I think or know somebody is angry at me.
14. When I see an opportunity for something I like I get excited right away.
15. I often act on the spur of the moment.

16. If I think something unpleasant is going to happen I usually get pretty "worked up."
17. I often wonder why people act the way they do.
18. When good things happen to me, it affects me strongly.
19. I feel worried when I think I have done poorly at something important.
20. I crave excitement and new sensations.
21. When I go after something I use a "no holds barred" approach.
22. I have very few fears compared to my friends.
23. It would excite me to win a contest.
24. I worry about making mistakes.

A.4.2.7 Shortened Intolerance of Uncertainty Scale

Not at all characteristic of me	Somewhat uncharacteristic of me	Neither characteristic nor characteristic of me	Somewhat characteristic of me	Entirely characteristic of me
1	2	3	4	5

1. Unforeseen events upset me greatly.
2. It frustrates me not having all the information I need.
3. One should always look ahead so as to avoid surprises.
4. A small, unforeseen event can spoil everything, even with the best of planning.
5. I always want to know what the future has in store for me.
6. I can’t stand being taken by surprise.
7. I should be able to organize everything in advance.
8. Uncertainty keeps me from living a full life.
9. When it’s time to act, uncertainty paralyses me.
10. When I am uncertain I can’t function very well.
11. The smallest doubt can stop me from acting.
12. I must get away from all uncertain situations.

A.4.2.8 Demographic questions

Age:_________

Gender that you current identify with:____________________

Year in College: Freshman Sophomore Junior Senior 5th Year
Senior Post-baccalaureate Other

Major:_____________________________

Approximate Overall Grade Point Average:

0.0-1.0 1.1-1.5 1.6-2.0 2.1-2.5 2.6-3.0 3.1-3.5 3.6-4.0

A.4.2.9 Free write: Thoughts and Responses to Experiment

Please take a few moments to answer some questions about the research you just participated in. Please answer as honestly as possible – your responses will not in any way influence whether or not you receive credit.

1. What did you think that the purpose of this study was?
2. What were your thoughts and feelings during the experiment?

APPENDIX B

ANALYSIS OF PARTICIPANT SELF-REPORT OF STEREOTYPIC BEHAVIOR

B.1 Comparison of Self-Reported and Observed Stereotypic Behavior

In the questionnaire portion of the experiment, the participants completed the fidgeting tendency scale (Mehrabian & Friedman, 1986) which included 40 items describing fidgeting behaviors. I included 9 additional behaviors to the scale. After the participants read through and responded to the behaviors listed in the fidgeting tendency scale, they were asked additional questions about the likelihood that they performed these behaviors during the study, what behaviors they performed, and the general severity of their most frequent habit. I tested how well these self-report measures related to their observed behavior.

In order to align observed behaviors with the behaviors described in the self-report measures, I created more broad categories of behaviors than the ones described previously. To create these categories, I started with the 49 items in the modified fidgeting tendency scale ($\alpha = .89$) and qualitatively created different categories for the types of behaviors described. The four resulting factors were: Hands/Arms ($\alpha = .77$), Torso/Leg ($\alpha = .75$), Head/Mouth ($\alpha = .73$), and Objects ($\alpha = .77$). This latter category was for behaviors that directly indicated using objects such as pens in the behavior. Seven items did not reliably correlate with others and were not used in subsequent analyses.

With these factors created, I categorized the participants' observed behaviors based on these factors and summed their total duration (seconds) for performing each category of behavior (Figure B.1). When asked whether they performed any of the behaviors described in the fidgeting tendency scale during the experiment, 65 (95.6%) participants said that they did, one said no, and two said that they were unsure.

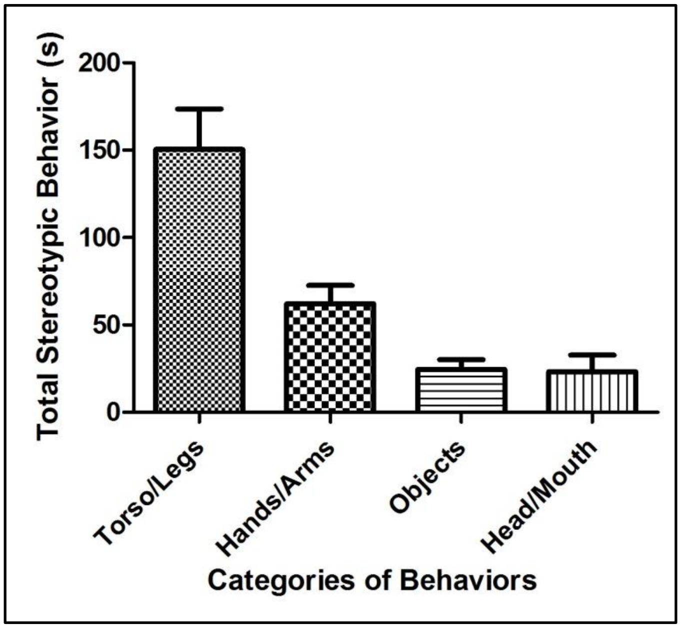

Figure B.1 Broad categories of stereotypic behaviors observed.

For the modified fidgeting tendency questionnaire, participants responded whether they performed the listed behaviors in the past 6 months via a 9-point Likert scale of *Very strongly agree* through *Very strongly disagree*. I tested whether participants' agreement with a behavior statement related to performing this type of behavior in the experiment. For all 4 factor-category pairs, there was no correlation between the participant response on the modified fidgeting tendency questionnaire and the duration of stereotypies observed during the experiment (Table B.1). Overall, the results suggest that participants' self-report measures of behaviors from the questionnaire do not correlate with behaviors observed in the experiment.

Table B.2 Pearson correlations between level of agreement with performing categories of behavior and the duration of observed behaviors.

Behavior Category	Behavior Duration (Mean±SD)	Questionnaire Response (Mean±SD)	Pearson Correlation (r) and significance (p)
Torso/Leg	150.49±190.60	3.87±1.27	r(68)=-0.225, p=0.065
Hands/Arms	62.09±86.67	3.97±0.97	r(68)=0.079, p=0.523
Objects	24.55±46.14	4.83±1.71	r(68)=-0.029, p=0.817
Head/Mouth	23.19±78.31	5.80±1.45	r(68)=-0.018, p= 0.885

Another way that the participants self-reported their behavior was through two open-response questions. I asked what their most frequent habit was as well as what behaviors they performed during the study. I qualitatively analyzed whether the participant response included behaviors that fit in each category. I used the 4 categories of behavior: Hands/Arms, Torso/Leg, Head/Mouth, and Objects and created *Yes* and *No* groups of participants for each of the behaviors. For both the most frequent habit and behaviors performed during the study, I performed one-way ANOVAs in order to test whether participants who indicated performing a category of behavior had different durations of these observed behaviors than participants who did not report performing a category of behavior.

For most behaviors, there were no significant differences between participants who reported performing each category of behavior and those who did not self-report performing a behavior, either for the most frequent habit or behaviors displayed during the study (Tables B.3, B.4). An exception to this was for Object use. Participants who reported that they manipulated objects during the study had a significantly higher total duration of Object behaviors than participants who did not report using objects during the study (F(1)=4.58, *p*= 0.036) (Figure B.2). Most, but not all, of the participants who used objects during the study accurately reported using these objects in their self-report answers. However, other than this scenario, the self-report measures collected either through the modified fidgeting tendency questionnaire and the open-ended questions did not relate to observed behavior.

Table B.3 Total duration of stereotypic behaviors in each behavior category, grouped by self-reports of each behavior category as a Most Frequent Habit or performed as Behavior during Study.

	Hands/Arms (Mean±SD)	Torso/Legs	Head/Mouth	Objects
Most Frequent Habit	Yes 64.88±93.82 n=53	Yes 188.95±235.74 n=17	Yes 19.27±17.67 n= 8	Yes 22.43±22.82 n=4
	No 52.21±55.83 n=15	No 137.67±173.85 n=51	No 23.72±83.21 n=60	No 24.68±47.31 n=64
Behavior during study	Yes 59.61±86.81 n= 63	Yes 167.00±191.41 n= 47	Yes 11.32±13.66 n=24	Yes 45.58±53.21 n=16
	No 93.33±87.67 n=5	No 113.55±188.03 n=21	No 29.67±96.61 n=44	No 18.08±42.22 n=52

Table B.4 One-way ANOVA results for differences in behavior performance as grouped by participants who reported performing each category of behavior. Bolded values represent p<0.05.

	Hands/Arms	Torso/Legs	Head/Mouth	Objects
Most Frequent Habit	F(1)= 0.247, p=0.621	F(1)= 0.922, p=0.341	F(1)= 0.022, p=0.881	F(1)= 0.009, p=0.925
Behavior during study	F(1)= 0.698, p=0.406	F(1)= 1.144, p=0.289	F(1)= 0.850, p=0.360	**F(1)= 4.581, p=0.036**

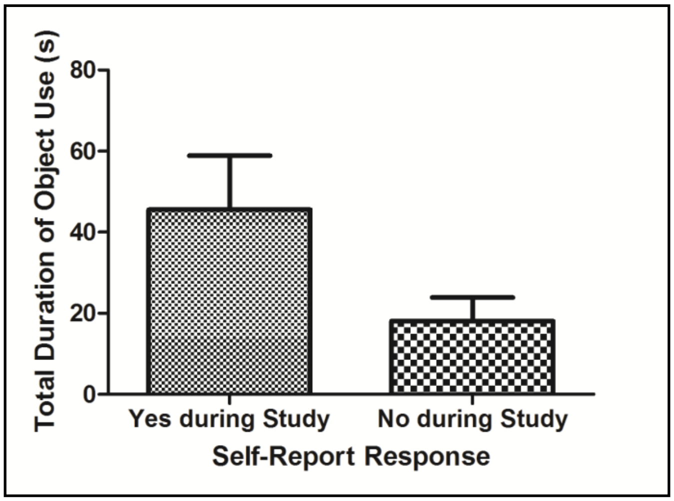

Figure B.5 Duration of object use during study by participants who reported using or not using objects.

With these broad categories of behavior generated, I also tested whether these types of stereotypies are related to each other with respect to the amount of time each participant performed these behaviors. Notably, these categories of behavior do not correlate with each other (Table B.5). These results suggest that people perform specific stereotypic behaviors that are possibly considered a habit to the participants rather than have a generalized tendency to perform many different kinds of stereotypies.

Table B.6 Pearson correlations between categories of behaviors.

	Torso/Legs	Hands/Arms	Objects
Torso/Legs	1		
Hands/Arms	r = 0.006, p=0.962	1	
Objects	r = 0.153, p=0.214	r = 0.074, p= 0.548	1
Head/Mouth	r = 0.012,p= 0.924	r= -0.129, p =0.294	r= -0.091, p=0.460

APPENDIX C

ANALYSIS OF THE EFFECTS OF NON-PREDICTABILITY FACTORS ON STEREOTYPIC BEHAVIOR

C.1 Gender

An independent samples t-test demonstrated that there were no significant differences in stereotypic behavior between males (*M*=49.81±34.09) and females (*M*=51.68±76.42) in the Wait paradigm (t(66)=0.98, p=0.923). There were also no significant differences between males (*M*=226.53±145.59) and females (*M*=209.02±210.28) in the DNMS paradigm (t(66)=0.318, p=0.751).

C.2 Reported ADHD and ASD diagnoses and Wore a watch

One factor that potentially warranted removal from subsequent analyses was that four participants reported having either or both an ADHD (n=4) and ASD (n=1) diagnosis. Secondly, eleven participants were recorded from the videos as wearing a watch. While wearing a watch may not generally affect stereotypic behavior, it may possibly affect responses to the experimental conditions. Both of these factors may affect stereotypic behavior.

I assessed whether any participants with these factors displayed levels of behavior different from the rest of the sample. Box plots (Figure C.1) of behavior demonstrated that there were statistical outliers; however, none of the outliers were participants that either wore a watch or had ADHD or ASD diagnoses. In addition, there were no significant differences between stereotypic behavior displayed by participants with ADHD or ASD and those without these diagnoses in the Wait paradigm (t(66)=1.15, p=0.253) (C.2). Because I had no indications that participants with an ADHD or ASD diagnosis or who wore a watch were meaningfully different from other participants with respect to stereotypic behavior during the study, I kept these participants in the analyses.

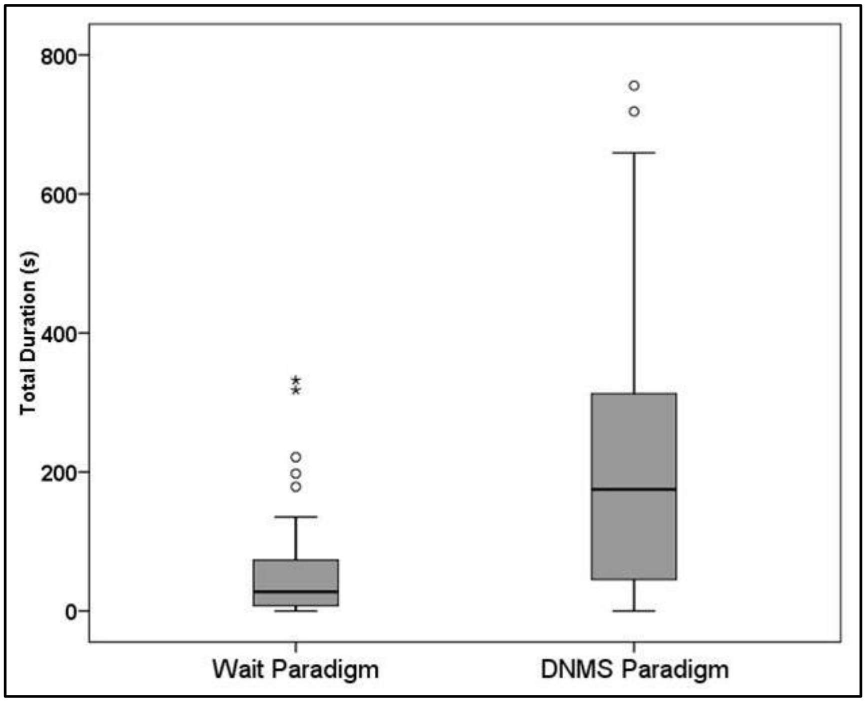

Figure C.1 Box plot distributions for total duration of stereotypies in Wait and DNMS paradigms.

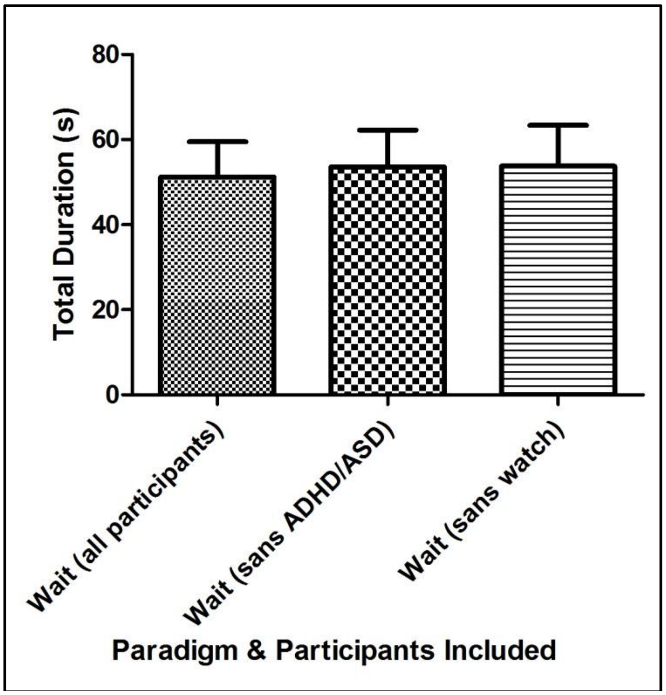

Figure C.2 Total duration of stereotypic behavior in Wait paradigm with all 68 participants included, with 4 participants eliminated (ADHD/ASD), and 11 participants eliminated (wearing a watch).

C.3 Performance on Task

The DNMS task contained seven trials. Participants viewed two stimuli and had to select the one that was not the stimulus previously presented to them. The participants received immediate visual and auditory feedback as to whether their choice was the right or wrong answer.

Most participants answered 5 (n=28) or 6 (n=24) questions correctly. Eight participants answered all 7 questions correctly, 7 participants answered 4 questions correctly, and 1 participant answered 3 questions correctly. No participants totaled less than 3 correct answers.

I used a one-way ANOVA with the number of correct responses as the between-subjects variable in order to test whether there was difference in stereotypic behavior duration between participants based on their performance in the task. There was no significant difference between the number of correct answers in the DNMS task and stereotypies in the DNMS paradigm (F(4,67)=1.17, *p*=0.332) (Table C.1). I also tested potential differences in behaviors in the Wait paradigm that may predict performance. However, there were also no statistically significant differences between performance on the task and stereotypic behaviors in the Wait paradigm (F(4,67)=0.880, *p*=0.481 (Table C.1)

Table C.1 The total amount of stereotypic behaviors performed in the Wait and DNMS paradigms as grouped by number of correct responses in the DNMS paradigm.

	Number Correct (out of 7 trials)				
	3	4	5	6	7
Wait Paradigm Behaviors (Mean ± SD) (s)	0±0 n=1	38.77±46.78 n=7	68.14± 73.37 n=28	37.22± 43.87 n=24	51.24± 114.34 n=8
DNMS Paradigm Behaviors (Mean ± SD) (s)	74.37±0 n=1	209.86± 238.27 n=7	267.05± 203.87 n=28	156.89± 154.89 n=24	215.62± 227.59 n=8

APPENDIX D

PARTICIPANT DISTRIBUTION BY CONDITION AND GENDER

D.1 Distribution of participants by gender across conditions for the Wait and DNMS paradigms.

Wait Paradigm	Predictable		Unpredictable	
DNMS Paradigm	Predictable with Clock	Unpredictable with Clock	Predictable without Clock	Unpredictable without Clock
Females	12	12	16	11
Males	5	4	1	7
Total	17	16	17	18

APPENDIX E

ETHOGRAM FOR EXPERIMENT 1 PARTICIPANT STEREOTYPIC BEHAVIOR

Behavior	Description
Stroke	Sustained contact between the hand or finger and another part of the body or substrate as the hand or finger is dragged across the surface. No minimum distance is required for a behavior to be considered a stroke, but the hand must move a visible distance across the surface. Once the hand is picked up off the surface, then the stroke has ended. The stroke has to involve at least 1 full finger, if it is a ½ finger (only the top joints) then that does not count as a stroke. *Counting Repetitions*: One drag across the surface is considered a stroke. More than one stroke may take place in an episode. Each hand performing strokes is counted separately (as 2 may be possibly stroking). An exception is if both hands are performing a stroking motion at the exact same time, simultaneously, then count this as 1 stroke even though both hands are involved.
Checks Watch	There is a clear watch or band on wrist, participant either moves arm or uses the other hand to angle the probable watch face towards the participant's face and eyes.
Bounce	Foot and/or leg that is not in contact with the ground moves up and down in the air either by movement at the ankle, knee, or hip joints. The foot or leg moves either up or down from the original position and then in the opposite direction. A bounce must consist of 1 up and 1 down (in either order) in order to be considered 1 bounce. Repetitions are counted. 1 bounce consists of 1 up/down, so 2 motions, as opposed to stroke in which one motion is counted.
Tap-Foot	Foot is in contact with the ground. The foot (either the front or back half) is elevated off the ground while the other half remains on the ground, and then the elevated portion is returned to the ground. The up and down movement is either by movement at the ankle, knee, or hip joints. A tap must consist of 1 up and 1 down (in that order) in order to be considered 1 tap. Repetitions are counted. 1 tap consists of 1 up/down, so 2 motions, as opposed to stroke in which one motion is counted.

Tap-Hand	Hand and/or forearm are in contact with either the body or another surface. The fingers, either individually or as a group, are elevated off the surface while the hand/forearm remains on the surface, and then the elevated portion is returned to the surface. The up and down movement is either by movement at wrist or elbow. A tap must consist of 1 up and 1 down (in that order) in order to be considered 1 tap. Repetitions are counted. 1 tap consists of 1 up/down, so 2 motions, as opposed to stroke in which one motion is counted.
Swing	Foot/feet on ground, and leg moves side-to-side while foot remains on the ground.
Other	Any other behaviors that are repeated 3 or more times. Indicate where the behavior was directed.

APPENDIX F

UMASS RHESUS MACAQUE DEMOGRAPHIC INFORMATION

Table F.1 ID, age, origin, and housing background for 14 rhesus macaques in Experiments 2 and 3.

ID (name in parentheses)	Sex	Age (years)	Origin	Housing
I18 (Linus)	Male (M)	24	NIH	Allentown; Single-housed
N01 (Friday)	M	14	UMass (offspring of V27 & V38)	Allentown; Single-housed
N02 (Lily)	Female (F)	12	UMass (offspring of V43 & V42)	Allentown; Single-housed
V27 (Ivan)	M	18	NIH	Pen; Grooming contact with V38
V38 (Taz)	F	18	NIH	Pen; Grooming contact with V27
V42 (Violet)	F	18	NIH	Pen; Grooming contact with V43
V43 (Coby)	M	18	NIH	Pen; Grooming contact with V42
ZA01 (Emmitt)	M	15	NIH	Pen; Pair-housed with ZA31
ZA02 (Nigel)	M	15	NIH	Allentown; Single-housed
ZA31 (Fozzie)	M	15	NIH	Pen; Pair-housed with ZA01
ZA54 (Bailey)	M	15	NIH	Pen; adjacent to ZA63
ZA56 (Kayla)	F	15	NIH	Allentown; Pair-housed with ZA65
ZA63 (Little G)	M	15	NIH	Pen; adjacent to ZA54
ZA65 (Zoey)	F	15	NIH	Allentown; Pair-housed with ZA56

APPENDIX G

ETHOGRAMS FOR RHESUS MACAQUE BEHAVIOR

G.1 Ethogram for stereotypies, anxious (yawn & scratch), aggressive (cage shake & threat), and tactile and oral exploration behaviors in Experiments 2 & 3.

Behavior	Description
Stereotypies	
Stereotypy	Any repetitive or ritualized pattern of behavior that serves no obvious function. It is not a part of play, sex, or grooming.
Stereotypy: Active-Pace	A whole-body motor pattern in which the animal locomotes in the same route and pattern for 3 or more consecutive cycles.
Stereotypy: Active-Other	A whole-body motor pattern in which the animal performs a same motion for 3 or more consecutive cycles. Examples include swinging or back flips.
Stereotypy: Self-directed: Eye-poke	The pressing of one corner of one eye with the thumb.
Stereotypy: Self-directed: Self-stroke	Monkey has one foot extended off ground or surface into the air. The foot then slowly makes contact with the body and slowly strokes the body from up to further down the body.
Stereotypy: Self-directed: Hair pull	Using the hands to grab and pull multiple hairs at once with no attempt of grooming evident.
Stereotypy: Other	Any behavior performed repetitively for at least 3 bouts with no more than 0.5 second between bouts. Examples include repetitive rubbing of hands or licking of an object.
Anxious Behaviors	
Yawn	A slow opening of the mouth to an extremely wide position exposing the teeth
Scratch	Any vigorous stroking of the hair by fingernails or toenails
Aggressive Behaviors	
Threat	A complex behavioral signal involving elements such as an open-mouth stare with teeth partially exposed, eyebrows lifted, ears flattened or flapping, rigid body posture, and a vocal element. The threat may contain all or some of these elements. Hand-slapping the floor in

	front of another monkey also is scored as a threat.
Cage Shake	Any vigorous shaking of the cage
Tactile and Oral Exploration	
Tactile-Oral Explore	Any tactile or oral manipulation of the cage or environment excluding contact with another animal. This category is strictly non-social. Involves an examination of the environment such as picking up, sniffing, or orally contacting, licking, biting, turning, rolling, gnawing objects, chains, or any part of the cage. Manipulation of food is scored in a separate category
Self-Injurious Behaviors	
Self-Bite	The animal inserts a part of its own body into its mouth and bites down on the body part vigorously. The animal may also move its mouth towards the body part rather than move the body part towards its mouth.

Table G.2 Other Behaviors sampled for in Modified Frequency sampling in Experiment 3.

*Only Locomotion and Visual Explore were analyzed as the other behaviors were too rare for analysis.

Behavior	Description
Visual Explore	The animal is sitting or standing motionless by itself with its eyes open. Passive behavior may be of short or long duration. Monkeys often break their activity with short and passive episodes. The pause has to be greater than 1 second in order to be considered visual explore. Make a note if the animal's eyes are closed in order to denote sleeping.
Locomotion	Two or more directed steps in the horizontal and/or vertical plane. Categories such as stereotypy, play, and aggression take precedence over scoring locomotion.
Vocalization	Any sound produced from the mouth and vocal apparatus
Aggress	Behavior involving an actual attack of another animal and can include biting, wrestling, chasing, hair pulling, jumping on another's back, etc. Play becomes aggression when one monkey fear grimaces and/or screams and the attacking monkey does not stop immediately.
Displacement	Takeover of an object, activity, or position of one animal by another. During the displacement, the displacer must touch the displacee or come within 2 feet. The displacer generally takes the place of the displacee or its activity or object.

Self Sex	Any tactile or oral manipulation of the genitals not involving grooming of them.
Crooktail	A strutting type of locomotion in which the tail is held high in the air and curled at the end.
Fear Grimace	A grin-like facial expression involving a retraction of the lips and exposing clenched teeth. This may be accompanied by flattened ears, stiff, huddled body posture, and screech vocalization.
Moan Lipsmack	Pursing the lips together and moving them together to produce a smacking sound, sometimes accompanied by moaning. Lipsmacking can occur during grooming or other social interactions.
Self-Mouth/Clasp	Any sucking of one owns body, usually fingers, toes, or genitals
Tac Oral Chow	Any tactile or oral manipulation of a piece of chow or food item.
Eat	The consumption of a food item. Eating shavings is scored as a separate category.
Drink	Any consumption of water from a water bottle. Drinking urine is not scored in this category.
Social Contact	Any passive contact not involving grooming, sex, aggression, or play. Physical contact means actual touching or within a monkey's arm length of each other.
Forage	Behavior involving the manipulation of a foraging substrate such as shavings, a foraging board, or a fleece grooming board.
Social Groom	Any picking, scraping, spreading, mouth-picking, or licking of an animal's hair by another animal.
Self Groom	Any picking, scraping, spreading, mouth-picking, or licking of an animal's own body hair. Also includes cleaning or chewing one's own fingernails and toenails.
Presents	Several postures often used to solicit grooming. Neck present involves lifting of the chin thereby exposing the neck. It entails exposing body surfaces in exaggerated ways to other animals.
Man/Eat Shavings	Any tactile or oral manipulation or consumption of wood shavings
Rump Present	A posture involving a stance on all fours with hind quarters elevated and tail raised. Animals may sometimes put their head between their legs. There may be brief tail flicks or by lifting tail to the side rather than in the air.
Mount	A posture in which an animal grabs the hind legs of another animal with its own hind feet (called a double-foot clasp) and places its hands on the lower back of the recipient. The animal may do everything but the double-foot clasp which is considered an incomplete mount. Attempted mounts can also be scored if the recipient is not in the correct rump-present posture.
Other	A behavior that does not fit the description of any other behavioral category. The behavior should be described when entered.

APPENDIX H

TESTS OF NORMALITY FOR EXPERIMENT 2 AND 3

Table H.1 Normality tests for durations of behavior observed in Experiment 2

Behavior Category	Shapiro-Wilk Test and significance	
	Predictable	Unpredictable
Stereotypies	W=0.718, p=0.001	W=0.730, p=0.001
Tactile-Oral Exploration	W=0.682, p<0.001	W=0.822, p=0.009
Yawn & Scratch	W=0.961, p=0.736	W=0.975, p=0.937
Cage shake & Threat	W=0.745, p=0.001	W=0.804, p=0.006

Table H.2 Normality tests for durations of behavior observed in Experiment 3

Behavior Category	Shapiro-Wilk Test and significance	
	Predictable	Unpredictable
Stereotypies	W=0.795, p=0.004	W=0.786, p=0.003
Tactile-Oral Exploration	W=0.510, p<0.0001	W=0.436, p<0.0001
Yawn & Scratch	W=0.870, p=0.042	W=0.788, p=0.004
Cage shake & Threat	W=0.588, p<0.0001	W=0.533, p<0.0001

APPENDIX I

THE EFFECTS OF SEX AND HOUSING ON BEHAVIOR INEXPERIMENTS 2 & 3

I.1 Experiment 2

Table I.1 The effects of sex and housing on stereotypic behavior frequency and average duration.

Behavior: Frequency				
Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	4.40±6.66	2.20±2.39	Main Effect (sex): F(1,12)=0.255, *p*=0.623 Interaction: F(1,12)=1.28, *p*=0.280
	Male	2.22±4.35	1.89±4.59	
Housing	Allentown	3.33±6.31	1.17±2.40	Main Effect (housing): F(1,12)=0.033, *p*=0.859 Interaction: F(1,12)=1.69, *p*=0.219
	Pen	2.75±4.71	2.63±4.72	
Behavior: Average Bout Duration				
Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	3.84±5.35	3.86±3.26	Main Effect (sex): F(1,12)=1.605, *p*=0.229 Interaction: F(1,12)=0.445, *p*=0.518
	Male	2.20±3.54	1.06±2.19	
Housing	Allentown	2.50±4.63	1.80±2.97	Main Effect (housing): F(1,12)=0.070, *p*=0.795 Interaction: F(1,12)<0.0001, *p*=0.986
	Pen	3.00±4.06	2.26±2.96	

Table I.2 The effects of sex and housing on yawn and scratch frequency and average duration. Bolded values represent p<0.05.

Behavior: Frequency

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	1.15±0.14 (log)	1.11±0.15 (log)	Main Effect (sex): F(1,12)=4.709, *p*=0.051+ Interaction: F(1,12)=0.280, *p*=0.607
	Male	1.39±0.25 (log)	1.39±0.26 (log)	
Housing	Allentown	1.22±0.10 (log)	1.18±0.10 (log)	Main Effect (housing): F(1,12)=1.95, *p*=0.188 Interaction: F(1,12)=0.243, *p*=0.631
	Pen	1.38±0.30 (log)	1.38±0.32 (log)	

Behavior: Average Bout Duration

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	2.31±0.68	2.91±1.32	**Main Effect (sex): F(1,12)=10.55, *p*=0.007**** Interaction: F(1,12)=0.001, *p*=0.979
	Male	4.21±1.08	4.80±1.42	
Housing	Allentown	2.82±0.52	3.41±0.70	Main Effect (housing): F(1,12)=3.30, *p*=0.094 Interaction: F(1,12)=0.001, *p*=0.977
	Pen	4.06±1.52	4.66±1.95	

Table I.3 The effects of sex and housing on tactile and oral exploration frequency and average duration.

Behavior: Frequency				
Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	4.20±2.86	4.20±4.09	Main Effect (sex): F(1,12)=0.375, p=0.552 Interaction: F(1,12)=1.02, p=0.333
	Male	11.67±23.83	9.00±20.12	
Housing	Allentown	2.83±3.37	3.17±4.22	Main Effect (housing): F(1,12)=0.897, p=0.362 Interaction: F(1,12)=2.12, p=0.171
	Pen	13.63±24.78	10.38±21.09	
Behavior: Average Bout Duration				
Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	2.19±1.55	1.11±1.09	Main Effect (sex): F(1,12)=0.002, p=0.966 Interaction: F(1,12)=2.49, p=0.140
	Male	1.68±2.54	1.73±2.89	
Housing	Allentown	0.87±1.32	0.79±1.07	Main Effect (housing): F(1,12)=1.72, p=0.214 Interaction: F(1,12)=0.400, p=0.539
	Pen	2.60±2.47	2.05±2.95	

I.4 The effects of sex and housing on cage shake and threat frequency and average duration. Bolded values represent p<0.05.

Behavior: Frequency

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	0.84±0.41 (log)	0.89±0.57 (log)	Main Effect (sex): F(1,12)=4.14, p=0.065 Interaction: F(1,12)=0.783, p=0.394
	Male	0.26±0.46 (log)	0.42±0.49 (log)	
Housing	Allentown	0.72±0.53 (log)	0.79±0.51 (log)	Main Effect (housing): F(1,12)=2.19, p=0.164 Interaction: F(1,12)=0.358, p=0.561
	Pen	0.28±0.44 (log)	0.44±0.55 (log)	

Behavior: Average Bout Duration

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	0.43±0.24 (log)	0.39±0.21 (log)	**Main Effect (sex): F(1,12)=10.24, p=0.008**** Interaction: F(1,12)=1.00, p=0.335
	Male	0.09±0.15 (log)	0.14±0.16 (log)	
Housing	Allentown	0.28±0.21 (log)	0.34±0.16 (log)	Main Effect (housing): F(1,12)=1.77, p=0.208 Interaction: F(1,12)=0.482, p=0.501
	Pen	0.16±0.27 (log)	0.16±0.21 (log)	

I.2 Experiment 3

I.2.1 Duration Sampling

Table I.5 The effects of sex and housing on duration of stereotypic behavior

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	17.99±19.44	23.33± 26.06	Main Effect (sex): F(1,12)=1.988, *p*=0.184 Interaction: F(1,12)=0.221, *p*=0.647
	Male	9.52±13.70	9.78±12.41	
Housing	Allentown	16.16±19.56	18.44±25.98	Main Effect (housing): F(1,12)=0.672, *p*=0.428 Interaction: F(1,12)=0.001, *p*=0.973
	Pen	9.83±13.06	11.75±12.02	

Table I.6 The effects of sex and housing on total duration of yawning and scratch behaviors. Data were log-transformed for ANOVAs.

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	4.13±4.79	4.41±0.78	**Main Effect (sex):** **F(1,12)=4.64, *p*=0.052+** Interaction: F(1,12)=0.024, *p*=0.878
	Male	11.14±9.19	15.12±10.60	
Housing	Allentown	4.66±1.90	5.97±1.16	Main Effect (housing): F(1,12)=2.871, *p*=0.116 Interaction: F(1,12)=0.045, *p*=0.835
	Pen	11.62±3.41	15.29±4.08	

Table I.7 The effects of sex and housing on total duration of tactile-oral exploration behaviors. Data were log-transformed for ANOVAs.

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	4.53±9.08	0.0±0.0	Main Effect (sex): F(1,12)=0.124, *p*=0.731 Interaction: F(1,12)=3.449, *p*=0.088
	Male	3.44±9.06	3.01±6.45	
Housing	Allentown	0.33±0.80	0.0±0.0	Main Effect (housing): F(1,12)=2.713, *p*=0.125 Interaction: F(1,12)=0.008, *p*=0.929
	Pen	6.45±11.07	3.39±6.79	

Table I.8 The effects of sex and housing on total duration of cage shake and threat behaviors. Bolded values represent p<0.05. Data were log-transformed for ANOVAs.

Between Subject Factor	Levels	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Mixed-design ANOVA results
Sex	Female	0.83±1.17	0.43±0.95	Main Effect (condition): F(1,12)=4.466, p=0.056 Main Effect (sex): F(1,12)=1.148, p=0.305 **Interaction: F(1,12)=7.301, *p*=0.019***
	Male	0.15±0.32	0.20±0.39	
Housing	Allentown	0.70±1.10	0.49±0.86	Main Effect (housing): F(1,12)=1.544, *p*=0.238 Interaction: F(1,12)=0.385, *p*=0.547
	Pen	0.16±0.33	0.12±0.34	

I.2.2 Modified Frequency Sampling

Table I.9 Main effects and interactions for stereotypies by sex, housing, and condition via two mixed-design ANOVAs.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=1.13, *p*=0.309
	Interaction	F(2,24)=1.072, *p*=0.358
Housing	Main Effect	F(1,12)=0.219, *p*=0.648
	Interaction	F(2,24)=0.431, *p*=0.655

Table I.10 Main effects and interactions for behavior rate by sex, housing, and condition via two mixed-design ANOVAs.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=0.355, *p*=0.562
	Interaction	F(1.167,14.005)=1.391, *p*=0.265
Housing	Main Effect	F(1,12)=0.197, *p*=0.665
	Interaction	F(1.304,15.653)=2.668, *p*=0.116

Table I.11 Main effects and interactions for behavior range by sex, housing, and condition via two mixed-design ANOVAs.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=0.134, *p*=0.721
	Interaction	F(2,24)=1.727, *p*=0.199
Housing	Main Effect	F(1,12)=0.268, *p*=0.614
	Interaction	F(2,24)=1.649, *p*=0.213

Table I.12 Main effects and interactions for yawn and scratch behaviors by sex, housing, and condition via two mixed-design ANOVAs. Bolded values represent p<0.05.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	**F(1,12)=5.66, *p*=0.035***
	Interaction	F(2,24)=2.391, *p*=0.113
Housing	Main Effect	F(1,12)=3.048, *p*=0.106
	Interaction	F(2,24)=0.991, *p*=0.386

Table I.13 Main effects and interactions for cage shake and threat behaviors by sex, housing, and condition via two mixed-design ANOVAs. Bolded values represent p<0.05.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=3.50, *p*=0.086
	Interaction	F(1.325,15.904)=1.874, *p*=0.191
Housing	Main Effect	**F(1,12)=4.95, *p*=0.046***
	Interaction	F(1.289,15.473)=0.587, *p*=0.496

Table I.14 Main effects and interactions for tactile and oral exploration behaviors by sex, housing, and condition via two mixed-design ANOVAs. Bolded values represent p<0.05.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=0.009, *p*=0.925
	Interaction	F(1.212,14.538)=0.402, *p*=0.575
Housing	Main Effect	**F(1,12)=4.864, *p*=0.048***
	Interaction	F(1.273,15.273)=1.306, *p*=0.283

Table I.15 Main effects and interactions for locomotion by sex, housing, and condition via two mixed-design ANOVAs.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=0.192, *p*=0.669
	Interaction	F(2,24)=0.613, *p*=0.550
Housing	Main Effect	F(1,12)=0.101, *p*=0.756
	Interaction	F(2,24)=1.123, *p*=0.342

Table I.16 Main effects and interactions for visual explore by sex, housing, and condition via two mixed-design ANOVAs. Bolded values represent p<0.05.

		Mixed-Design ANOVA results
Between Subject Factor	Effect	AM Data, Predictable, Unpredictable
Sex	Main Effect	F(1,12)=4.419, *p*=0.057
	Interaction	F(2,24)=0.94, *p*=0.911
Housing	Main Effect	F(1,12)=0.541, *p*=0.476
	Interaction	**F(2,24)=5.704, *p*=0.009***

APPENDIX J

ANALYSIS OF PREDICTABLE DELAY LENGTH IN PREDICTABLE AND UNPREDICTABLE SESSIONS IN EXPERIMENT 2

For the average duration, I both analyzed these behaviors with respect to my hypothesis, but I also used it to compare behaviors that occurred with similar delay lengths across conditions. There were delays of 15 seconds in both the Predictable and Unpredictable conditions: trials 1, 5, and 9. With similar delay lengths in these trials, I compared the duration for behaviors in these trials between conditions.

Three trials in each Unpredictable session had 15-second delays, which was the delay length of all trials in a Predictable session. These three trials were always the first (1), middle (5) and last (9) trial of the Unpredictable session. I calculated the average durations for behaviors in trials 1, 5, and 9 (Figure J.1) in order to specifically test the behavioral response to the predictable delay length when presented in both Predictable and Unpredictable sessions. These results demonstrate a consistency in the overall pattern of behaviors rather than a significantly different pattern of behaviors between Predictable and Unpredictable sessions even when the delay length is predictable. However, there were no differences between conditions for any of the behavioral categories measured.

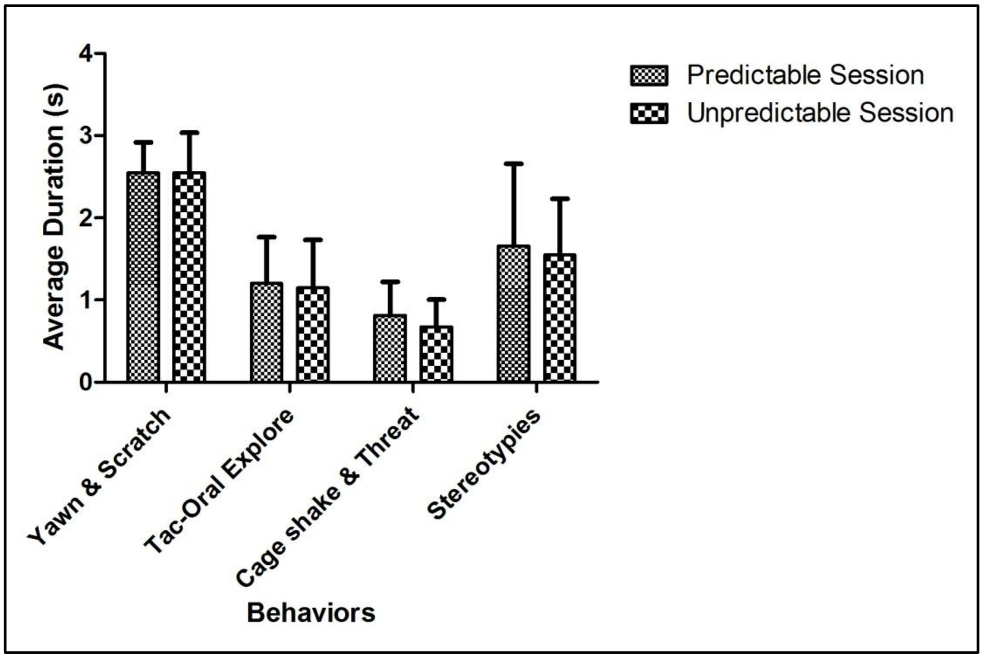

Figure J.1 The average durations of behaviors observed in trials 1, 5, and 9, by condition. Bars represent ±1 SEM.

Table J.1 Comparisons of behavior observed between condition in trials 1, 5, and 9.

Behavioral Category	Predictable Condition (Mean±SD)	Unpredictable Condition (Mean±SD)	Wilcoxon Signed-Rank Test
Yawn & Scratch	2.55±1.38	2.55±1.82	*Z*=0.031, *p*=0.975
Tactile-Oral Exploration	1.20±2.11	1.15±2.18	*Z*=0.392, *p*=0.695
Cage shake & Threat	0.81±1.52	0.68±1.23	*Z*=0.140, *p*=0.889
Stereotypies	1.66±3.75	1.55±2.53	*Z*=0.135, *p*=0.893

APPENDIX K

ANALYSIS OF POTENITAL HABITUATION IN EXPERIMENT 3

K.1 Overview

I analyzed the durations of the four categories of behaviors (stereotypies, aggressive, anxious, and tactile-oral exploration) in order to investigate whether the monkeys habituated to Experiment 3 over the 8 sessions. I used a Wilcoxon Signed-Ranks test to analyze whether there were differences between conditions when either the first Predictable and Unpredictable trials are compared or the last Predictable and Unpredictable trials are compared. There were no significant differences between behaviors observed in the first Predictable and Unpredictable conditions or in the last sessions (Table K.1). Figure K.1a of stereotypies suggests that macaques may have experienced habituation. However, this was observed across conditions, so the monkeys habituated to the entire paradigm rather than to certain conditions (Figure K.1).

Table K.1 Wilcoxon Signed-Ranks Test of First Predictable and Unpredictable sessions and the last Predictable and Unpredictable sessions.

Behavior	First Sessions	Last Sessions
Stereotypies	Z=0.135, *p*=0.893	Z=0.447, *p*=0.655
Yawn & Scratch	Z=0.0, *p*=1.00	Z= 0.078, p=0.937
Cage Shake & Threat	Z=1.34, *p*=180	Z=0.0, *p*=1.00
Tactile-Oral Explore	Z=1.00, *p*=0.317	Z=0.535, *p*=0.593

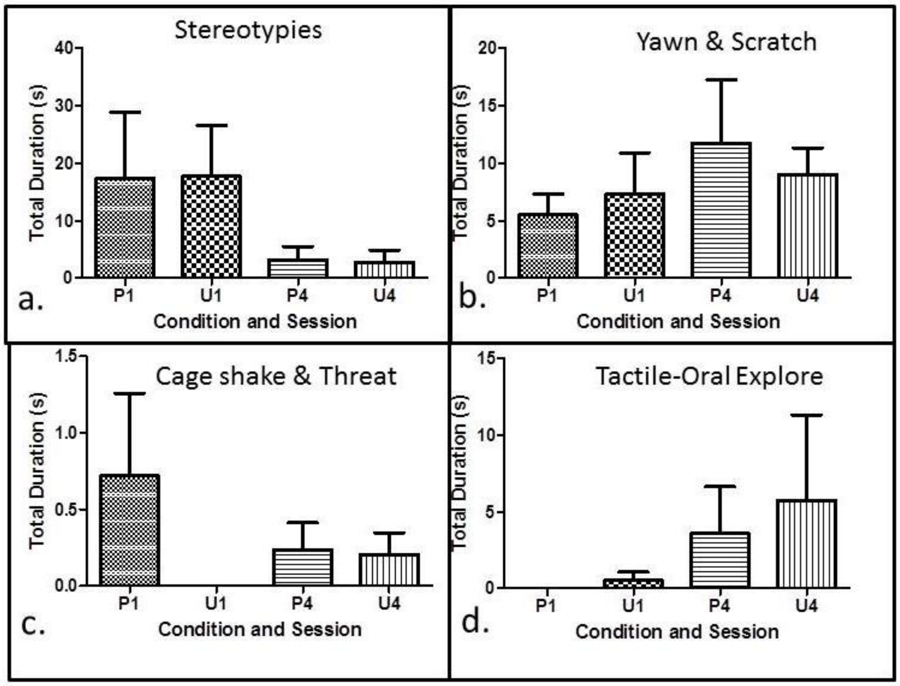

Table K.1 Total duration of stereotypic (a), yawn and scratch (b), cage shake and threat (c), and tactile-oral explore (d) observed in the first (1) and last (4) sessions of the Predictable and Unpredictable conditions.

REFERENCES

Altmann, J. (1974). Observational study of behaviors: Sampling methods. *Behaviour,* 49, 227-267.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition. Arlington, VA, American Psychiatric Association.

Asakura, S. (1958). A social relation between monkey and man. *Primates, 1*(2), 99-110.

Barash, D. P. (1974). Human ethology: Displacement activities in a dental office. *Psychological Reports*, *34*(3), 947-949.

Barkley R. A., & Murphy, K. R. (1998) *Attention Deficit Hyperactivity Disorder: A Clinical Workbook*. New York: Guilford Press.

Bauer, S. A., & Baker, K. C. (2016). Persistent effects of peer rearing on abnormal and species-appropriate activities but not social behavior in group-housed rhesus macaques (*Macaca mulatta*). *Comparative Medicine*, *66*(2), 129-136.

Bayne, K., Dexter, S., & Suomi, S. (1992). A preliminary survey of the incidence of abnormal behavior in rhesus monkeys (*Macaca mulatta*) relative to housing condition. *Lab Animal*, *21*, 38–46.

Bennett, K., Reichow, B., & Wolery, M. (2011). Effects of structured teaching on the behavior of young children with disabilities. *Focus on Autism and Other Developmental Disabilities*, *26*(3), 143-152.

Beran, M. J., & Smith, J. D. (2011). Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition*, *120*(1), 90-105.

Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(2), 111.

Berkson, G., Rafaeli-Mor, N., & Tarnovsky, S. (1999). Body-rocking and other habits of college students and persons with mental retardation. *American Journal on Mental Retardation*, *104*(2), 107-116.

Bloomsmith, M. A., & Lambeth, S. P. (1995). Effects of predictable versus unpredictable feeding schedules on chimpanzee behavior. *Applied Animal Behaviour Science, 44*(1), 65-74.

Bloomsmith, M. A., Hasenau, J., & Bohm, R. P. (2017). Position Statement: “Functionally appropriate nonhuman primate environments” as an alternative to the term “Ethologically appropriate environments.” *Journal of the American Association for Laboratory Animal Science*, *56*(1), 102-106.

Brakewood, C., Macfarlane, G. S., & Watkins, K. (2015). The impact of real-time information on bus ridership in New York City. *Transportation Research Part C: Emerging Technologies*, *53*, 59-75.

Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: caveats and guidelines. *Ethology*, *116*(5), 472-480.

Cannon, T. H., Heistermann, M., Hankison, S. J., Hockings, K. J., & McLennan, M. R. (2016). Tailored enrichment strategies and stereotypic behavior in captive individually housed macaques (*Macaca* spp.). *Journal of Applied Animal Welfare Science*, *19*(2), 171-182.

Carleton, R. N., Norton, M. P. J., & Asmundson, G. J. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders*, *21*(1), 105-117.

Carriere, J. S., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: Individual differences in mind wandering and inattention predict fidgeting. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(1), 19-31.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319.

Champoux, M., Metz, B., & Suomi, S. J. (1991). Behavior of nursery/peer-reared and mother-reared rhesus monkeys from birth through 2 years of age. *Primates*, *32*(4), 509-514.

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience & Biobehavioral Reviews*, *28*(3), 317-332.

Dettmer, A. M., Woodward, R. A., & Suomi, S. J. (2015). Reproductive consequences of a matrilineal overthrow in rhesus monkeys. *American Journal of Primatology*, *77*(3), 346-352.

Dickhäuser, O., Reinhard, M. A., & Englert, C. (2011). How task experience influences students' performance expectancies: The role of certainty. *Psychological Reports*, *109*(2), 380-388.

Díez-León, M., Bursian, S., Galicia, D., Napolitano, A., Palme, R., & Mason, G. (2016). Environmentally enriching American mink (*Neovison vison*) increases lymphoid organ weight and skeletal symmetry, and reveals differences between two sub-types of stereotypic behaviour. *Applied Animal Behaviour Science, 177*, 59-69.

Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind defining boredom in terms of attention. *Perspectives on Psychological Science*, *7*(5), 482-495.

Ecker, C. (2017). The neuroanatomy of autism spectrum disorder: An overview of structural neuroimaging findings and their translatability to the clinical setting. *Autism*, *21*(1), 18-28.

Edwards, M. J., Lang, A. E., & Bhatia, K. P. (2012). Stereotypies: a critical appraisal and suggestion of a clinically useful definition. *Movement Disorders*, *27*(2), 179-185.

Fahlman, S. A., Mercer-Lynn, K. B., Flora, D. B., & Eastwood, J. D. (2013). Development and validation of the multidimensional state boredom scale. *Assessment*, *20*(1), 68-85.

Fenton, W. N. (1953). Iroquois Eagle Dance an offshoot of the Calument Dance. *Bulletin*. Washington, D.C.: Smithsonian Institution. Retrieved from http://ehrafworldcultures.yale.edu/document?id=nm09-029

Field, M. J. (1970). Search For Security: An Ethno-Psychiatric Study Of Rural Ghana. *Norton Library*. New York, N.Y.: W. W. Norton & Company, Inc. Retrieved from http://ehrafworldcultures.yale.edu/document?id=fe12-027

Fisher, L., Ames, E. W., Chisholm, K., & Savoie, L. (1997). Problems reported by parents of Romanian orphans adopted to British Columbia. *International Journal of Behavioral Development*, *20*(1), 67-82.

Foley, P., & Kirschbaum, C. (2010). Human hypothalamus-pituitary-adrenal axis responses to acute psychosocial stress in laboratory settings. *Neuroscience & Biobehavioral Reviews*, *35*(1), 91-96.

Fowler, S. C., Covington III, H. E., & Miczek, K. A. (2007). Stereotyped and complex motor routines expressed during cocaine self-administration: results from a 24-h binge of unlimited cocaine access in rats. *Psychopharmacology*, *192*(4), 465-478.

Freymond, S. B., Bardou, D., Briefer, E. F., Bruckmaier, R., Fouche, N., Fleury, J., ... & Bachmann, I. (2015). The physiological consequences of crib-biting in horses in response to an ACTH challenge test. *Physiology & behavior*, *151*, 121-128.

Golub, M. S., Hogrefe, C. E., Widaman, K. F., & Capitanio, J. P. (2009). Iron deficiency anemia and affective response in rhesus monkey infants. *Developmental Psychobiology*, *51*(1), 47-59.

Gomot, M., Bernard, F. A., Davis, M. H., Belmonte, M. K., Ashwin, C., Bullmore, E. T., & Baron-Cohen, S. (2006). Change detection in children with autism: an auditory event-related fMRI study. *Neuroimage*, *29*(2), 475-484.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504-528.

Gottlieb, D. H., Capitanio, J. P., McCowan, B. (2013a). Risk factors for stereotypic behavior and self-biting in rhesus macaques (*Macaca mulatta*): Animal's history, current environment, and personality. *American Journal of Primatology*, *75*, 995–1008.

Gottlieb, D. H., Coleman, K., & McCowan, B. (2013b). The effects of predictability in daily husbandry routines on captive rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science*, *143*(2), 117-127.

Gottlieb, D. H., Maier, A., & Coleman, K. (2015). Evaluation of environmental and intrinsic factors that contribute to stereotypic behavior in captive rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science*, *171*, 184-191.

Hansen, D. J., Tishelmian, A. C., Hawkins, R. P., & Doepke, K. J. (1990). Habits with potential as disorders prevalence, severity, and other characteristics among college students. *Behavior Modification*, *14*(1), 66-80.

Hayes, A. F. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York: Guilford Press.

Hopper, L. M., Freeman, H. D., & Ross, S. R. (2016). Reconsidering coprophagy as an indicator of negative welfare for captive chimpanzees. *Applied Animal Behaviour Science*, *176*, 112-119.

Kickstarter.com: Most Funded. (2017). Retrieved from: https://www.kickstarter.com/discover/advanced?sort=most_funded

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76-81.

Kitanaka, J., Kitanaka, N., Tatsuta, T., Hall, F. S., Uhl, G. R., Tanaka, K., ... & Takemura, M. (2009). Sigma1 receptor antagonists determine the behavioral pattern of the methamphetamine-induced stereotypy in mice. *Psychopharmacology*, *203*(4), 781-79

Krantz, P. J., MacDuff, M. T., & McClannahan, L. E. (1993). Programming participation in family activities for children with autism: parents' use of photographic activity schedules. *Journal of Applied Behavior Analysis*, *26*(1), 137-138.

Kumar, S., & Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature*, *392*(6679), 917-920.

Lang, M., Krátký, J., Shaver, J. H., Jerotijević, D., & Xygalatas, D. (2015). Effects of anxiety on spontaneous ritualized behavior. *Current Biology*, *25*(14), 1892-1897.

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping*. New York: Springer publishing company.

Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders: a review of research in the last decade. *Psychological Bulletin*, *137*(4), 562-593.

Levine, R. (1997). A Geography of Time: The temporal misadventures of a Social Psychologist, or how every Culture Keeps Time just a little Bit Differently. New York: BasicBooks.

Loukidou, L., Loan-Clarke, J., & Daniels, K. (2009). Boredom in the workplace: More than monotonous tasks. *International Journal of Management Reviews*, *11*(4), 381-405.

Lutz, C. K., & Novak, M. A. (2005). Environmental enrichment for nonhuman primates: theory and application. *ILAR Journal*, *46*(2), 178-191.

Lutz, C., Well, A., & Novak, M. (2003). Stereotypic and self-injurious behavior in rhesus macaques: a survey and retrospective analysis of environment and early experience. *American Journal of Primatology*, *60*(1), 1-15.

Lutz, C. K., Coleman, K., Maier, A., & McCowan, B. (2011). Abnormal behavior in rhesus monkeys: risk factors within and between animals and facilities. *American Journal of Primatology*, *73*, S41.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, *4*(1), 84.

Mason, G. J., & Latham, N. R. (2004). Can't stop, won't stop: is stereotypy a reliable animal welfare indicator? *Animal Welfare*, *13*, S57-S70.

McLachlan, M., & McLachlan, M. (2016). Fidget Cube: a Vinyl Desk Toy. Retrieved from: https://www.kickstarter.com/projects/antsylabs/fidget-cube-a-vinyl-desk-toy

Mehrabian, A., & Friedman, S. L. (1986). An analysis of fidgeting and associated individual differences. *Journal of Personality*, *54*(2), 406-429.

Melnick, D. J., Pearl, M. C., & Richard, A. F. (1984). Male migration and inbreeding avoidance in wild rhesus monkeys. *American Journal of Primatology*, *7*(3), 229-243.

Mercer-Lynn, K. B., Flora, D. B., Fahlman, S. A., & Eastwood, J. D. (2011). The measurement of boredom: Differences between existing self-report scales. *Assessment*, *20*(5), 585-596.

Mercer-Lynn, K. B., Bar, R. J., & Eastwood, J. D. (2014). Causes of boredom: The person, the situation, or both?. *Personality and Individual Differences*, *56*, 122-126.

Munichor, N., & Rafaeli, A. (2007). Numbers or apologies? Customer reactions to telephone waiting time fillers. *Journal of Applied Psychology*, *92*(2), 511.

Nett, U. E., Goetz, T., & Daniels, L. M. (2010). What to do when feeling bored?: Students' strategies for coping with boredom. *Learning and Individual Differences*, *20*(6), 626-638.

Nimuendaju, C., & Lowie, R. H. (1946). Eastern Timbira. *University Of California Publications In American Archaeology And Ethnology*. Berkeley and Los Angeles: University of California Press. Retrieved from http://ehrafworldcultures.yale.edu/document?id=so08-001

Novak, M. A., Kinsey, J. H., Jorgensen, M. J., & Hazen, T. J. (1998). Effects of puzzle feeders on pathological behavior in individually housed rhesus monkeys. *American Journal of Primatology*, *46*(3), 213-227.

Novak, M. A., Tiefenbacher, S.T., Lutz, C., & Meyer, J. S. (2006). Deprived environments and stereotypies: insights from primatology. In G. Mason & J. Rushen (Eds.), *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (153-189). Wallingford, UK: CABI.

Novak, M. A., Peterson, E. J., Rosenberg, K., Varner, E. K., Worlein, J. M., Lee, G. H., Bellanca, R. U., & Meyer, J.S. (2015). Extreme behavioral phenotypes, hypothalamic-pituitary adrenal axis (HPA) activity and anxiety in rhesus macaques (*Macaca mulatta*) [Abstract]. *American Journal of Primatology*, *77* (S1), 99.

Pàmies, M. D. M., Ryan, G., & Valverde, M. (2016). What is going on when nothing is going on? Exploring the role of the consumer in shaping waiting situations. *International Journal of Consumer Studies*, 40(2), 211-219.

Pruetz, J. D., Bloomsmith, M. A. (1992). Comparing two manipulable objects as enrichment for captive chimpanzees. *Animal Welfare, 1*, 127-137.

Rafaeli-Mor, N., Foster, L., & Berkson, G. (1999). Self-reported body-rocking and other habits in college students. *American Journal on Mental Retardation, 104*(1), 1-10.

Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities, 26*(6), 527-547.

Richard, A. F., Goldstein, S. J., & Dewar, R. E. (1989). Weed macaques: the evolutionary implications of macaque feeding ecology. *International Journal of Primatology, 10*(6), 569-594.

Ridley, R. M., & Baker, H. F. (1982). Stereotypy in monkeys and humans. *Psychological Medicine, 12*, 61-72.

Rosati, A. G., Stevens, J. R., Hare, B., & Hauser, M. D. (2007). The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. *Current Biology, 17*(19), 1663-1668.

Schindler, S., Reinhard, M. A., & Dickhäuser, O. (2016). Boon and bane of being sure: the effect of performance certainty and expectancy on task performance. *European Journal of Psychology of Education, 31*(2), 245-253.

Schino, G., Perretta, G., Taglioni, A. M., Monaco, V., & Troisi, A. (1996). Primate displacement activities as an ethopharmacological model of anxiety. *Anxiety, 2*(4), 186-191.

Silk, J. B. (2009). Nepotistic cooperation in non-human primate groups. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1533), 3243-3254.

Singer, H. S. (2009). Motor Stereotypies. *Seminars in Pediatric Neurology, 16*, 77-81.

Southwick, C. H. (1967). An experimental study of intragroup agonistic behavior in rhesus monkeys (*Macaca mulatta*). *Behaviour, 28*(1), 182-209.

Stein, D. J., Denys, D., Gloster, A. T., Hollander, E., Leckman, J. F., Rauch, S. L., & Phillips, K. A. (2009). Obsessive–compulsive disorder: diagnostic and treatment issues. *Psychiatric Clinics of North America, 32*(3), 665-685.

Suomi, S. J., Harlow, H. F., & Kimball, S. D. (1971). Behavioral effects of prolonged partial social isolation in the rhesus monkey. *Psychological Reports, 29*(3 suppl), 1171-1177.

Sweeny, K., & Andrews, S. E. (2014). Mapping individual differences in the experience of a waiting period. *Journal of Personality and Social Psychology*, *106*(6), 1015.

Terwiel, B. J. (1975). Monks And Magic: An Analysis of Religious Ceremonies In Central Thailand. *Monograph Series*. Lund, Sweden: Studentlitteratur ; Curzon Press. Retrieved from http://ehrafworldcultures.yale.edu/document?id=ao07-019

Timmins, R.J., Richardson, M., Chhangani, A. & Yongcheng, L. (2008). *Macaca mulatta*. The IUCN Red List of Threatened Species 2008: e.T12554A3356486. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T12554A3356486.en. Downloaded on 17 June 2016.

Titiev, M. (1944). Old Oraibi: A Study of The Hopi Indians Of The Third Mesa. *Papers*. New York: Kraus Reprint Co. Retrieved from http://ehrafworldcultures.yale.edu/document?id=nt09-00

Tustin, R. D. (1995). The effects of advance notice of activity transitions on stereotypic behavior. *Journal of Applied Behavior Analysis*, *28*(1), 91-92.

Twohig, M. P., & Varra, E. M. (2006). Treatment of drug-induced stereotypy. *The Behavior Analyst Today*, *7*(2), 206.

Ulyan, M. J., Burrows, A. E., Buzzell, C. A., Raghanti, M. A., Marcinkiewicz, J. L., & Phillips, K. A. (2006). The effects of predictable and unpredictable feeding schedules on the behavior and physiology of captive brown capuchins (*Cebus apella*). *Applied Animal Behaviour Science*, *101*(1), 154-160.

Vandeleest, J. J., McCowan, B., & Capitanio, J. P. (2011). Early rearing interacts with temperament and housing to influence the risk for motor stereotypy in rhesus monkeys (*Macaca mulatta*). *Applied Animal Behaviour Science*, *132*(1), 81-89.

van Tilburg, W. A., & Igou, E. R. (2017). Can boredom help? Increased prosocial intentions in response to boredom. *Self and Identity*, *16*(1), 82-96.

Vogel-Walcutt, J. J., Fiorella, L., Carper, T., & Schatz, S. (2012). The definition, assessment, and mitigation of state boredom within educational settings: A comprehensive review. *Educational Psychology Review*, *24*(1), 89-111.

Waitt, C., & Buchanan-Smith, H. M. (2001). What time is feeding?: How delays and anticipation of feeding schedules affect stump-tailed macaque behavior. *Applied Animal Behaviour Science*, *75*(1), 75-85.

Watkins, K. E., Ferris, B., Borning, A., Rutherford, G. S., & Layton, D. (2011). Where is my bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. *Transportation Research Part A: Policy and Practice*, *45*(8), 839-848.

Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the Positive and Negative Affect Schedule-Expanded Form.* Ames: The University of Iowa.

Whiting, J. W. M.. 1970. “Kwoma Journal.” *Hraflex Books*. New Haven: Human Relations Area Files, Inc. http://ehrafworldcultures.yale.edu/document?id=oj13-002.

Woods, D. W., & Miltenberger, R. G. (1996). Are persons with nervous habits nervous? A preliminary examination of habit function in a nonreferred population. *Journal of Applied Behavior Analysis*, *29*(2), 259-261.

Woods, D. W., Miltenberger, R. G., & Flach, A. D. (1996). Habits, tics, and stuttering prevalence and relation to anxiety and somatic awareness. *Behavior Modification*, *20*(2), 216-225.

Young, F. M. (1947). The incidence of nervous habits observed in college students. *Journal of Personality*, *15*(4), 309-320.

Zvolensky, M. J., Lejuez, C. W., & Eifert, G. H. (2000). Prediction and control: Operational definitions for the experimental analysis of anxiety. *Behaviour Research and Therapy*, *38*(7), 653-663.