Distribution Agreement

Signature:

____________________________ _________________

Cyrus N. Chi Date

Implementation of Discriminative Stimuli in the Evolutionary Theory of Behavior Dynamics

By

Cyrus N. Chi
Doctor of Philosophy

Psychology

Jack J McDowell, Ph.D.
Advisor

Phil Wolff, Ph.D.
Committee Member

Michael Treadway, Ph.D.
Committee Member

Darryl Neill, Ph.D.
Committee Member

Samuel L. Morris, Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D, MPH
Dean of the James T. Laney School of Graduate Studies

Date

Implementation of Discriminative Stimuli in the Evolutionary Theory of Behavior Dynamics

Cyrus N. Chi
B.A., Rutgers, The State University of New Jersey, 2001
M.S., Rutgers, The State University of New Jersey, 2005
M.A., Emory University, 2019

Advisor: Jack J McDowell, Ph. D

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Psychology
2024

Abstract

Implementation of Discriminative Stimuli in the Evolutionary Theory of Behavior Dynamics

By Cyrus N. Chi, M.S., M.A.

McDowell's evolutionary theory of behavioral dynamics (ETBD) is a complexity theory that treats behaviors within an organism as 'agents' that interact with each other according to evolutionary principles. The theory has been used to animate artificial organisms (AOs) that produce behaviors that are considered the predictions of the theory. The theory's predictions have been found to be congruent with a number of quantitative findings in environments with reinforcers and punishers. However, the theory as it currently exists does not have a paradigm for engaging with discriminative stimuli in the environment. In order to enhance the theory, elements of Estes' stimulus sampling theory were adapted into a form compatible with the ETBD and added into the ETBD's functional loop. AOs animated by the modified ETBD were tested in concurrent schedule, and multiple schedule environments. When AOs were found to not appropriately behave in the similar manner to live organisms, additional modifications based on behavioral principles (e.g. reinforcement based attention, background reinforcement) were added in order to improve AO functioning. The results show that the modified ETBD was able to replicate the previous finding on concurrent schedule behavior and predict learning based on discriminative stimuli, but not all features of live organism behavior was able to be reproduced with the modified ETBD described here. The principles that functioned well (e.g. entropy-based observation) and the additional principles deemed necessary (i.e. durability of learning, selectivity for stimuli) to model discriminative stimulus behavior are discussed.

Keywords: ETBD, Stimulus Sampling Theory, Stimulus Control, Complexity Theory

Implementation of Discriminative Stimuli in the Evolutionary Theory of Behavior Dynamics

By

Cyrus N. Chi
B.A., Rutgers, The State University of New Jersey, 2001
M.S., Rutgers, The State University of New Jersey, 2005
M.A., Emory University, 2019

Advisor: Jack J McDowell, Ph. D

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Psychology
2024

Acknowledgements

I thank my family for all of their wisdom and encouragement.

I thank my wife for being there with me through everything and for her unending patience.

Most of all, I thank Dr. McDowell for his guidance, and for giving me the chance to grow in his natural garden.

Contents

I. Introduction

This dissertation is focused on expanding McDowell's computational Evolutionary Theory of Behavior Dynamics (ETBD; McDowell, 2004) to allow for the implementation of discriminative stimulus control. The ETBD, as it currently stands, accounts only for response-consequence phenomena and cannot account for antecedent-response-consequence phenomena like what is seen in multiple schedules and many real-world situations. This gap was partially bridged by adapting the concepts in Estes' mathematical Stimulus Sampling Theory (SST; Estes, 1950) into forms that are compatible with complexity theory, and then incorporating these into the ETBD. The modified ETBD was tested using multiple schedule experiments and stimulus generalization experiments to determine if the theory could produce animal-like behavior in those paradigms.

The dissertation is divided into three components. The first component focused on adapting SST into a computational form and merging it with the ETBD. Experiment series one was completed to verify that the modified ETBD can generate artificial organisms that duplicate an iconic finding of the unmodified ETBD. The second component tested the behavior of artificial organisms that are animated by the modified ETBD in a multiple schedule. Experiment series two aimed to establish that artificial organisms can discriminate between stimuli similarly to live organisms within a multiple schedule. The third component tested if artificial organisms would be able to produce animal-like stimulus gradients. Experiment series three ran artificial organisms through an experimental procedure similar to the one used by Guttman and Kalish (1956). Each experiment was accompanied by two modifications each, for a total of six modifications. The modifications are further changes to the artificial organisms or the experimental environment made in response to phenomena observed in interim results. Although all modifications increased understanding of the underlying challenges, not all changes implemented in the modifications were efficacious enough to be included in further testing. After the second and third experiments were completed, due to the presence of modifications, versions of the first experiment were run in order to certify that the artificial organisms that were restructured as specified by the modifications were still able to perform in an expected manner on concurrent schedules.

This document consists of seven sections. The first section reviews multiple schedules, the phenomena of stimulus generalization, the ETBD, and SST along with their significant findings. The second section discusses integration of the ETBD and SST. The third section discusses general methods used in all studies. The fourth section covers experiment series one. It contains the experimental procedures, analysis methods, results, and discussion for this study. The fifth section covers experiment series two. It contains the experimental procedures, analysis methods, results, and discussions for this study.. The sixth section contains a general discussion of all three experimental series and their results in context of the field. Finally, section seven discusses critical components necessary for further work into stimulus control for the ETBD. For the sake of brevity, the methods, results, and discussion for the stimulus generalization experiments (experimental series three) were placed into Appendix A.

Introduction to Multiple Schedules

Multiple schedules are a behavioral experimental arrangement in which two or more schedules of reinforcement are arranged sequentially with discriminative stimuli presented concurrently with each schedule (Pierce & Cheney, 2013). During a multiple schedule, each component schedule is presented to the subject sequentially for a predetermined amount of time or number of reinforcers. After all of the component schedules are presented, the sequence repeats until the end of the multiple schedule. For example, if the multiple schedule has two component schedules, schedule A and schedule B, then the organism will be presented with schedule A, then B, then A, and so on until the end of the experiment.

Through exposure to a multiple schedule, an organism is expected to demonstrate a differential response to the discriminative stimuli (Skinner, 1938). In Figure 1, the hypothetical behavioral results from a multiple schedule is shown. The example shows the session-by-session development of an organism's behavior during a multiple schedule with (i) a variable interval schedule that arranges reinforcement for the first response after, on average, 2 minutes (VI 2-min) and (ii) an extinction schedule (EXT). In the figure, the discriminative stimulus for the VI 2-min is shown as S^D and the extinction stimulus for the EXT is shown as S^Δ. As the number of sessions increases, the behavior between the two component schedules diverges markedly, showing that the organism is successfully differentiating between the two schedules by means of their discriminative stimuli.

In addition to research into discriminative stimuli, multiple schedules have been used to study behavioral contrast (Reynolds, 1961), behavioral momentum (Nevin, 1992), response rate in the presence of drugs in animal models (Gonzalez, 2002), and stimulus generalization (Hanson, 1959). Stimulus generalization will be discussed further in a later section.

Clinical Relevance of Multiple Schedules. One of the earliest demonstrations of multiple schedules used in training human behavior was by Bijou and Orlando (1961). They trained 46 institutionally housed children diagnosed with intellectual disabilities to have differential rates of lever-pressing based on the presence of a red or blue light. This experiment showed that multiple schedules had the potential to be an efficient teaching tool for complex behaviors that may require multiple steps. Since this experiment, multiple schedules have been used regularly by behavior analysts to teach specific skills, assess preferences, and control rates of behavior.

In school settings, multiple schedules have been used to manage disruptively high-rate requests for teacher attention by controlling the availability of attention from instructors. In one experiment, using multiple colored leis as discriminative stimuli, an instructor would respond to specific groups of students based on the discriminative stimulus present and ignore the other students (n = 20; Nava et al., 2016; n = 2; Niedfeld et al., 2020). In this experiment, question asking behavior would be put on extinction if the teacher was wearing a lei with a certain color, and reinforced with attention if the teacher was wearing a lei with a different color. Student behavior came under discriminative stimulus control, leading to reduced classroom disruption and increased instruction time. This training has been found to be effective with students in preschool (Vargo et al., 2014) up to sixth grade in both small and large group settings (Cammilleri et al., 2008).

Multiple schedules have also been used with children with autism spectrum disorder (ASD) to promote adaptive behavior and assess for preferences in non-verbal or younger children. For example, toe walking is a behavior that is found in up to 20% of ASD cases (Ming et al., 2007) and can require invasive surgery to surgically counteract the shortening of the Achilles tendon, or serial casting to stretch the muscles around the heel. Multiple schedule conditioning has been found to substantially reduce toe walking behavior (n = 1; Hodges et al., 2018), preventing the need for more extreme measures. In their study, Hodges and her collaborators used a wristband as a discriminative stimulus and used both rewards and punishments (e.g. rewarding praise like ‘nice walking’, and punishments like saying “no toe walking; the wristband is on”) while the child was wearing the wristband and gave neither rewards nor punishments while the wristband was off. After the entire protocol, with the wristband on, the child was able to walk 1000 steps in a community setting with toe-walking occurring during only 0.03% of the steps. The authors noted that prior to the treatment, rewards and punishments were also used to try to control toe-walking, but were found to be ineffective. The presence of the discriminative stimulus, the wristband, was integral for establishing stimulus control over the toe-walking behavior. Treatment using multiple schedules has also been found to be effective in increasing the use of mands (i.e. verbal requests for reinforcement; Akers et al., 2019) rather than other maladaptive behaviors, and reducing compulsive-like behavior in ASD patients (Chok & Harper, 2016).

Multiple schedules have also been used extensively in Functional Communication Training (FCT; Carr & Durand, 1985) as a method for schedule thinning. In FCT, the function of a challenging behavior is identified and a functionally equivalent communication response is taught/established while the challenging behavior is placed on extinction. Once the training is complete, the amount of reinforcement given during the training is often unsustainable in an *in vivo* setting, like the home or in school. This can lead to a delay of reinforcement or, in the worst case, the absence of reinforcement after a mand. This in turn can lead to recovery of the disruptive behavior, resurgence (i.e. treatment relapse), or the appearance of other problematic behavior. A multiple schedule can be used to gradually reduce the amount of reinforcement needed while also keeping the disruptive behavior at a low level (n = 4; Fisher et al., 1998; Hanley et al., 2001). This has been shown to be effective when addressing (i) multiple target behaviors (e.g. aggression, disruptions, screaming, self-injurious behavior, etc.,), (ii) both positively and negatively reinforced behaviors, (iii) multiple levels of intellectual ability, and (iv) children of all ages (n = 20; Fisher et al., 2020; Greer et al., 2016; Shamlian et al., 2016).

Behavioral thinning in FTC has been further developed to create individualized behavioral thinning multiple schedules (n = 5; Call et al., 2018). In a typical behavioral thinning multiple schedule, there is no systematic method used for determining the rate of behavioral thinning. Call and colleagues altered the duration of the component schedules based on the participant’s response patterns during treatment. This could potentially reduce the time spent during behavioral thinning for compliant patients, and reduce the possibility of relapse in more difficult patients.

The research into multiple schedules is one area of behavioral research that has been found to be useful outside of the laboratory. However, most clinical studies using multiple schedules have limited sample sizes (typically ranging from 1 to 20). Experiments that increase sample sizes, or streamline the design or testing of various protocols, could be of benefit to the field. Extension of the ETBD to allow modeling of behavior on multiple schedules would be able to meet this need. The number of AOs that can be run is only limited by computer power and time. In addition, since AOs that are animated by the ETBD can be placed into any schedule environment, clinical protocols can be easily tested and modified to increase effectiveness or to reduce treatment time.

Introduction to Stimulus Generalization

This introduction to stimulus generalization covers the basic areas of stimulus generation relevant to this dissertation. There is a large amount of stimulus generalization-related phenomena that are not relevant to the current series of experiments, although they may one day be excellent targets for future investigation using AOs animated by the ETBD. This section will focus on single discrimination stimulus generalization gradients, and discuss (i) basic definitions and methods, (ii) the inputs, (iii) outputs, and (iv) factors in stimulus generalization.

Stimulus generalization is a manifestation of stimulus control. Stimulus control is observed when an event (the controlling stimulus) alters the probability that a response will occur (Pierce & Cheney, 2013, p. 254) due to the differential correlation between the event and the consequences the response may engender. When the probability of the response increases, then the stimulus is called a discriminative stimulus. When the probability of the response decreases, then the stimulus is called an extinction stimulus. While this is a relatively straight-forward dynamic, it is not the end of the story. When another novel stimulus that is similar to the controlling stimulus on some dimension appears (e.g. the color pink and the color red, or a 1000 hertz tone and a 990 hertz tone), organisms will still have a change in response, but not as strongly as it would if the controlling stimulus was presented. Then, as the distance between the novel stimulus and discriminative stimulus increases, the organism's response to the novel stimulus will (i) monotonically decrease, in the case of a discriminative stimulus, or (ii) monotonically increase, in the case of an extinction stimulus. At some point the distance will be great enough that there is no change in response to the novel stimulus. This pattern of behavior is what is referred to as stimulus generalization. Put simply, "Stimulus generalization is the tendency, after experience with a given stimulus, to react in the same way to other, similar stimuli" (Staddon & Reid, 1990).

Stimulus generalization has long been investigated. Although some have traced stimulus generalization back to Aristotle's principle of association by resemblance (Shepard, 1987), the most common starting point is generally considered to be Pavlov's work with salivating dogs. When he trained dogs to salivate in response to a tone, he found that similar tones also acquired similar properties to the original tone, the conditioned stimulus. Those properties were also found to diminish as the new tone's similarity to the original tone diminished (Pavlov, 1927, p.

113). From that point onward, there were many experiments done to explore stimulus generalization, but there was a significant amount of controversy about the cause of stimulus generalization, and some questioned whether it existed at all in the form described (Razran, 1949). It is likely that methodological challenges led to a portion of the controversy. It wasn't until Guttman and Kalish's study, which was based on a format recommended by Skinner (1950, p. 204), that a reliable method of empirically generating stimulus generalization gradients was established.

In Guttman and Kalish's paper (1956), they used 24 naive pigeons that were restricted to 80% of their free-feeding body weight. Using a modified operant chamber, the birds were trained to eat from a magazine and, through differential reinforcement of successive approximations (i.e., shaping), were trained to peck at a lit key. The pigeons were split into 4 groups, and each groups was trained to peck at the lit key to receive reinforcement on one of four wavelengths: 530, 550, 580, and 600Mµ. After a series of trainings, they were then run through a generalization test, which was carried out under extinction. During the generalization test, each bird was exposed to 12 series of 11 different wavelengths of light presented in random order. During each series, the key in the Skinner box was lit with one of the 11 different wavelengths of light for 30 seconds, followed by a 10 second black out period. Then the pattern would repeat until all 11 wavelengths were presented. After this, additional training to the initial four wavelengths was done, followed by a repeat of the generalization test. For most groups, the 11 wavelengths were divided such that, along with the wavelength of the discriminative stimulus, and five wavelengths were taken from wavelengths below the conditioned stimuli and five wavelengths were taken from the wavelengths above the conditioned stimuli. The wavelengths chosen were all between 10-20 Mµ (i.e., millimicron or nanometer) apart from each other.

Figure 2 shows some of the results from their experiment. The mean generalization gradients for the first generalization test (Figure 2, left subplot) show a clear peak at the wavelength of the conditioned stimuli for each group with a decrementing number of responses as the distance from the conditioned stimulus increases. Individual generalization gradients for 12 of the pigeons (Figure 2, right subplot) all show the same general shape, although there are individual differences in height, and slope. The solid line in the right subplot is the result from the first generalization test, while the dashed line is the result of the second generalization test. For the 12 pigeons shown here, there is a wider amount of variability in height during the second generalization test, as well as a marked decrease in number of responses in most cases, although not all. As Guttman and Kalish note, the overall decrease in behavior in the second generalization test is likely due to the fact that the generalization testing was done under extinction. While there are other findings from the study, the most critical part of their study that has likely led to most of its 985 citations on Google Scholar is the “orderliness and reproducibility of the generalization process” (Guttman & Kalish, 1956, p. 84).

In their study, Guttman and Kalish used particular wavelengths of light as their discriminative stimulus. The wavelengths around the conditioned stimulus are considered to be along the ‘stimulus dimension.’ These stimulus dimensions are generally considered to be innate, based on sensory organs or neural processing (Ganz & Riesen, 1962; Honig & Urcuioli, 1981) but they have also been found to be influenced by adopted generalization rules in humans (Zaman et al., 2022). There are generally two categories of stimulus dimension; the intensity dimension and the rearrangement dimension. There are also stimuli which do not neatly fall into the two categories. The stimuli along an intensity dimension, as the name suggests, vary based on the amount of sensory stimulation they provide the test subject. Experiments using a the volume of white noise (Zielinski & Jakubowska, 1977), the brightness of a fake egg (Baerends & Drent, 1982), or the chemical concentration of an odor (Bhagavan & Smith, 1997) are engaging subjects with intensity-based stimuli. Stimulus gradients in this category have a markedly different shape from the one described above. When stimuli at higher intensity than the conditioned stimulus are presented, the rate of behavior will tend to monotonically increase (in the case of a discriminative stimulus), rather than decrease. Stimuli along the rearrangement dimension in contrast to the intensity dimension, have a constant amount of sensory stimuli presented, but differ in terms of arrangement. For example, studies which use wavelengths light (Guttman, 1959; Zidar et al., 2019), sound frequencies (Galizio, 1985), or locations in space (Cheng et al., 1997), and line orientations (Bloomfield, 1967; Soto & Wasserman, 2010) all fall into this category. Other stimulus dimensions are harder to categorize, like complex shapes (Ferraro & Grisham, 1972), visual patterns (Pearce et al., 2008), or the calls of monkeys (Masataka, 1983). Although some stimulus dimensions have been suggested to be fit into one of the two categories based on the similarity of their stimulus generalization gradient shape (Ghirlanda, 2002), other stimulus dimensions do not fall neatly into either category.

Stimulus gradients are typically generated using response rate or response probability as the dependent variable. This response data is then processed in order to determine the relationships between the independent variables and the dependent variables. The most common ways the data are statistically organized include: (i) the area under the gradient, (ii) the height or elevation of the condition with the greatest responding (i.e. the peak), (iii) the slope, (iv) form, and (v) symmetry. The data are also averaged across multiple organisms due to the amount of variation between organisms, or normalized to obtain relative generalization gradients. Sometimes these methods are used in tandem to describe a phenomena. For example, the condition with the greatest responding during testing is typically the one that presents the discriminative stimulus. However, in some discrimination studies that have both an extinction stimulus and a discriminative stimulus present on the same dimension, the peak will occur on a novel stimulus and not when the conditioned stimulus is presented. In addition, the peak will be on the far side of the conditioned stimulus, away from the extinction stimulus. The phenomenon is called peak shift, or response bias (Tinbergen, 1951). The slope is often described in a qualitative manner, with it being considered “steep” or “shallow” (Lotfizadeh et al., 2012). Although this simplification can be somewhat sufficient for direct comparisons with straight slopes, other more rounded shapes (Figures 3 and 4; Blough, 1972; Thomas & King, 1959) can at times be found. The form of the stimulus generalization gradient has been found to be Gaussian in nature in the majority of studies (Ghirlanda & Enquist, 2003) although there are some theories that proposed an exponential form (Shepard, 1987) and others that suggested that both forms might be possible (Staddon & Reid, 1990). In terms of symmetry, the standard stimulus generalization gradient is thought to be symmetrical, but there are two areas where asymmetries have been found. The first possible asymmetry is the location of peak response, described earlier as peak shift. The second is where the slope on one side of the stimulus generalization gradient is steeper than the other, leading one side of the gradient to have a larger area. This is sometimes called area shift (Rilling, 1977). While these perspectives are generally considered and analyzed separately, there have been new techniques that allow more complex analysis like cluster analysis (Zaman et al., 2019) or Bayesian parameter estimation (Lee et al., 2021).

The exploration of the factors that underlie stimulus generalization is an ongoing process. Beside the strict, operational definitions described above, researchers have endeavored to relate the physical phenomena to conceptual or cognitive roots. The psychophysical conceptual framework suggested that the rate of responding was a combination of the organism's sensitivity to variation along the test dimension, and their sensitivity to the threshold for responding (Honig & Urcuioli, 1981; Lashley & Wade, 1946). This view was supported by work done with pigeons. It was found that there was a correspondence between features of the pigeon's sensitivity for hue discrimination (Wright, 1974) and variations in stimulus generalization gradient features (Blough, 1972; Jitsumori, 1978). However, these are not the only factors that impact stimulus generalization. Existing conditions before exposure to a stimulus, called establishing operations or motivating operations (e.g. food and or water deprivation of the organism, or pre-exposure to a stimulus that is later punished, etc.; Laraway et al., 2003; Vervliet et al., 2010), have also been found to lead to changes in generalization gradient height and slope angle (Lotfizadeh et al., 2012). On the other end of the spectrum, consequences (e.g. reinforcers and punishers) have also been found to lead to shallower generalization gradients, with punishers leading to the shallowest gradients and the most generalization (Schechtman et al., 2010). In addition, the magnitude or intensity of punishment following a stimulus can also impact stimulus generalization. Volunteers who were exposed to higher magnitude aversive stimuli had greater generalization than individuals exposed to lower magnitude aversive stimuli (Dunsmoor et al., 2017). Training length has also been found to impact generalization gradients. Studies in both humans using time-compressed speech (Banai & Lavner, 2014) and with bees using visual patterns (Stach & Giurfa, 2005) have found increased responses to similar stimuli, as well as broadening of the range of generalization.

There has also been work looking at individual differences. One study was able to find a weak relationship between behavioral type and red junglefowl chicks' stimulus generalization between novel stimuli (Zidar et al., 2019). A behavioral type, or coping style, is “a coherent set of behavioral and physiological stress responses which is consistent over time and which is characteristic to a certain group of individuals” (Koolhaas et al., 1999). The behavioral type can also be thought of as a categorization method for individual differences within a population for organisms. In humans, a significant amount of research has gone into exploring the impact of cognitive rules on stimulus generalization (Lee et al., 2018; Lovibond et al., 2020; Maes et al., 2015; Wong et al., 2020; Wong & Lovibond, 2020; Zaman et al., 2022). In brief, two major rules were identified within human subjects; the ‘linear’ rule and the ‘similarity’ rule. Subjects who followed the ‘linear’ rule, engaged with the task in a comparative fashion (e.g. the more red the target is, the more likely a reward will be given), and were more likely to show a monotonically increasing generalization gradient, like those seen in intensity-based stimulus dimensions. Subjects who followed the ‘similarity’ rule engaged with the task in a matching fashion (e.g. the more similar the target is to this exact hue of red-yellow, the more likely it is that a reward will be given) and were more likely to show the more standard Gaussian shaped stimulus generalization gradient. This rule-based or inductive response was also found to be distinct from the ability of the subjects to differentiate between different conditions (Lovibond et al., 2020).

Stimulus generalization is a fundamental phenomenon that is integral to many, if not most behavioral experiments. For example, any study that uses differential reinforcement of successive approximation to teach animals to interact with an operant apparatus is leveraging stimulus generalization to set up the experiment. For this reason, it is important that this form of stimulus control is considered when adapting the ETBD to be able to recognize discriminative stimuli. Due to the complex factors that influence the form, slope, and shape of stimulus generalization gradients, it is unlikely that the ETBD would be able to account for all possible phenomena that influence them. However, since the ETBD is a complexity theory, there is always the possibility of emergent phenomena that cannot be easily derived from basic rules of the ETBD. The next section will discuss the ETBD itself and the rules that underlie its predictions.

Evolutionary Theory of Behavioral Dynamics

The ETBD is a behavioral theory that utilizes artificial organisms (AOs) to generate its predictions. These AOs interact with a simulated environment by emitting behaviors that may elicit reinforcement from the environment. The presence or absence of reinforcement influences the behavioral repertoire of the AOs for the next generation. A genetic algorithm (McDowell, 2004) is used to determine how the repertoire changes based on environmental feedback for each time tick, or generation. The genetic algorithm created by McDowell is based on selectionist principles (i.e., selection, reproduction, and mutation) and was designed to be an computational instantiation of selection by consequences (Skinner, 1981). In the following paragraphs, the support for the theory's predictions will be briefly discussed, followed by a description of the AOs, the underlying rules of the genetic algorithm, and a brief discussion of the environment in which the AO's behavior is modeled. The support for the theory's predictions are not meant to be a comprehensive overview. For a more complete review of the theory's current state, see (McDowell, 2019).

The pattern of behavior emitted by an AO cannot be predicted by mathematical manipulations of the underlying rules, which means that the ETBD is a complexity theory (McDowell & Popa, 2009). Instead, the patterns of behavior can be considered emergent phenomena that are generated by the repeated operation of the low-level rules (e.g. selection, replication, etc.). The behavioral records of AOs during an experiment represent the predictions of the theory; these records have been compared with behavioral records of live organisms.

The ETBD has been found to accurately predict the behavior of live organisms in many environments (McDowell, 2004; McDowell, 2013). A few examples of these confirmed predictions include the behavioral adaptations to rapidly changing reinforcement rate ratios in pigeons (Kulubekova & McDowell, 2013; McDowell & Klapes, 2020) and monkeys (Chi, 2019), behavioral features during extinction in rats (McDowell, 2013), the distribution of time between responses under low response cost situations (Kulubekova & McDowell, 2008), continuous choice behavior on concurrent random-interval random-interval (RI RI) schedules (McDowell et al., 2008), concurrent RI RI schedules with covarying rates and magnitudes of reinforcement (McDowell et al., 2012), concurrent random-ratio random-ratio (RR RR) schedules (McDowell & Klapes, 2018), behavioral responses to punishment in concurrent RI RI schedules (McDowell & Klapes, 2019), behavioral responses to changes in reinforcement context (McDowell & Klapes, 2020), and changeover rate pattern in concurrent schedules as reinforcement ratios vary (McDowell, 2013). One particular paper on the ETBD merits a more detailed review due to its relevance. McDowell and colleagues (2006) reported on a possible theoretical method of generating stimulus control phenomena in chain schedules. This paper will be discussed in greater detail in a later section.

The Artificial Organisms. In the simplest terms, an AO is a simulated creature that emits behavior. The pool of possible behaviors that the AO can emit is defined by a population of 100 behaviors that are contained within that AO. Behaviors in the ETBD are defined abstractly, instead of having concrete physical definition (e.g. a number value instead of ‘raising a paw’). Furthermore, each behavior can be represented in two ways. The first is by its genotype, which is a string of 10 binary digits. The second is by its phenotype, which is the decimal integer value of the behavior's genotype. For example, a behavior with the genotype "0000000100" would have the phenotype, "4." There are 1024 possible behavioral phenotypes for the AOs based on the limitation of the 10 digit binary genotype. When an AO is first created, 100 behaviors are chosen at random from the phenotype range, 0 to 1023.

The Rules of the ETBD. Time in the theory is represented by ticks or generations. During one generation, an AO starts with an initial population of behaviors that is subjected to all the rules of the ETBD. This produces a new population of potential behaviors for the next generation, and the process repeats. The rule set is as follows:

1. Emission:

   At the start of a generation, the AO *emits* a behavior at random from its current population and the experimental environment produces a reinforcer if the emitted behavior meets the criteria set by the schedule.
2. Selection:

   If a reinforcer is …

   a. not obtained, then behaviors in the population are chosen at random to become 'parents' until 100 pairs of behaviors are chosen.
   b. obtained, then behaviors in the population are probabilistically chosen based on their *fitness*, that is, their numerical closeness to the emitted behavior in phenotypic space. The probability that a particular behavior will be chosen to be a parent is based on a linear probability density function that is centered on the emitted behavior that led to reinforcement. The linear probability density function is

$$p(x) = -\frac{2}{9\mu^2}x + \frac{2}{2\mu}, \qquad (1)$$

where *p(x)* is the probability that a particular behavior will be chosen, *x* is the absolute value of the difference between that behavior's phenotype and the emitted behavior's phenotype (i.e., the behavior's fitness), and μ is the mean of the density function. Using a small μ leads to higher selectivity (i.e., parent behaviors are more likely to be close in phenotype to the emitted behavior), while a large μ leads to lower selectivity. Using this method, behaviors in the population are chosen until 100 pairs of 'parents' are obtained.

3. Recombination/Replication:

   The 100 chosen pairs of behaviors are recombined using bitwise recombination at the genotypic level. Each bit in the genotype of a parent behavior is matched with the corresponding bit in the other parent behavior's genotype. Each parent has a 50% chance to have their bit incorporated into the new behavior. The result is 100 new "child" behaviors.

4. Mutation:

   Finally, a proportion of behaviors chosen at random from the new population is subjected to mutation using the bitflip-by-individual method. According to this method, one bit in the behavior's genotype is changed to the opposite state. The proportion of behaviors subjected to mutation is referred to as the mutation rate.

5. The new population then replaces the initial population, and the steps repeat until the program is terminated.

A flow chart of the rules is shown in Figure 5.

The Environmental Components. The environment that the AO interacts with in the standard model is an abstract representation of an operant chamber, with only operandi, reinforcement, and punishment. Operandi are activated when a behavior is emitted within a target class. Typically, behaviors within a target class are continuous in the phenotype space. The continuous aspect of the target class is significant because behaviors that are phenotypically close to one another are considered similar to one another in terms of fitness. The behaviors that activate a particular target class can also be randomly chosen and non-sequential to represent background reinforcement (McDowell & Klapes, 2020). Reinforcement is given directly to the organism without need for additional behavior (e.g. eating a food pellet from a recently opened hopper) once the criteria for the reinforcement schedule are met.

One notable missing component in the environment for the purposes of this study is external stimuli. Although lab operandi and food pellets are often considered stimuli in an experiment with a live organism, an AO has no mechanism to 'sense' that these items are present in its environment, and changes its behavior only based on reinforcement. Adding the ability to detect external stimuli and having changes in behavior occur due to the discriminative stimuli in an AO's environment would allow the exploration of many phenomena that the ETBD is currently incompatible with.

Stimulus control in the ETBD. In a 2006 paper on stimulus control, McDowell and colleagues examined one possible avenue for incorporating discriminative stimuli into the ETBD to allow the theory to predict behavior in chain schedules. Chain schedules are compound schedules of reinforcement in which a sequence of behaviors must be completed before a reinforcer can be delivered. Each part of the sequence also has a unique discriminative stimulus associated with it. Chain schedules have a couple of features that made them incompatible with the original ETBD. First, the original ETBD did not have a way to define a discriminative stimulus. Second, the unreinforced parts of the sequence would never experience any form of reinforcement, so no learning could occur. To overcome the first point of incompatibility, McDowell et al. had each discriminative stimuli represented by a different behavioral population. The AO was able to move along the sequence by emitting behavior within a designated target class, which satisfied the conditions for changing discriminative stimuli. To overcome the second point of incompatibility, they used the principle that "discriminative stimuli associated with reinforcement themselves acquire reinforcing properties." They instantiated this principle by (i) using the reinforcement magnitude of the reinforcer at the end of the sequence as a base, (ii) generating an associative strength value using the associative learning process described by the Rescorla-Wagner theory (a value between 0 and 1), (iii) assigning the discriminative stimuli a reinforcing value based on the reinforcement magnitude of the subsequent link divided by associative strength. For example, in the terminal link where the AO gets the reinforcer, the reinforcement magnitude would be unchanged. In the second-to-last link in the sequence, the reinforcement value would be the reinforcement magnitude in the last link divided by the associative strength. The reinforcement value in the third-to-last link in this chain would be based on reinforcement magnitude in the second-to-last link, divided by the associative strength, and so on.

This form of stimulus control in the ETBD was primarily focused on chain schedule phenomena and it would be difficult to use to explore other forms of stimulus control. In this system the AOs would, by definition, recognize all discriminative stimuli as unique. There was no way to look at AO responses to "similar" stimuli as required to explore stimulus generalization gradients, or to have confusion between one stimuli and another. AOs using this system would also only recognize one stimulus at a time. This methodology effectively sidesteps the process of parsing and recognizing a stimulus as discriminative, excluding the exploration of phenomena like blocking, where one discriminative stimulus prevents the learning of another discriminative stimulus.

In summary, the form of stimulus control described in the 2006 paper linked discriminative stimuli effectively with reinforcement, but did not explore how discriminative stimuli are learned, and did not have a framework from which to show relationships between stimuli. Stimulus sampling theory is an alternative method of looking at the interaction between stimuli and behavior that could be a solution to the two problems described above, and will be discussed further below.

Stimulus Sampling Theory

Stimulus Sampling Theory (Estes, 1950) is an early example of a mathematical learning theory (Howard, 2014). Through SST, Estes sought to describe behavioral phenomena using the "statistical properties of environmental events" instead of using cognitive constructs within the organism (e.g. memory traces, internal clocks). SST was meant to address two concerns. The first is the difficulty of falsifying hypotheses based on cognitive constructs. Estes (1955) said the following about cognitive, or "hypothetical" constructs:

> The difficulty with this type of construct is that it is always much easier to postulate than to unpostulate. Few hypothetical entities are so ill-favored that once having secured a foothold they cannot face out each new turn of empirical events with the aid of a few ad hoc assumptions. (p.145)

To avoid the pitfall described above, SST operationalizes each assumption using “experimentally manipulable variables” that operate within a discrete time framework. The consequences of the assumptions unfold using “strict mathematical reasoning” (Estes, 1950). Through the use of experimentally manipulable variables, like stimuli, a quantitative relationship between the independent variables and the dependent variables can be formed. The second concern that SST addresses is parsimony. Cognitive constructs, due to their flexibility, can be used to explain vast arrays of behavioral phenomena. However, some of the attributed results may be potentially explained by environmental stimulus variables that are already present in all behavioral situations (Estes, 1955). Estes sought to prove that many behavioral phenomena did not require cognitive constructs to be understood, and could be explained using an abstract environment composed of ‘stimulus elements’ and simple rules of interaction.

The SST model was primarily used to derive mathematical relationships between stimuli and behavioral responses in classical conditioning and instrumental learning experiments. In the initial paper, Estes (1950) sought to explain basic learning phenomena: learning to associate a buzzer and shock, reductions in latencies during travel in simple mazes, and lever-pressing behavior by rats (Bower, 1994). Various versions of SST were used to explain probability learning in humans (Estes & Johns, 1958; Estes & Lauer, 1957), forgetting and spontaneous recovery (Estes, 1955), positive reward in human learning (Estes, 1969b), and punishment (Estes, 1969a).

SST is conceptually based on the notion that although any organism is at every moment exposed to multiple stimuli, only a subset of the stimuli present will impact the organism’s behavior. That subset is a randomly drawn sample from the entire population of possible stimuli. Once that moment passes, the stimuli that were present in that sample will be replaced with new stimuli for the next observation period. SST also organizes an organism's behavior into response classes. Response classes are groups of behaviors that are identical in terms of their environmental impact. For example, a rat may push a lever down in any number of ways, but all of those ways would be considered a part of one response class by SST.

Stimulus elements are always associated with only one response class at a time. During each experimental trial, one response is emitted. A behavior is chosen probabilistically based on the proportion of stimulus elements associated with it within the sampled stimuli. For example, consider a situation where five stimulus elements are sampled, each with an equal chance of occurring. Three of the five stimulus elements are associated with head scratching while the other two are not. In this situation, there is a sixty percent chance that the organism will scratch its head. During a learning or training trial, the response class is considered to be forced or induced on the subject by the experimental design. After a response is emitted, all stimulus elements present in the sample are considered conditioned to the specific response class that was just emitted, and unconditioned from any other response classes.

The theory hypothesizes that the following sequence occurs during every experiment:

1. Initially, there is a pool of stimulus elements (S) that is comprised of all possible elemental stimuli within a given environment. These elements are currently not associated with any of the conditioned responses under observation (although there may be other unrelated response classes already present).
2. At each time tick within an experiment, a subset of S is sampled randomly.
    a. If this occurs during a learning trial or a training period, the behavioral response is forced to be within the correct response class. All of the sampled stimulus elements are changed to be associated with the response class that is currently being trained.

  b. If this occurs during an experimental trial, the probability of the behavior coming from a particular response class is proportional to the number of stimuli in the sample that are associated with that response class. All of the sampled stimulus elements are changed to be associated with the response class that was emitted during this trial.

3. Repeat step two until the end of the experiment.

A visualization of the steps is presented in Figure 6. It is important to note that the steps described above are not used iteratively as they are in the ETBD. The reason why these steps are not used in an iterative fashion is unknown, but likely due to the fact that SST was created well before the advent of personal computing. The principles and steps described above are the principles from which mathematical interpretations are derived. The mathematical forms of the theory will not be described in this dissertation because the principles and steps described above are sufficient for the creation of an iterative computational version of SST.

Principles of SST. Estes breaks down the theory into the following four assumptions:

a. Any environmental situation, as constituted at a given time, determines for a given organism a population of stimulus events from which a sample affects the organism’s behavior at any given instant. In statistical learning theories, the population is conceptualized as a set of stimulus elements from which a random sample is drawn on each trial.
b. Conditioning and extinction occur only with respect to the elements sampled on a trial.
c. The behaviors available to an organism in any given situation may be categorized into mutually exclusive and exhaustive response classes.
d. At any time, each stimulus element in the population is conditioned to exactly one of these response classes. (p.146-147; Estes, 1955).

II. Theoretical Design

In the following sections, benefits of incorporating elements of SST into the ETBD in the present study are reviewed, the manner in which this incorporation was implemented is discussed, and the scope of the present study is described.

Benefits of Incorporating Elements of SST into the ETBD

The ETBD and SST have been used to explain numerous phenomena, but both require further development in order to explain the emergence of discriminated responding under multiple schedules and stimulus generalization gradients.

Of the two models, SST is arguably the more limited. It is capable of handling stimulus variability, stimulus sampling, and stimulus-response associations (Bower, 1994). However, as a purely mathematical model designed before the age of desktop computing, SST is capable only of predicting the average behavior of an organism under simpler circumstances than may be modeled using the ETBD. The response portion of SST is particularly limited. Since all the stimuli present become associated with the reinforced response, it is expected that this would lead to exclusive preference for one response class, even if there are reinforcers for the alternative response class. SST also has multiple known weaknesses and has been found to not sufficiently model experiments involving blocking, stimulus compounding, and the "overlap problem" (i.e. the neutralization of shared elements after further discrimination has occured between two stimuli; Bower, 1994).

The ETBD, in its standard form, is insensitive to discriminative stimuli and focused on operant contingencies. This makes it difficult to apply outside of situations with stable stimuli. As discussed earlier, the inability to account for antecedent, contextual factors severely limits the range of phenomena that can be modeled and accounted for by the ETBD. Most, if not all, multicellular organisms have means of detecting stimuli in their environment and are able to respond to stimuli to receive reinforcement or to avoid punishment. The operant chamber environment in which the ETBD is typically thought to be enclosed within can also have discriminative stimuli which are currently not being utilized by AOs animated by the ETBD. This makes adding the ability to detect and react to discriminative stimuli a natural extension of the ETBD.

There are multiple ways in which the ETBD and SST may be synergistic if used together. Though SST was initially designed as a mathematical learning model, it has the elements required to become a complex system, namely multiple interacting agents (i.e. the stimuli) and simple rules for interaction that are capable of working iteratively, similar to the ETBD. Both theories function using a discrete time scale and have clear rules that generate a record of behavior. Finally, by increasing the complexity of the ETBD with the added stimulus elements of SST, there is the potential for emergent phenomena that cannot be predicted by observing the operation of the base rules alone. These may only be discovered after the model is created and run. It is predicted that the combined model will be able to do the following:

1. Model antecedent-consequence contingencies
2. Model response-reinforcement contingencies
3. Permit AOs to respond differentially between the presence and absence of discriminative stimuli
4. Preserve behavioral patterns while stimuli are absent
5. Improve the model’s animal-like behavior (e.g. potential need for training, similar to live organisms).

Implementation

In this section, the method of incorporating elements of SST into the ETBD is documented along with the rationale. The goal of the implementation is to create a modified ETBD that will be able to meet the six capabilities listed above. The major components of the model (i.e. the environment and the organism) will be described first, followed by their interaction.

The Modified Environment. The AO will be placed in an environment with multiple stimulus elements. The environment consists of the global environment, which contains all stimulus elements, and a local environment, which contains a subset of stimulus elements that can be potentially observed by the AO. If there is a light that changes from red to green within an experimental environment, that will be represented by removing the red light stimulus elements from the local environment and inserting green light stimulus elements into the local environment. If there are neutral stimulus elements within the environment, they will be retained.

The Artificial Organism. As the experiment is initialized, the AO will have no behavioral populations. AOs will have the ability to sample five stimulus elements at random from the local environment. When the AO first encounters a stimulus element it has not yet experienced, the AO will generate an unique behavioral population of 100 behaviors at random for each stimulus, following the recommendations of McDowell et al. (2006). Going forward, this unique behavioral population will be linked to the stimulus the AO was exposed to and used for future interactions. Each behavior will have a decimal phenotype and a 10-digit binary genotype, like the basic ETBD instantiated AO.

Rules of the modified ETBD. The rules of the system are the following:

1. Initialization. The stimulus elements within the global environment will be initialized. A subset of these stimulus elements will be additionally assigned to be part of the local environment. The number of stimulus elements, their initial starting locations (e.g. local or not local), and their labels (e.g. wall, red light, green light, etc) will be determined by the experiment. The name of a stimulus element does not confer any properties to it, but are used to represent how they will be used in the experiment.
2. Emission. The AO will randomly sample five stimulus elements present within the local environment. The sampled stimulus elements will be checked for a linked behavioral population. If the stimulus is found to not have a linked behavioral population, one will be created, until all sampled stimulus elements have a linked behavioral population. The initial behaviors for the populations will be randomly chosen. From the group of linked behavior populations, one behavioral population will be chosen at random. From the chosen behavioral population, one behavior will be emitted at random. If the emitted behavior is in a target or background class, the reinforcement schedules for that class will be checked to determine if a reinforcer is available for the AO to obtain.
3. Selection. Based on the presence or absence of reinforcement the following may occur to all behavioral populations linked to an observed stimulus element:
    a. If a reinforcer is not obtained, then behaviors are chosen at random within each population to become ‘parents’ until 100 pairs of behaviors are chosen.
    b. If a reinforcer is obtained, then behaviors within each population are probabilistically chosen based on their fitness. The probability that a particular behavior will be chosen to be a parent is based on a linear density function

(Equation 5) that is centered on the emitted behavior. Using this method, behaviors in the population are chosen until 100 pairs of ‘parents’ are chosen.

4. Recombination. Within each population, the chosen pairs of behaviors are recombined using bitwise recombination on a genotypic level. Each bit in one parent behavior’s genotype is matched with the corresponding bit in the other parent behavior’s genotype. Each parent has a 50% chance to have their bit incorporated into the new behavior. This results in 100 new behaviors.
5. Mutation. Finally, within each population, a portion of the behaviors are randomly chosen to undergo mutation using the bitflip-by-individual method. In the behaviors chosen to undergo mutation, one bit in the behavior’s genotype is changed to the opposite state.
6. Each new population then replaces the previous population.
7. Based on the experimental design, determine if any stimulus elements will be moved in or out of the local environment. This step is not used in this experiment, but could be used to move stimuli in or out of the AO’s “view,” since the AO can only sample from the local environment.
8. Steps 2-7 are repeated until the end of the program.

A flowchart illustrating these rules appears as Figure 7. The flow chart begins in the bottom left corner, with the initialization of the stimulus elements in the global and local environments. The global environment contains all stimulus elements in the model. In this case, there are five green stimulus elements, five red stimulus elements and five wall stimulus elements, representing a green light, a red light, and neutral environmental elements (e.g. walls, the ceiling, etc.). The local environment contains the stimulus elements that are currently observable by the AO. In this case, only the elements representing a red light and stable stimulus elements are in the local environment. The AO randomly samples five stimulus elements from the local environment. The observed stimulus elements from the local environment (i.e., three red stimulus elements and two gray stimulus elements) are contained within the yellow circle at the bottom left. Each stimulus element is linked with a specific population of behaviors (only two linkages are shown in the figure). One population from the observed environment is chosen at random and one random behavior is emitted from that population. All five populations separately go through the steps in the standard ETBD (i.e., selection, recombination, and mutation) to generate a new population of behaviors. If the emitted behavior leads to reinforcement (S^{r+}), all of the behaviors in their respective populations will go through fitness-based selection, rather than random selection. After the next time tick, there is a check to determine whether the environment will change due to a change in the reinforcement schedule. In this case, a schedule change causes a change in the discriminative stimulus, requiring the red stimulus elements to be moved out of the local environment, and the green stimulus elements to be moved into the local environment. Afterward, the cycle repeats.

In the bigger picture, the modified ETBD attempts to preserve the functionality of both the ETBD and the SST as much as possible. For the ETBD, all the basic processes of the ETBD (i.e., emission, selection, recombination, and mutation) are preserved, but extra behavioral populations and their relationships to stimulus elements are added. For the SST, the random interaction with the local environment was maintained, but the behavioral population replaced the simpler response class. This is considered a modified ETBD and not a modified SST or a modified SST-ETBD due to the fact a majority of the ETBD and the driving function of its predictions (the evolutionary principles), are still intact, while only a component of SST is maintained.

The Present Study

In the present study, I aimed to achieve the following objectives:

Aim 1: *Implement a version of the ETBD with the stimulus elements of SST.*

This form of stimulus sampling has never been combined with the ETBD previously. Using the implementation outlined above, a modified ETBD algorithm was developed. The functionality of the algorithm was tested using a set of concurrent schedules to verify that the algorithm maintains the same capabilities as the original ETBD (Experiment 1).

Aim 2: *Demonstrate that AOs animated by the modified ETBD can distinguish between different stimuli and modulate their behavior similar to living organisms.*

AOs were tested using a multiple RI EXT schedule, to determine if it was able to develop discriminative responding, similarly to live organisms (Experiment 2).

Aim 3: *Implement stimulus generalization through manipulation of environmental stimuli.*

AOs will be placed into environments that have been known to demonstrate stimulus generalization gradients. (Appendix A).

III. General Methods

Subjects, Apparatus, and Materials

AOs animated by the modified ETBD were evaluated in a three-experiment series. Dr. J. J McDowell originally conceptualized and developed the source code for the ETBD and the experimental environment in VB.net. Other contributors to the ETBD codebase include Olivia

Calvin, Bryan Klapes, Saule Kulubekova, Andrei Popa, Ryan Higginbotham, and Steve Riley. The modified ETBD was implemented in Python 3.7.11 using the Spyder 5 IDE. Python libraries used include Scipy, Numpy, Numba, PANDAS, OS, Multiprocessing, Time, Glob, Pathlib, Argparse, Json, Sys, Statsmodels, and Collections. Simulations were run either on 1) Windows 10 with an Intel Core i7-5600U CPU at 2.60Ghz with 16GB RAM for individual simulations or 2) Amazon Web Services Linux Servers (Oregon-west, instance type c5a.8xlarge with 32 CPUs) for larger multiprocessing batch simulations. The modified ETBD is publicly available at https://github.com/CyrusChi/pyETBD-SST. All of the experiments ran thirty AOs through each testing condition, except for one part of Experiment Three, Concurrent Schedules Review. This part of experiment three was run using only 10 AOs for each condition in order to reduce computation time. See Tables 1, 2, and 3 for summaries of the experimental parameters.

IV. Experiment Series One

Experiment Series One Methods

Phase One. The first phase was designed to test the modified ETBD AOs using the exact procedure from McDowell et al., 2008. The simulated environment contained two targets; target one and target two. The two targets ranged from 471 to 511, and 512 to 552 in the phenotype space, respectively. The AOs were tested on 11 concurrent schedules: RI 20 RI 120, RI 30 RI 110, RI 40 RI 100, RI 50 RI 90, RI 60 RI 80, RI 70 RI 70, RI 80 RI 60, RI 90 RI 50, RI 100 RI 40, RI 110 RI 30, and RI 120 RI 20. The reinforcement magnitude for both targets was 40. The experiment followed the same procedure as phase one in the canonical ETBD concurrent schedules experiment (McDowell et al., 2008). There was only one stimulus element present in the local environment, with one linked behavioral population with 100 behaviors.

Phase Two. The second phase was designed to determine the impact of adding additional stimulus elements. AOs were tested with different numbers of stimulus elements present in the local environment. All stimulus element quantities between one stimulus element to 15 stimulus elements were tested. The organism sampled five random stimulus elements from the local environment per generation during the observation step. All other components of the testing environment remained the same as phase one.

The first experiment in phase two highlighted a specific challenge; as the number of stimulus elements increased, the AO's sensitivity to reinforcement decreased below the levels expected for live organisms. In order to identify the issue leading to the decreased reinforcement sensitivity, it was important to consider how increasing the number of stimulus elements affects the basic functional loop of the ETBD. First, the observed stimulus element and its linked behavioral population are chosen at random. Theoretically, the random observation method weakens the relationship between the last emitted behavior and the next emitted behavior since they can be drawn from completely independent behavioral populations. The more stimulus elements present in the local environment, the more likely an unrelated stimulus element will be chosen. Second, the increase in the number of stimulus elements multiplicatively increases the number of behaviors in the total pool and reduces the impact of any one reinforcement event. This could potentially increase the amount of variation in the behavior emitted. In order to compensate for these issues, two changes were implemented and are discussed further in the two sections labeled modification one and modification two.

Modification One. While random selection of stimulus elements is a critical part of SST, it likely weakens the AO's ability to respond to rapid changes in reinforcement. One way to strengthen this response is to create a bias in which stimulus elements are chosen based on which linked behavioral population has been most impacted by reinforcement. This can be done using entropy. As first described by Shannon (1948), entropy is a measure of uncertainty within data. When data are evenly distributed across a range, then the entropy will be at its highest. Conversely, if the data are all clustered at exactly at the same point, then the entropy will be at its lowest. When used on a behavioral population, the level of entropy corresponds roughly to how clustered the behavioral population is around a certain set of phenotypes.

For modification one, in order to calculate the entropy at a relevant resolution, the phenotype space was first broken into 25 bins with 40-41 phenotypes per bin. This was chosen in order for a target class to be contained within no more than two bins. First, for AOs with this modification implemented, five stimulus elements were chosen at random from the local environment as usual. Next, the linked behavioral populations were separated into 25 bins equally divided over the phenotype space. Finally, resulting histograms were normalized and the Shannon entropy for the linked behavioral populations was calculated using the entropy function from the scipy.stats python library. The equation for entropy is

$$\mathrm{E} = -\sum_{n=1}^{\infty} (a_n * \log(a_n)), \tag{2}$$

where a_n is the number of behaviors in each bin. The stimulus elements which have entropies that are within a certain percentage (2% and 5% for the current testing groups) of the lowest entropy were used as the observed pool of stimulus elements. This process will be called entropy-based observation X% for the rest of the document. All other elements of the modified ETBD were run as previously described.

There are conceptual and technical reasons to use entropy to produce bias in observation. It is already known that rewards associated with visual stimuli can increase attentional behavior towards those stimuli (Chelazzi et al., 2013; Cowie et al., 2020). In the modified ETBD, a frequently rewarded behavioral population will likely have behaviors clustered around the rewarded phenotypes and thus have lower entropy compared to non-rewarded stimulus elements. Due to this phenomenon, entropy can be used as a proxy for bias in attentional behavior. Using entropy also allows information within the model to be used formulaically to determine the chosen observation targets. This reduces the need for additional variables to be set by the experimenter.

Entropy as a concept has not been used often in experimental psychology although it has been used more regularly in neuropsychology. Entropy has been used to create a theoretical explanation of consciousness (Mason, 2019), neuronal branching (Islas et al., 2020), neuron networks (Song et al., 2021; Viol et al., 2019), and as a component of fMRI image processing (Akdeniz, 2017; Gupta et al., 2017; Somwanshi et al., 2016). The difficulty with using entropy in experimental psychology is likely related to the following assumptions: First, data needs to be transmitted in many individual units rather than in a single unit. Second, the elements of the data need to be unstructured and functionally interchangeable (Luce, 2003). Fortunately, the behavioral populations within the ETBD meet these assumptions thus paving the way for the use of entropy.

Modification Two. The second potential challenge caused by the increase in the number of stimulus elements is the increase in emitted behavior variation. One way to reduce the variation in emitted behavior is by increasing the number of behaviors in each population. This should theoretically lead to an increase in the regularity of the behavioral population in the same fashion that increasing the sample size of an experiment can reduce the variability found in experimental results. Since AOs run with entropy-based observation 2% showed the most animal-like matching sensitivity levels in during modification one, that observation style was used in this experimental set, and in the experiments going forward. AOs were run with entropy-based observation 2%, ten stimulus elements were contained in the local environment, and the behavioral population size was varied starting with a population of 50 behaviors and ending with a population of 500 behaviors, incrementing 50 behaviors at a time. All other elements of the modified ETBD were run as previously described and following the experimental procedure described in phase one.

Data Analysis for Experiment One

The logarithmic transformation of the power function matching law,

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

was fitted to the response rate ratios (B_1 / B_2) and reinforcement rate ratios (R_1 / R_2) from each AO, and the percentage of variance accounted for was calculated. Equation 3 was found in previous studies to is known to describe concurrent schedule performance in human and animals well (McDowell et al., 2008). The residuals were examined for trends using the random cubic trend test (RCTT, McDowell et al., 2016).

Experiment Series One Results

Phase One. This experiment was designed to test the ability of the modified ETBD code to reproduce basic ETBD functionality. The values obtained after Equation 3 was fitted to the response and reinforcement ratios: (i) exponents *a*, (ii) bias parameters *b*, (iii) the proportion of the variance accounted for, and (iv) the r squared values of the residuals for the RCTT are listed in Table 4 for all 30 AOs. The average of the 30 least squared plots of Equation 3 is shown in Figure 9. The results in Table 4 show that the matching equation provided an excellent description for the steady-state behavior of the AOs, explaining between 96% and 99% of the variance, with an average of 98% of the variance accounted for. The residuals were fitted with a cubic polynomial. The r^2 from the best fitting cubic polynomial was compared with the critical r^2 value for cubic polynomials ($n = 11$), 0.651 (Table 2 in McDowell et al., 2016). The residuals for all 30 AOs had substantially smaller r^2 values than the critical r^2 value, which indicates the absence of trends in their residuals. As shown in Table 4, the exponent, *a,* varied from 0.99 to 0.80, with an average of 0.88 which demonstrates a mild degree of undermatching. The bias parameter, *b*, varied from 0.91 to 1.05, with an average of 0.99, indicating that there was little to no bias in the responding on the concurrent schedules. These results indicate that the AOs animated by the modified ETBD (i) generate behavior that is fitted well by the matching law equation (ii) with parameters that are comparable to those obtained from experiments with live organisms.

Phase Two. This experiment was designed to determine the impact of additional stimulus elements on AO behavior. The results of the baseline experiment and modification one are shown in Figure 10, and in Tables 5, 6, and 7. In Figure 10, the ‘base’ style (indicated by triangular markers) is when five stimulus elements were chosen at random from the local population. The ‘EN05’ style (indicated by square markers) is when five stimulus elements were chosen at random and the stimulus element with the lowest entropy and any stimulus elements with an entropy within 5% of the lowest entropy were chosen. The ‘EN02’ style (indicated by circular markers) is when five stimulus elements were chosen at random and the stimulus element with the lowest entropy and any stimulus elements with an entropy within 2% of the lowest entropy were chosen. As seen in Figure 10, matching sensitivity of the ‘base’ style of observation quickly dropped to near indifference levels, and stabilized after four or more stimulus elements were used at around 0.5 sensitivity. The ‘EN05’ and the ‘EN02’ styles both showed a trend towards lower sensitivity levels as the number of stimulus element increased, but the reduction in sensitivity as the number of stimulus elements increased was substantially slower than in the ‘base’ style. At ten stimulus elements, the ‘EN05’ style had an average sensitivity of 0.68 and a SEM of 0.005, while ‘EN02’ had an average sensitivity of 0.75 and a SEM of 0.006. The bias parameter, *b*, fluctuated at or near 1 and the PVAF was over 95% for all observation styles and all quantities of stimulus elements tested.

Modification two explored the impact on behavior population size on the matching sensitivity. The relationship between behavior population size and matching sensitivity is shown in Figure 11, and Table 8. As shown in Figure 11, the sensitivity increases with population size until the population size reaches 200 behaviors, and then transitions to a slower upward trend in sensitivity for populations with 200 behaviors to populations with 500 behaviors. Since the 90% confidence intervals were found to be marginally overlapping, it is uncertain if the upward trend is significant. The bias parameter, *b*, fluctuated at or near 1 and the PVAF was over 96% for all behavior population sizes tested.

In addition to the observation styles mentioned above and the behavior population size, other potential modifications were also tested, including: reinforcement magnitude, linear minimum (the lower limit for the number of behaviors that can be potentially chosen as parents in the case of reinforcement), and changing the number of stimulus elements initially chosen at random during observation. These alternative modifications were not found to have any impact on AO matching law sensitivity parameter levels.

The first experimental series found that, for environments containing ten stimulus elements, the ‘EN02’ observation style combined with a behavior population greater than 200 were likely necessary for animal-like behavior from AOs animated by the modified ETBD. The same values obtained after Equation 3 was fitted to the response and reinforcement ratios are listed in Table 9 for all 30 AOs. The average of the 30 least squared plots for AOs using the conditions mentioned above are shown in Figure 12. The results shown in Table 9 reveal that the matching equation (a) was an excellent fit for the steady-state behavior of the AO, and (b) accounting for between 94% and 100% of the variance, with an average of 99% of the variance accounted for. Comparing the r^2 from the best fitting cubic polynomial to the residuals, with the critical r^2 for detecting a systematic trend ($n = 11$), 0.651, the experimental r-squareds are substantially smaller. Thus all 30 AOs passed the RCTT, indicating an absence of trends in their residuals. As shown in Table 9, the exponent, *a*, varied from 0.86 to 0.71, with an average of 0.80 which demonstrates a moderate degree of undermatching. The bias parameter, *b*, varied from 0.95 to 1.07, with an average of 1.00, indicating that there was little to no bias in the responding on the concurrent schedules. These results indicate that AOs animated by the modified ETBD using the ‘EN02’ observation style, containing a 200-behavior population size, and in an environment with 10 stimulus elements are able to generate behavior that is fitted well by the matching law with parameters that are comparable to the parameters obtained from experiments with live organisms.

Due to the increasing computational cost that comes with increasing the behavior population size, the smallest behavior population size (200) that showed animal-like matching sensitivity values was used for further testing.

Experiment Series One Discussion

The phase one result shows that the modified ETBD is able to replicate the finding from McDowell et al. (2008) on concurrent schedules with only one stimulus element and one behavioral population. The sensitivity values obtained are slightly higher, but within the expected range. As a proof of concept, the modified ETBD with additional stimulus elements was able to show animal-like behavior after the changes implemented by modifications one and two. The ETBD as it was originally conceived was designed to have only one behavioral population which responded to reinforcement obtained from the environment. The continuity between recent reinforcement and the features of the behavioral population was likely broken by adding additional behavioral populations and using random sampling to choose between them. Adding entropy to the random sampling as a stimulus selection method was able to restore at least some of the continuity that was originally lost. It was also found that a very selective entropy function, one that selected between the top 2% of the behavioral populations present, was necessary to have an acceptable sensitivity level when ten stimulus elements were present in the local environment. Increasing the behavior population size to 200 was also used to increase the sensitivity level further. This series of modifications, while successful in obtaining animal-like sensitivity levels, points to the need to better characterize the relationship between matching sensitivity, the entropy function, and the population size.

As discussed earlier, Shannon entropy was designed to reflect the amount of information within a system. In communication theory, “information is a measure of one’s freedom of choice when one selects a message” (Weaver, 2017). For an AO, the choosing of a behavior to emit would be its message selection. For example, if the behavioral population within an AO consisted of 100 behaviors but all of the behaviors were of phenotype ‘500,’ then the Shannon entropy for the population would be zero or nearly zero, since there is essentially only one choice for emitted behavior. As the number of possible options increases, the entropy for the behavioral population will also increase. When sorting through different stimulus elements and their linked behavioral populations, the ones with the lowest entropy would be the ones with the most biased behavioral populations. Within the ETBD, the most likely way to bias a population is through reinforcement. Therefore, the entropy of a population can be used as a measure of how reinforced that population was in the past.

Entropy, as it has been implemented in this project, has both strengths and weaknesses as a measure of reinforcement. One of the major strengths is that entropy allows information already contained within the organism to determine which stimulus element will be chosen. This is in contrast to creating an external counter, or some other ‘reservoir system’ (Catania, 2005) to determine which stimulus element should be chosen next. For example, if we are looking to keep continuity between behaviors and reinforcement, we could conceivability create a standing tally of which stimulus elements have had an emitted behavior that led to reinforcement and increase the possibility of choosing stimulus elements with a higher tally value. However, this method would essentially be capitulating and saying that the neural instantiation of the ETBD in live organisms does not handle such functions and requires an external mechanism (e.g. a cluster of neurons that do nothing but count) in order to generate this kind of behavior. Based on the results of this study, this capitulation would be premature.

The current implementation of entropy depends on target size and is indifferent to the location of targets. Part of the entropy function bins all of the behavior in phenotype space into bins with 40 phenotypes each. This is designed so that all behaviors within a target class fall within at most two bins, causing a minimum entropy level when all behaviors are within a target class. This design specification (i.e. the 40 phenotype bin size) could potentially impact target classes that are smaller than half the size of the bin, since multiple target classes could potentially fall into a single bin. If this occurs, the entropy value will not change if behaviors are clustered around just one target class, or spread between two target classes within one bin. Target classes larger than the bin size would likely have a lesser effect on the entropy comparison because the proportion of behaviors can still be captured correctly. However, the baseline entropy will likely be higher because more than two bins can be filled by one target class. That being said, target classes larger than 40 are unlikely to be optimal (Li et al., 2018) and have yet to be used with the ETBD, so the impact of this limitation is likely low. It is likely a good practice to have the entropy bin size be equal to the smallest target class size in order to prevent the over stacking scenario described above.

Another feature of this implementation of entropy is that it does not take into account the adjacency of the bins. For example, if all behaviors were contained within two bins, the entropy value would be the same no matter where the bins were located in the phenotype space. For this particular study, this issue did not impact the ability to generate animal-like behavior. However, there may be an unwanted interaction to occur between entropy and various target locations. It has been found that the hamming distance between targets is computationally equivalent to a changeover delay (Popa & McDowell, 2010). Due to this, in a three target scenario, it is possible for two reinforced targets to have identical entropy values while having different theoretical changeover delays from the third target class. When designing future experiments using entropy and more than two targets, this interaction may need to be managed.

As seen in Figure 10, there is gradual decrease in sensitivity as the number of stimulus elements increases for both entropy conditions. This could be due to the random sampling function. During each generation of the emission step, the AO would randomly sample five stimulus elements from the local environment. When the local environment contains five or less stimulus elements, all stimulus elements are chosen. When the local environment contains six or more stimulus elements, the chances of the organism selecting the most relevant stimulus element begins to decrease, and continues to decrease as the number of stimulus elements increases. It is possible that this phenomenon would disappear if the number of stimulus elements sampled each generation were increased, but further study is required to confirm such a hypothesis.

The increase in the number of behaviors in the behavioral populations was also found to increase the matching sensitivity of AOs. As shown in Figure 11, the rate of sensitivity grows quickly until a 200 behavior population size, and then seems to trend upward at a much slower rate for populations with between 200 and 500 behaviors. Interestingly, this is different from how the unmodified ETBD responds based on previous unpublished data as described by J. J McDowell (personal communication, October 18, 2022). In the unmodified ETBD, the matching sensitivity increases as the behavioral population increases and does not seem to reach an asymptote even at 600+ behaviors. In this way, the modified ETBD may be more robust than the original instantiation because matching sensitivity in the modified ETBD is not affected by changes in population size, as long as the population size is 200 or above. The decreased sensitivity in populations under 200 could be potentially due to statistical fluctuations due to small population sizes. Once the population size is great enough, the AO’s matching sensitivity becomes more stable. While it is unknown how behavioral population size directly links to its neural instantiation within the brains of live organisms, having a range of possible behavioral population sizes improves the robustness of the theory.

V. Experiment Series Two

Experiment Series Two Methods

This experiment was designed to evaluate the impact of using unique stimulus elements in multiple (mult) schedules. After a training phase using a single RI 10 schedule, AOs were run on a mult RI 70 EXT schedule. The target class was 40 consecutive integers in the phenotype space, in the same position as target one in the first experiment. There was a pool of fifteen stimulus elements in the global environment. Five stimulus elements represented the presence of a green stimulus, another five stimulus elements represented the presence of a red stimulus, and another five stimulus elements represented neutral elements in the environment. At any one time, there were a total of 10 stimulus elements in the local environment. Depending on the schedule, the red and green stimulus elements were either present or not present, while five neutral stimulus elements were always present. As the experiment switched between schedules, the stimulus elements for the stimulus that was designated to be “on” were moved into the local environment, and the stimulus elements for the stimulus that was “off” were moved out of the local environment. Modifications one and two were implemented in experiment two, using an entropy-based observation 2% and a behavioral population of 200.

During the initial experiment-two runs, it was found that: 1) behaviors within the target range were higher than expected for the extinction component of the multiple schedule and 2) AOs exhibited extremely fast learning speeds incomparable to living organisms. At this stage, a number of theoretical causes for these phenomena were posited: One possible cause for high rates of behavior during extinction could be the presence of a high operant level, and one possible cause for the fast learning speeds could be the distribution and quantity of stimulus elements. Both of these potential causes were further investigated in modifications three and four.

Modification Three. Modification three was designed to investigate if the high amount of behavior during extinction schedules is due to a high operant level. In the absence of reinforcement, the behavioral population is expected to randomly distribute over the entire phenotype range due to random recombination and mutation. This includes random emission of behavior within the target class (i.e. the operant level), which is a behavior that is not expected of a live organism. This is because live organisms, when not being rewarded for behavior on experimental operandi, will engage in other rewarding behaviors unrelated to the experimental set up. In rodents, this might be grooming, or scratching behavior. In a laboratory setting, this alternate behavior is commonly expected to be a response to non-experimentally driven stimuli or background reinforcement. Recent work by McDowell and Klapes (2020) noted that the ETBD, as it stood at the time, did not account for the principle that "all behavior is choice" and might have difficulty in reproducing the behavior of live organisms during multiple schedules. In order to rectify this, the authors added this principle into the ETBD using a specific background reinforcement methodology.

For this study, background reinforcement was added using a similar methodology. First, a preliminary list of possible background reinforcement target phenotypes was generated. The phenotypes for the two typical target classes were excluded to prevent overlap with experimental operandi. Second, the 60 closest phenotypes above and below the target classes were also excluded. This second step was implemented to reduce the chance that behaviors within the target class would be selected as parents if one of the background behaviors received reinforcement. Using the linear probability density function, we can calculate that the probability for a parent to be chosen reaches approximately zero at approximately three times the reinforcement magnitude. Therefore, if the background reinforcement magnitude is 20, then it would be nearly impossible for a background reinforcement event to draw a behavior from the target class to be a parent. The other background reinforcement magnitudes used in this study are 40 and 60. At those magnitudes, there is some possibility of drawing a parent from the target class during a background reinforcement event. However, this buffer range is still expected to reduce the probability of such behaviors being chosen as parents since the behaviors within the target class will be at least 60 phenotypes away from the reinforced behavior. Third, all potential background behaviors that were within one hamming distance of the two typical target classes were also removed. This was done to reduce the chance that one mutation event could move a behavior in the background class into the target class. Finally, 200 behaviors were chosen at random from the final list to be used as the background target. The background target list was chosen randomly for each AO.

The amount of behavior within the target class in extinction only and background reinforcement only conditions were measured to be used as comparison points. Background reinforcement was delivered on random interval schedules using RIs of 1, 5, 10, 20, 40, 60, 80, 100, 120, and no reinforcement. The reinforcement magnitude for background reinforcement was set at 40.

Concurrent Schedules Review. In order to determine if the background reinforcement impacted the ability of the modified ETBD to demonstrate animal-like behavior on concurrent schedules, the experiment-one schedule set was run with the inclusion of background reinforcement. 10 stimulus elements were present in the local environment. Background reinforcement was set up in an identical fashion to the previous experiment. The random intervals used for background reinforcement were 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, and 120. The reinforcement magnitudes used were 20, 40, and 60 for all reinforcement rates.

Modification Four. Modification four was designed to explore the impact of stimulus element quantity on both (i) rates of behavior within the target class and (ii) the fast rate of learning.

One possible cause for the high amount of behavior during extinction is the presence of the wall (i.e., neutral) stimuli during both the reinforcement and extinction schedules during the multiple schedules. If that is the case, removal of wall stimuli and use of only red and green stimuli should reduce the number of behaviors in the target class during extinction.

Another challenge identified in the second experiment is the fast rate of learning. One possible solution is increasing the number of stimulus elements present. By increasing the number of stimulus elements, the chances that the next behavior emitted comes from a trained behavior population decreases.

In order to explore both of these potential solutions, two experimental groups were designed. One experimental group used X red and X green stimulus elements (where X = 10, 20, 30, 40, 50, 100, 1k, 5k, 10k, and 100k) without wall stimulus elements. In this group, X stimuli of one color were present in each schedule, alternating between red and green. Another experimental group used X wall, X red, and X green stimulus elements (where X = 500, 2,5k, 5k, 25k, and 50k). In this group, each schedule had X wall and either X red stimulus elements or X green stimulus elements. For example, when X = 500, one schedule had 500 wall stimulus elements and 500 red stimulus elements, while the next schedule had 500 wall stimulus elements and 500 green stimulus elements. This allowed the total number of stimulus elements during one schedule to be 1000, which was compared to the previous experimental group's results.

Data Analysis for Experiment Two

The plot of the data was examined for differentiation between the two components of the schedule using basic descriptive statistics. A repeated measures ANOVA and one contrast was conducted on the final results. Based on the Tukey Sidak Bonferroni method (Keppel & Wickens, 2004, p. 119), a family-wise alpha of 0.05 is appropriate for 19 degrees of freedom with two comparisons. Logarithmic equations were fitted to the data describing the impact of background reinforcement on behavior emitted in the target classes.

In order to characterize learning in AOs animated by the ETBD, a metric based on emitted behavior is required. The matching law states that, over the course of a schedule with a specific rate of reinforcement, an organism's behavior is proportional to the reinforcement provided. This relationship between behavior and reinforcement is developed through exposure to the environment so it is to be expected that the relationship between behavior and reinforcement starts completely uncorrelated (unless the organism was previously exposed to the environment) and gradually strengthens with time. Assuming that the reinforcement rate does not change during the course of a schedule, the pattern of behavior should be random at the beginning but should gradually settle into a pattern that is proportional to the reinforcement provided. One method of looking at the behavioral pattern is by combining the data into a histogram. The emitted behavior across 30 AOs was collected in 50 generation chunks and a relative histogram of those behaviors was generated. This histogram was compared to the relative histogram of the overall behavior in the schedule using a Pearson's correlation. See

Figure 8 for an example of a histogram of emitted behavior over 50 generations compared to the histogram of emitted behavior across an entire schedule.

Experiment Series Two Results

This series examined the impact of unique stimulus elements in multiple schedules. The number of behaviors per schedule is shown in Figure 13. The figure shows that behavior during reinforced schedules and extinction schedules remained relatively constant across schedules. During schedules that had the reinforcement schedule present, the average number of behaviors per schedule was 1871.8, 90%CI [1860.90, 1882.83]. During extinction schedules, the average number of behaviors was 796.5, 90% CI [792.72, 800.28]. A repeated measures ANOVA found a statistically significant difference between the behavior of at least two groups ($F(19,551) = 1293.15$, $p < 0.05$). The effect size, calculated as omega squared (ω^2), was 0.976, indicating a large effect. A planned contrast between behavior during all reinforced schedules and all extinction schedules was also performed and a statistically significant difference was found between these two groups ($F(1,551) = 24389.87$, $p < 0.05$). The effect size was 0.468, again indicating a large effect. The behavior within the target class during the extinction schedule was likely caused by random behavioral emissions (i.e., the operant level) that occur during the absence of reinforcement.

Modification Three. This experiment was designed to reduce the operant level by adding in background reinforcement. An experiment was run with no reinforcement present and it had 797.43 average behaviors in one schedule, with 90% CI [784.97, 809.89]. The impact of adding background reinforcement of varying strengths is shown in Figure 14. The impact of background reinforcement rates on average rates of behavior during schedules with reinforcement (circle), and average rates of behavior during extinction schedules (triangle) are shown. The baseline rates of behavior without background reinforcement for reinforced schedules and extinction schedules are represented by a dashed horizontal line and a solid horizontal line, respectively. Logarithmic trends for each group are marked by curved dotted lines. Relative to the no background reinforcement baseline, the decrease in average behavior for reinforced and extinction schedules is similar for both groups, with the reinforced schedule average behavior rates decreasing slightly faster than the extinction schedule behavior. The average behavior during the reinforced schedule had a high of 1668.4 behaviors 90% CI [1657.3, 1679.6] when the background reinforcement was on a RI 120 schedule, and a low of 455.2 behaviors 90% CI [449.1, 461.3] when the background was on a RI 1 schedule. The average behavior during the extinction schedule had a high of 745.3 behaviors 90% CI [741.1, 749.4] when the background reinforcement was on a RI 120 schedule, and a low of 187.1 behaviors 90% CI [184.1, 190.0] when the background was on a RI 1 schedule. The data shows that the addition of background reinforcement reduces the average rate of behavior by approximately 75% for both reinforced and extinction schedules.

Concurrent Schedules Review. Since background reinforcement was not tested during the initial experiment series, the impact of background reinforcement on the matching sensitivity was examined. The relationship between background reinforcement rate and matching sensitivity is shown in Figure 15. The data shows reasonable matching sensitivity levels for a background RI of 20 or above. At reinforcement rates faster than RI 20, the three reinforcement magnitudes diverge. The highest reinforcement magnitude, 20, showed a trend towards increasing matching sensitivity, while the lowest reinforcement magnitude, 60, showed a trend towards decreasing sensitivity. At reinforcement rates slower than RI 20, the results of the three reinforcement magnitudes have overlapping confidence intervals and the matching sensitivity is within the expected range. In addition, the average matching sensitivity, bias, and percentage of variance accounted for were nearly equivalent for all three reinforcement magnitudes, as shown in Table 10. This data shows that the three reinforcement magnitudes used did not have a large impact on the results across the RI 20 to RI 120 background reinforcement range and that the values obtained are similar to the values obtained in the first experimental series. The values of the percentage of variance accounted for was also above 90% and within the acceptable range.

Modification Four. This experimental change was designed to address the rapid speed of learning by either removing wall stimuli or by increasing the number of stimulus elements. The average number of behaviors per condition (reinforcement and extinction schedules) were found to be nearly equivalent for the 10 Red / 10 Green stimulus element condition (‘10/10 no wall’) and the 5 Red / 5 Green / 5 Wall condition (‘5/5/5 wall’), as shown in Figure 16. The reinforcement schedules had an average around 1177 behaviors per schedule, while the extinction schedules had an average around 568 behaviors per schedule. The rates of initial learning for the reinforcement schedules show rapid adaptation in both conditions, with the Pearson’s correlation coefficient reaching a plateau around or slightly above 1000 generations. During the extinction schedules, the ‘10/10 no wall’ and ‘5/5/5 wall’ condition diverge, with the ‘10/10 no wall’ condition showing a similar rapid adaptation pattern similar to the reinforcement schedules, and the ‘5/5/5 wall’ condition showing a noisy but relatively high and stable relationship between the number of generations and the amount of correlation. This data is shown in Figure 17.

As the number of stimulus elements increases in the local environment, in the ‘no wall’ conditions, the number of behaviors per schedule increases more slowly across the reinforcement schedules and decreases more slowly across extinction schedules. At the highest two stimulus element quantities tested, ‘50k/50k no wall’ and ‘100k/100k no wall,’ the increase in behavior during the reinforcement schedules and the decrease in behavior during the extinction schedules was low enough that they did not reach the plateau reached in the other conditions. This data is shown in the top portions of Figure 18 and 19. The rates of learning for the ‘no wall’ condition (the top portions of Figure 20, and 21) showed a high degree of variability between data points. Linear extrapolations were used to show the general trends in the data. The trends show that the rate of learning decreases as the quantity of stimulus element in the local environment increases.

As the number of stimulus element increases in the local environment, in the ‘wall’ conditions, the number of behaviors per schedule increases more slowly across the reinforcement schedules. The rate at which behavior increases is more gradual in the ‘wall’ conditions (shown in the bottom chart in Figure 18) compared to the rate at which behavior increases in the ‘no wall’ conditions (shown in the top chart in Figure 18). Since none of the stimulus element quantities tested converge by the tenth presentation, it is undetermined if they would converge if more schedules were run. During the extinction schedules, as presented in the bottom chart in Figure 19, increasing the quantity of stimulus elements led to an increase in amount of behavior relatively uniformly across all schedules. The three test groups with the largest number of stimulus elements present in the local environment cluster around 700 to 750 behaviors per schedule. The rates of learning for the ‘wall’ condition (the bottom portions of Figures 20 and 21) showed a high degree of variability between data points, like the ‘no wall’ condition. Linear extrapolations were used to show the general trends in the data. The results show that during the first reinforcement schedule, the rate of learning decreases as the quantity of stimulus element in the local environment increases. During the first extinction schedule, the three test groups with the lowest number of stimulus elements in the local environment had lower rates of learning as the quantity of stimulus elements increased. The two groups with the highest stimulus element quantities showed negative rates of learning, indicating that the behavioral histograms actually became slightly less similar to the overall distribution of behavior over the first 2000 generations.

Bringing experiment two together, the most optimal conditions of the ones studied combine a background reinforcement of RI 20 with the ‘5/5/5 wall’ condition. The ‘10/10 no wall’ condition showed no learning during the extinction schedule, which deviates from the expected behavior of live organisms. Increasing the number of stimulus elements in the local environment was successful in slowing down the rate at which behavior decreased across the experiment during extinction schedules, but also suppressed the rate at which behavior increased across the experiment during the reinforced schedules. In addition, increasing the quantity of stimulus elements also increased the amount of variation in learning, leading to much less consistent moment to moment behavior in reinforced and extinction schedules. The background reinforcement level of 20 was used since it was the highest rate of reinforcement (thus having the strongest suppressive effect) that still allowed animal-like behavior on concurrent schedules. The behavior per schedule data is shown in Figure 22. As shown in Figure 22, behavior during reinforced schedules and extinction schedules remained relatively constant across schedules. During schedules that had a reinforcement schedule present, the average behavior was 1178.56 behaviors per schedule, 90%CI [1169.15, 1187.97]. During extinction schedules, the average behavior was 569.54 behaviors per schedule, 90% CI [565.42, 573.66]. A repeated measures ANOVA was performed and a statistically significant difference was found between the behavior of at least two groups ($F(19,551) = 644.95$, $p < 0.05$). The effect size, calculated as omega squared (ω^2), was 0.953, indicating a large effect. A planned contrast between behavior during all reinforced schedules and all extinction schedules was also performed and a statistically significant difference was found between these two groups ($F(1,551) = 9596.92$, $p < 0.05$). The effect size was 0.460, again indicating a large effect.

Experiment Series Two Discussion

Comparing the data from the idealized result of discrimination on multiple schedules (as seen in Figure 1) and the result obtained by the modified ETBD with background reinforcement (as seen in Figure 22), the differences are clear. While the AO was able to learn to perform specific behaviors while the discriminative stimulus was present, the rate of learning and unlearning was also a challenge for the modeling of animal behavior by AOs animated by the modified ETBD. Like the ideal case, reinforced schedules had a high number of responses. However, this rate of response was constant and does not have the gradual increase in number of responses expected in the later schedules. The extinction schedules had a lower number of responses that stayed constant through the entire experiment and did not decay, unlike the ideal case.

The stable amount of behavior during the reinforced schedules was expected. AOs animated by the standard ETBD have been found in prior experiments to learn extremely quickly. While the modified AOs had the potential for more moderate speeds of learning, it was clear from the results that the AOs animated by the modified ETBD also learned quickly when only a low number of stimulus elements were present. Based on the initial learning results shown in Figure 17, when only 10 stimulus elements are present in the local environment, the AO is able to respond appropriately to the scheduled reinforcement by approximately the 1000th generation. This speed of learning is not problematic, since live organisms also learn how to respond to schedules of reinforcement very quickly (Corrado et al., 2005).

The increase in behavior in the ideal case during the later reinforced schedules is likely due to within-trial behavioral contrast (Clement et al., 2000), which the AO was not able to replicate. It is possible that (a) behavioral contrast might need its own unique implementation in the modified ETBD in order to function, and/or (b) the baseline requirements for this to appear as an emergent phenomenon of the ETBD have not yet been met. This study does not reveal which of these possibilities is the case.

The rapid learning of the AOs during the extinction schedules was expected for the same reason as the rapid learning was expected during reinforced schedules. While the high operant level during extinction was suppressed somewhat by background reinforcement, this suppression did not allow the AOs to have levels of behavior similar to live organisms. This suggests that there is some missing mechanism here that differentiates AO behavior from live organism behavior. It is possible that this could be a within-trial behavioral contrast effect, or a learned avoidance, that is currently not incorporated into the modified ETBD. In the modified ETBD, in the absence of experimentally given reinforcement, AOs will behave randomly. This is simply not the case for live organisms (Goodman et al., 2022). It is also possible that some form of latent learning may be required to capture extinction behavior appropriately.

Another aspect of live organism behavior is that the amount of behavior in early extinction schedules is similar to the amount of behavior during reinforced schedules. Since no behaviors during an extinction schedule are rewarded, the behavior must be reinforced by other means. It is possible that some feature of the environment is acting like a discriminative stimulus (e.g. the experimental operandi, or the experimental chamber itself), and the organism’s response to that discriminative stimulus has yet to be extinguished. During the extinction schedules with five red, and five neutral wall stimuli present, it was clear that there is learning occurring while there is no learning occurring in the experiments with only 10 red stimuli. This shows that having neutral stimuli in the environment is likely very important for simulating live organism behavior.

The speed of learning and unlearning, as well as the lack of durability of learning, also adds to the difficulty of modeling such behavior. Looking at Figure 19, the extinction schedule data collected from an environment without walls generally collapses together into a narrow zone, while the behavior from environments with walls generally do not collapse together and remain relatively parallel. The initial drop in the number of behaviors in the environment without walls was likely due to background reinforcement pulling behaviors away from the target class. In contrast, the amount of behaviors in the environments with walls show how the walls can increase the number of behaviors in a schedule, based on how many walls were present. The following paragraph is a detailed description of how this may have occurred.

During the reinforced schedule, the behavioral populations linked to wall and green stimulus elements are being exposed to the fitness-based pathway in the ETBD (as shown in Figures 5 and 7) when a reinforcer is given. This leads the behavioral populations becoming biased towards the target class and also reduces the entropy rating for those populations. During the following extinction schedules, because the wall stimulus elements are carried over between schedules, they will have lower entropy than the naïve red stimulus elements, which makes them more likely to be selected as observation targets. Since this is an extinction schedule, the wall stimulus elements will quickly unlearn the bias towards the target class that developed earlier, but not before emitting some behavior in the target class. The rate of behavior in the target class during the extinction schedule will decrease and eventually stabilize at the operant level. Then the reinforced schedule occurs again causing a bias in the wall stimulus elements. The cycle repeats, so the walls continue to have the same impact on behavior during the extinction schedules, schedule after schedule.

The attempt to decrease the speed of learning by increasing the number of stimulus elements was partially successful but also created some issues. The speed of unlearning during extinction was decreased, as predicted, but the speed of learning during reinforced schedules was also significantly slowed. In addition, as the number of stimulus elements in the local environment increased, the variability in learning rate also increased. This is likely due to the increased chance of picking a set of stimulus elements that had not yet been exposed to any reinforcers. As the number of stimulus elements increased, the lack of continuity created by the random sampling method of SST became more pronounced. Some other mechanism of choosing which stimulus elements to observe may be necessary here, or a different conceptualization of stimulus elements. For example, extending entropy-based observation to cover the entire local environment instead of using random sampling might be helpful in reducing the discontinuity in learning. Another possible mechanism that might support learning is categorization. Categorization of different stimuli by organisms is a major area of study and has clear impacts on animal behavior (Seger, 2008). If AOs can sort stimuli by category, this could reduce the burden of learning tremendously, particularly in complex environments with many stimulus elements.

Taken together, the modified ETBD was able to show learning of discriminative stimuli on multiple schedules, but lacked some of the key features of animal behavior. The critical element that was missing is likely to be learning durability. AOs were able to learn quickly, creating bias in the behavioral populations towards the target class. However, this bias was quickly extinguished without reinforcement during the extinction schedule. During the course of a multiple schedule, it is expected that live organisms learn: (i) to associate the extinction stimulus with lack of reward and (ii) that the neutral stimuli in the environment are irrelevant because they do not predict either reinforcement or extinction. The current modified ETBD did not accomplish either of these goals. The durability of learning during reinforced schedules was not tested in this paradigm because there was always reinforcement present during those schedules, but it is unlikely that the learning about the discriminative stimulus (e.g. the green stimulus element) would last any longer than the learning about the extinction stimulus. If the discriminative stimulus was tested during extinction as is done during stimulus generalization experiments, this challenge would likely need to be overcome.

VI. General Discussion

This study, and many studies like this one have one major underlying question. How many phenomena can be explained by the ETBD? The selection of behavior via evolutionary principles (Skinner, 1981) and how it is theorized to be instantiated in the brain via neuronal group selection (Edelman, 1978) are both conceptual precursors to the complexity theory-based genetic algorithm known as the ETBD. Unlike the preceding theories, the ETBD is able to make predictions and falsifiable claims about the behavior of organisms, and its development is ongoing. One of the greatest strengths of the ETBD as a model is its parsimony. Through the use of very simple rules, multiple unique behavioral phenomena have been given not only a descriptive presentation (e.g. the matching law) but also a potential causal explanation (e.g. behavioral adaptation based on evolutionary principles) for their existence. Another strength of the ETBD is its transparency. Theories often have a substantial number of vague, conceptual assumptions and principles that must be true for them to have explanatory power. The ETBD has its own conceptual principles, however those principles are operationalized in a concrete and tangible function. These principles are then computed in a sequential, generative manner which lead to the predictions of the theory. It can be said that every phenomenon that can be explained by the ETBD no longer holds any mystery. This is one reason why expansion of the ETBD into new domains is so appealing. As the results of this study show, even the partial successes of the ETBD are revealing.

This study aimed to address the area of stimulus control, using a modified version of the ETBD that draws on elements of Este's SST. Each step of the way required additional modifications to be written for the modified ETBD to function appropriately, and the results were often only partially successful descriptions of live organism behavior. However, this study showed promise in terms of highlighting some important design principles when attempting to expand the ETBD, the robustness of the ETBD, and some areas of stimulus control that require extra attention that are often ignored when dealing with live organisms.

Not all modifications attempted within this study were equal. The most successful modification in the study was the inclusion of entropy. The conceptual basis of the entropy function comes from the finding that organisms pay more attention to stimuli from which they have received reinforcement in the past. The robustness to different parameter manipulations in the modified ETBD with entropy and background reinforcement was unexpected. The ability to still obtain animal-like behavior in many population sizes and background reinforcement levels in both experiment series two and three strongly suggests that entropy has a place in the ETBD. It also made use of the information stored within behavioral populations to manage how the AO attended to its local environment. The modifications to the ETBD that seem to be the most effective often follow this mold (Klapes et al., 2018; Riley, 2022). Modifications to the ETBD that operationalize principles of behavior, in a parsimonious and integrated fashion within the ETBD often seem to work better than complex systems that have entirely independent mechanisms. Another possible source for inspiration is studies on related neural processing, like research into instantiations of stimulus generalization in the brain (Fujita et al., 2020; Ramos, 2014), reward prediction error (Kishida et al., 2016; Schultz et al., 1997), and categorization (Seger, 2008) to name a few.

Two key areas of focus emerged over the course of the study. The most prominent area of focus was a need for the AOs to have a more durable response when in the presence of a discriminative stimulus or an extinction stimulus. In terms of durability, there are two challenges. The standard way of testing stimulus generalization is usually done in extinction, with no rewards being given for any behavior. To have AOs behave like live organisms, they must make this same mistake (i.e. responding without reward). Live organisms do eventually respond correctly, but at varying rates. This variability and speed of adaptation to the environment is also a possible target for modeling for the ETBD.

This project also made clear that an extinction stimulus response is likely more than just background reinforcement, and may require a separate mechanism for its function (Dunsmoor et al., 2017; Schechtman et al., 2010). When placed within an extinction environment, the AOs will quickly reduce their behavior in the target class until it reaches the operant level. This level of behavior is essentially the same level of behavior the AO will have when it is placed in a neutral environment without any reinforcement or punishment. The learning about extinction is not preserved and the AO can be expected to behave in a naïve fashion when next exposed to a reinforcer or a punisher. It’s clear that, to allow AOs to respond appropriately to an extinction stimulus when re-exposed, this learning needs to be preserved.

A secondary, but also important area of focus, is the need for greater selectivity of stimuli. There are two phenomena in this study that point to this need. First, to respond appropriately to a multiple schedule like the one in this study, the AO needs to learn, erroneously, that it should respond during the extinction schedule. Over the course of multiple repetitions, the AO should additionally learn that it will get no reinforcers during the extinction schedule and slowly decrease the amount of behavior in the target class until nothing is left. Second, the AO needs to learn to stop associating the wall or neutral stimuli with the reinforcer and only associate the discriminative stimulus with the reinforcer. The second phenomenon is more subtle and less understood than the first. In this study, its absence was shown by AOs when they re-associated wall stimuli with behaving in the target class during the reinforced schedule, regardless of how many extinction schedules they went through. The second phenomenon is often taken for granted during studies of stimulus generalization. Any stimulus control subject likely goes through this kind of process, even without extinction being present. Pavlov, when he was first describing stimulus control wrote:

> When conditioned reflexes are being established in dogs for the first time, it is found that the whole experimental environment, beginning with the introduction of the dog to the experimental room, acquires at first conditioned properties.... Later on, when the special reflex to a single definite and constant stimulus has appeared, all the other elements of the environment gradually lose their special conditioned significance. However this inhibition is at first easily dis-inhibited by any extra stimulus. (Pavlov, 1927, p. 115)

Unlike Pavlov’s work, most studies of stimulus control do not describe this loss of “special conditioned significance.” The only other reference to a similar sounding phenomenon was by Stach and Giurfa, who described bees eliminating redundant visual information unnecessary to solve the task (Stach & Giurfa, 2005). This is also likely related to the “overlap problem” of SST described earlier. SST was unable to model the learning of new information once a system was fully trained, and had no mechanism to remove extraneous stimuli. There is a need for AOs to be able to differentiate and prioritize particular stimuli when in a reinforced environment without punishment. There is also additional need for flexibility when the environment changes or if there are multiple types of stimuli in the local environment.

Limitations

There are multiple serious questions that this study does not attempt to answer, one of which is the following: What is the appropriate number of stimuli that should be present in the local environment? This study primarily utilized 10 stimulus elements to prove that stimulus control and stimulus generalization gradients were possible using the modified ETBD. In a physical sense, the limit to the number of stimulus elements could be considered nearly infinite. On the other hand, the number of stimulus elements in a functional sense is likely limited by things like the sensory limits of the organism (Wright, 1972), the computational power of an organism’s neural circuitry (Miller, 1956), or whether or not it had lunch earlier in the day (Edwards et al., 2019). Incorporating discriminative stimuli into the ETBD will likely require some form of stimulus element. The basic stimulus element designed by Estes in the SST is one of the more basic forms of this concept. Some of the environmental design challenges commented on in this paper (e.g. designating hierarchy among stimuli, dimensionalizing stimuli, etc.) can be potentially sidestepped by using cameras or microphones to act as sensory inputs, but this also still requires some form of translation into data that can be input into AOs creating again de facto stimulus elements.

The list of already investigated factors that affect stimulus control and stimulus generalization is long. Thus this study does not take into account a majority of those factors and assumes an abstract environment where many of those factors do not apply. While these factors can be manipulated or controlled in live organisms, this cannot occur in AOs without a concrete operationalization of the underlying mechanisms which connect the physical world phenomena to how they modify stimulus control behavior. However, the ignorance of the model to these other factors does not make it meaningless. Dr. Box's famous quote, "All models are wrong, but some are useful" (Box, 1979) applies here. All models are by definition, simplifications of the phenomena expressed in the physical world, and are thus wrong. However, their predictive power can be useful in many real-world applications.

Conclusions

This project combined the ETBD with elements of SST in order to determine if the combination of the two theories would be able to predict live organism behavior in stimulus control experimental paradigms. The modified ETBD, with additional modifications, was able to predict learning discriminative stimuli in multiple schedules, and generate stimulus generalization gradients similar to those generated by live organisms using Guttman and Kalish's experimental protocol (1956). The experiment also identified areas where the modified ETBD diverged from animal behavior in both paradigms and highlighted the major components that need to be addressed in order to have more a complete model of stimulus control.

VII. Future Directions

This section will focus on different possible methods to handle the two challenges highlighted above, durability and selectivity of learning, which were not addressed by this study.

Durability of learning. The durability method used in this study (the reproduction selection modifier based on entropy combined with the population mutation modifier) was unable to reproduce animal-like behavior on concurrent schedules. In addition, the learning had the tendency to collapse quickly when a small number of behaviors began to spread out over the phenotype space from the convergence point. A second durability issue was present for extinction stimuli. During extinction, the number of behaviors would drop to the operant level, rather than to zero, leading to a non-animal-like rate of behavior on the target class.

One possible mechanism to improve durability of learning without stopping the reproductive and recombinatory aspects of the ETBD, is to store this information in a location outside of the behavioral population. One potential way of doing this is to store the information in the circular reinforcement probability landscape. If a behavioral population is rewarded for emitting a particular behavior, the probability for that behavior to become a parent in future generations can be increased slightly. This can adjust the probability of a single behavior, or affect a range of phenotypes around it based on a linear probability density function. This set of weights can also be used as a proxy for past behavior. It is potentially possible to adjust the probability more or less based on the how probable that behavior was to be emitted, similar to prediction error (Schultz et al., 1997). To apply this paradigm to extinction, reward prediction error can be used to reduce the probability of becoming a parent based on loss of reinforcement. This can potentially also be used to differentiate between novel stimuli and older stimuli, giving an exposure effect, or a latent learning effect. In order for this to function properly, an internal reward/punishment paradigm needs to be implemented, along lines similar to what was done in previous studies (Riley, 2022).

Another possible durability method is to link a stability-based selection modifier to each stimulus element. One of the challenges of the selection modifier function used in this study was the loss of ability to respond to changes in the reinforcement environment once learning had occurred, until the population mutation rate spread out the behavioral population enough to allow selection, reproduction and recombination to function again. To deal with this challenge, a dynamic window capturing reinforcement history (e.g. a reinforcement context kernel) can be used to detect changes in the reinforcement environment and the ability of the organism to obtain a regular rate of reward. The following paragraphs describe an example reinforcement context kernel.

The inputs for this system are the stability of the rate of reward (S), and the rate of the reward (R) itself and the outputs are the size of the reinforcement history capturing window, and the percentage of the behavioral population that will be selected to become parents. Stability is captured by comparing two time periods of recent behavior. For example, one comparison point could be the last three generations, and the other would be the three generations before that. The proportion of reward for each set will be calculated and the difference between the two will be the measure of stability. If the difference is small, then the stability is considered high, and vice versa. See Figure 23 (top) for an example of this calculation and (bottom) an example of how data is collected for one generation.

The size of the window is a measure of the sensitivity of the kernel. The smaller the window, the more impacted it is by reinforcement, or lack of reinforcement. A larger window will have the opposite effect. When S is high, then it is assumed that the behavioral population is responding to the environment appropriately and will reduce the selection percentage in order to reduce the amount of change the behavioral population will have between generations. The window size will also decrease in order to increase sensitivity to new changes in reinforcement. When S is low, then it is assumed that the behavioral population is not responding to the environment correctly and thus the selection percentage is increased to allow for faster adaptation to the environment. The window size will also increase in order to compare the reinforcement context over a longer period of time and to reduce the sensitivity of the kernel to new reinforcement or non-reinforcement. The steps required to change the window size are shown in Figure 24.

The challenge for this particular system is the large number of parameters that need to be set by the designer. There is a need for initial parameters (e.g. starting window length, starting selection modifier) and dynamic parameters (e.g. step speed for the selection percentage, and speed of change for the window size, the thresholds for when to change the window size or leave it the same, and the selection boundary). The dynamic parameters may also be asymmetric, (e.g. requiring more activations to decrease the window size than to increase it). An example of a stimulus generalization gradient generated from using a reinforcement context kernel is shown in Figure 25.

Selectivity for a particular stimulus. Since there are no prior designs for selectivity of a particular stimulus for the ETBD or the SST, the following section will be a discussion of the principles that are likely involved in such a process. There are multiple possible ways to bring selectivity into the modified ETBD, but they all require differential treatment of stimulus elements without explicitly provided punishment. In a stimulus discrimination task, the following sequence occurs. First, many, if not all, of the stimulus elements must be able to become conditioned stimuli that lead to the conditioned response. Then there needs to be a selective inhibition of behavior on all stimulus elements that are not the discriminative stimulus, which grows stronger with time. The inhibition of behavior on all other stimulus elements will need to grow until the discriminative stimulus is left as the sole arbiter of the conditioned response. This could potentially utilize punishment (Klapes, 2020), within-trial behavioral contrast mechanisms (Zentall & Singer, 2007), or changes to the probabilities on the circular selection landscape as described earlier in order to accomplish the inhibition of behavior.

A secondary issue that is not typically discussed in discriminative stimuli experiments is the attention of the test subject. Once the discriminative stimulus is the sole arbiter of the conditioned response, the AO will need to observe the discriminative stimulus first, in order to have the conditioned response. If the environment is complex enough, random chance may not be sufficient to direct the attention of the AO to the discriminative stimulus, so a biasing method is necessary. This could be a change to the stimulus element environment to increase the probability that the discriminative stimulus is visible, a change to how the AO scans and chooses what to observe, or the act of observing itself could be added to the behavioral pool as an action. The last method is undeniably the most complex and might require additional principles to function appropriately. However, this may also be an important component of animating a robot with the ETBD.

References

Akdeniz, G. (2017). Complexity Analysis of Resting-State fMRI in Adult Patients with Attention Deficit Hyperactivity Disorder: Brain Entropy. *Computational Intelligence and Neuroscience*, *2017*, Article 3091815. https://doi.org/10.1155/2017/3091815

Akers, J. S., Retzlaff, B. J., Fisher, W. W., Greer, B. D., Kaminski, A. J., & DeSouza, A. A. (2019). An Evaluation of Conditional Manding Using a Four-Component Multiple Schedule. *Analysis of Verbal Behavior*, *35*(1), 94-102. https://doi.org/10.1007/s40616-018-0099-9

Baerends, G. P., & Drent, R. H. (1982). The herring gull and its egg: Part II. The responsiveness to egg-features. *Behaviour*, *82*, 416.

Banai, K., & Lavner, Y. (2014). The effects of training length on the perceptual learning of time-compressed speech and its generalization. *Journal of the Acoustical Society of America*, *136*(4), 1908-1917. https://doi.org/10.1121/1.4895684

Beckers, T., Miller, R. R., De Houwer, J., & Urushihara, K. (2006). Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *Journal of Experimental Psychology-General*, *135*(1), 92-102. https://doi.org/10.1037/0096-3445.135.1.92

Bhagavan, S., & Smith, B. H. (1997). Olfactory conditioning in the honey bee, Apis mellifera: Effects of odor intensity. *Physiology & Behavior*, *61*(1), 107-117. https://doi.org/10.1016/s0031-9384(96)00357-5

Bijou, S. W., & Orlando, R. (1961). RAPID DEVELOPMENT OF MULTIPLE-SCHEDULE PERFORMANCES WITH RETARDED-CHILDREN. *Journal of the Experimental Analysis of Behavior*, *4*(1), 7-16. https://doi.org/10.1901/jeab.1961.4-7

Bloomfield, T. M. (1967). A peak shift on a line-tilt continuum. *Journal of the Experimental Analysis of Behavior*, *10*(4), 361-+. https://doi.org/10.1901/jeab.1967.10-361

Blough, P. M. (1972). Wavelength generalization and discrimination in pigeon. *Perception & Psychophysics*, *12*(4), 342-&. https://doi.org/10.3758/bf03207219

Bower, G. H. (1994). A turning-point in mathematical learning-theory [Review]. *Psychological Review*, *101*(2), 290-300. https://doi.org/10.1037/0033-295x.101.2.290

Box, G. E. P. (1979). Robustness in the Strategy of Scientific Model Building. In R. L. LAUNER & G. N. WILKINSON (Eds.), *Robustness in Statistics* (pp. 201-236). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-438150-6.50018-2

Call, N. A., Clark, S. B., Mevers, J. L., Parks, N. A., Volkert, V. M., & Scheithauer, M. C. (2018). An individualized method for establishing and thinning multiple schedules of reinforcement following functional communication training. *Learning and Motivation*, *62*, 91-102. https://doi.org/10.1016/j.lmot.2017.03.006

Cammilleri, A. P., Tiger, J. H., & Hanley, G. P. (2008). Developing stimulus control of young children's requests to teachers: classwide applications of multiple schedules [Article]. *Journal of Applied Behavior Analysis*, *41*(2), 299-303. https://doi.org/10.1901/jaba.2008.41-299

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, *18*(2), 111-126. https://doi.org/10.1901/jaba.1985.18-111

Catania, A. C. (2005). The operant reserve: A computer simulation in (accelerated) real time [Article; Proceedings Paper]. *Behavioural Processes*, *69*(2), 257-278. https://doi.org/10.1016/j.beproc.2005.02.009

Chelazzi, L., Perlato, A., Santandrea, E., & Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Research*, *85*, 58-72. https://doi.org/10.1016/j.visres.2012.12.005

Cheng, K., Spetch, M. L., & Johnston, M. (1997). Spatial peak shift and generalization in pigeons. *Journal of Experimental Psychology-Animal Behavior Processes*, *23*(4), 469-481. https://doi.org/10.1037/0097-7403.23.4.469

Chi, C. (2019). *Monkey See Computer Do: Simulation of Dynamic Behavior via the Evolutionary Theory of Behavior Dynamics (Unpublished Master's Thesis)* Emory University]. Atlanta, Georgia.

Chok, J. T., & Harper, J. M. (2016). Heart Rate Assessment and Use of a Multiple Schedule Treatment for an Individual with Obsessive Compulsive-like Behavior. *Journal of Developmental and Physical Disabilities*, *28*(6), 821-834. https://doi.org/10.1007/s10882-016-9511-3

Clement, T. S., Feltus, J. R., Kaiser, D. N., & Zentall, T. R. (2000). "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*(1), 100-106. https://doi.org/10.3758/bf03210727

Corrado, G. S., Sugrue, L. P., Seung, H. S., & Newsome, W. T. (2005). Linear-Nonlinear-Poisson models of primate choice dynamic [Article]. *Journal of the Experimental Analysis of Behavior*, *84*(3), 581-617. https://doi.org/10.1901/jeab.2005.23-05

Cowie, S., Gomes-Ng, S., Hopkinson, B., Bai, J. Y. H., & Landon, J. (2020). Stimulus control depends on the subjective value of the outcome. *Journal of the Experimental Analysis of Behavior*, *114*(2), 216-232. https://doi.org/10.1002/jeab.622

Dunsmoor, J. E., Kroes, M. C. W., Braren, S. H., & Phelps, E. A. (2017). Threat Intensity Widens Fear Generalization Gradients. *Behavioral Neuroscience*, *131*(2), 168-175. https://doi.org/10.1037/bne0000186

Edelman, G. M. (1978). Group selection and phasic re-entrant signaling: a theory of higher brain function. In V. Mountcastle (Ed.), *The Mindful Brain*. MIT Press.

Edwards, T. L., Lotfizadeh, A. D., & Poling, A. (2019). Motivating operations and stimulus control. *Journal of the Experimental Analysis of Behavior*, *112*(1), 1-9. https://doi.org/10.1002/jeab.516

Estes, W. (1969a). Outline of a theory of punishment. In B. Campbell & R. Church (Eds.), *Punishment and aversive behavior*. Appleton-Century-Crofts.

Estes, W. (1969b). Reinforcement in human learning. In J. Tapp (Ed.), *Reinforcement and Behavior* (pp. 63-95). Academic Press.

Estes, W. K. (1950). Toward a statistical theory of learning [Article]. *Psychological Review*, *57*(2), 94-107. https://doi.org/10.1037/h0058559

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression [Article]. *Psychological Review*, *62*(3), 145-154. https://doi.org/10.1037/h0048509

Estes, W. K., & Johns, M. D. (1958). Probability-learning with ambiguity in the reinforcing stimulus [Article]. *American Journal of Psychology*, *71*(1), 219-228. https://doi.org/10.2307/1419209

Estes, W. K., & Lauer, D. W. (1957). Conditions of invariance and modifiability in simple reversal learning [Article]. *Journal of Comparative and Physiological Psychology*, *50*(2), 199-206. https://doi.org/10.1037/h0047344

Ferraro, D. P., & Grisham, M. G. (1972). Discrimination and generalization of complex visual shape variations in pigeons. *Perceptual and Motor Skills*, *35*(3), 915-927. https://doi.org/10.2466/pms.1972.35.3.915

Fisher, W. W., Fuhrman, A. M., Greer, B. D., Mitteer, D. R., & Piazza, C. C. (2020). Mitigating resurgence of destructive behavior using the discriminative stimuli of a multiple schedule. *Journal of the Experimental Analysis of Behavior*, *113*(1), 263-277. https://doi.org/10.1002/jeab.552

Fisher, W. W., Kuhn, D. E., & Thompson, R. H. (1998). Establishing discriminative control of responding using functional and alternative reinforcers during functional communication training. *Journal of Applied Behavior Analysis*, *31*(4), 543-560. https://doi.org/10.1901/jaba.1998.31-543

Fujita, Y., Yagishita, S., Kasai, H., & Ishii, S. (2020). Computational Characteristics of the Striatal Dopamine System Described by Reinforcement Learning With Fast Generalization. *Frontiers in Computational Neuroscience*, *14*, Article 66. https://doi.org/10.3389/fncom.2020.00066

Galizio, M. (1985). Human peak shift - analysis of the effects of 3-stimulus discrimination-training. *Learning and Motivation*, *16*(4), 478-494. https://doi.org/10.1016/0023-9690(85)90028-1

Ganz, L., & Riesen, A. H. (1962). Stimulus generalization to hue in dark-reared macaque. *Journal of Comparative and Physiological Psychology*, *55*(1), 92-&. https://doi.org/10.1037/h0044987

Ghirlanda, S. (2002). Intensity generalization: Physiology and modelling of a neglected topic. *Journal of Theoretical Biology*, *214*(3), 389-404. https://doi.org/10.1006/jtbi.2001.2452

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, *66*, 15-36. https://doi.org/10.1006/anbe.2003.2174

Gonzalez, F. A. (2002). Improved multiparameter models of drug effects on response rate under multiple variable interval schedules: evidence from rat studies. *Psychopharmacology*, *163*(3-4), 310-318. https://doi.org/10.1007/s00213-002-1136-9

Goodman, J., Gabriele, A., Ornelas, R. A. G., & Packard, M. G. (2022). Behavioral and Neural Mechanisms of Latent Extinction: A Historical Review. *Neuroscience*, *497*, 157-170. https://doi.org/10.1016/j.neuroscience.2022.06.001

Greer, B. D., Fisher, W. W., Saini, V., Owen, T. M., & Jones, J. K. (2016). Functional communication training during reinforcement schedule thinning: An analysis of 25 applications. *Journal of Applied Behavior Analysis*, *49*(1), 105-121. https://doi.org/10.1002/jaba.265

Gupta, L., Jansen, J. F. A., Hofman, P. A. M., Besseling, R. M. H., de Louw, A. J. A., Aldenkamp, A. P., & Backes, W. H. (2017). Wavelet Entropy of BOLD Time Series: An Application to Rolandic Epilepsy. *Journal of Magnetic Resonance Imaging*, *46*(6), 1728-1737. https://doi.org/10.1002/jmri.25700

Guttman, N. (1959). Generalization gradients around stimuli associated with different reinforcement schedules [Article]. *Journal of Experimental Psychology*, *58*(5), 335-340. https://doi.org/10.1037/h0045679

Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus-generalization. *Journal of Experimental Psychology*, *51*(1), 79-88. https://doi.org/10.1037/h0046219

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training [Article]. *Journal of Applied Behavior Analysis*, *34*(1), 17-38. https://doi.org/10.1901/jaba.2001.34-17

Hanson, H. M. (1959). Effects of discrimination-training on stimulus-generalization [Article]. *Journal of Experimental Psychology*, *58*(5), 321-334. https://doi.org/10.1037/h0042606

Hergenhahn, B. R., & Olson, M. H. (1993). *An Introduction to Theories of Learning* (4th ed.). Prentice Hall.

Hodges, A. C., Wilder, D. A., & Ertel, H. (2018). The use of a multiple schedule to decrease toe walking in a child with autism. *Behavioral Interventions*, *33*(4), 440-447. https://doi.org/10.1002/bin.1528

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956) - 25 years of research on stimulus-generalization [Article]. *Journal of the Experimental Analysis of Behavior*, *36*(3), 405-445. https://doi.org/10.1901/jeab.1981.36-405

Howard, M. W. (2014). Mathematical learning theory through time [Article]. *Journal of Mathematical Psychology*, *59*, 18-29. https://doi.org/10.1016/j.jmp.2013.09.003

Islas, C., Padilla, P., & Prado, M. A. (2020). Information Processing in the Brain as Optimal Entropy Transport: A Theoretical Approach. *Entropy*, *22*(11), Article 1231. https://doi.org/10.3390/e22111231

Ito, M., & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *Journal of Neuroscience*, *24*(13), 3313-3324. https://doi.org/10.1523/jneurosci.4364-03.2004

Jitsumori, M. (1978). Wavelength discrimination function derived from post-discrimination gradients in pigeon. *Japanese Psychological Research*, *20*(1), 18-28. https://doi.org/10.4992/psycholres1954.20.18

Kahnt, T., Park, S. Q., Burke, C. J., & Tobler, P. N. (2012). How Glitter Relates to Gold: Similarity-Dependent Reward Prediction Errors in the Human Striatum. *Journal of Neuroscience*, *32*(46), 16521-16529. https://doi.org/10.1523/jneurosci.2383-12.2012

Keppel, G., & Wickens, T. D. (2004). *Design and Analysis: A Researcher's Handbook* (4th ed.). Pearson Education, Inc.

Kishida, K. T., Saez, I., Lohrenz, T., Witcher, M. R., Laxton, A. W., Tatter, S. B., White, J. P., Ellis, T. L., Phillips, P. E. M., & Montague, P. R. (2016). Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(1), 200-205. https://doi.org/10.1073/pnas.1513619112

Klapes, B. (2020). *Predicting and Testing a Contemporary Quantitative Model of Punishment* [Doctoral Dissertation; Emory University]. Atlanta, GA.

Klapes, B., Riley, S., & McDowell, J. J. (2018). Toward a contemporary quantitative model of punishment. *Journal of the Experimental Analysis of Behavior*, *109*(2), 336-348. https://doi.org/10.1002/jeab.317

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., De Jong, I. C., Ruis, M. A. W., & Blokhuis, H. J. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience and Biobehavioral Reviews*, *23*(7), 925-935. https://doi.org/10.1016/s0149-7634(99)00026-3

Kulubekova, S., & McDowell, J. J. (2008). A computational model of selection by consequences: Log survivor plots [Article; Proceedings Paper]. *Behavioural Processes*, *78*(2), 291-296. https://doi.org/10.1016/j.beproc.2007.12.005

Kulubekova, S., & McDowell, J. J. (2013). Computational model of selection by consequences: Patterns of preference change on concurrent schedules [Article]. *Journal of the Experimental Analysis of Behavior*, *100*(2), 147-164. https://doi.org/10.1002/jeab.40

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them some further refinements. *Journal of Applied Behavior Analysis*, *36*(3), 407-414. https://doi.org/10.1901/jaba.2003.36-407

Lashley, K. S., & Wade, M. (1946). The pavlovian theory of generalization. *Psychological Review*, *53*(2), 72-87. https://doi.org/10.1037/h0059999

Lee, J. C., Hayes, B. K., & Lovibond, P. F. (2018). Peak Shift and Rules in Human Generalization. *Journal of Experimental Psychology-Learning Memory and Cognition*, *44*(12), 1955-1970. https://doi.org/10.1037/xlm0000558

Lee, J. C., Mills, L., Hayes, B. K., & Livesey, E. J. (2021). Modelling generalisation gradients as augmented Gaussian functions. *Quarterly Journal of Experimental Psychology*, *74*(1), 106-121. https://doi.org/10.1177/1747021820949470

Li, D., Elliffe, D., & Hautus, M. J. (2018). A multivariate assessment of the rapidly changing procedure with McDowell's Evolutionary Theory of Behavior Dynamics. *Journal of the Experimental Analysis of Behavior*, *110*(3), 336-365. https://doi.org/10.1002/jeab.478

Lotfizadeh, A. D., Edwards, T. L., Redner, R., & Poling, A. (2012). Motivating Operations Affect Stimulus Control: A Largely Overlooked Phenomenon in Discrimination Learning. *Behavior Analyst*, *35*(1), 89-100. https://doi.org/10.1007/bf03392268

Lovibond, P. F., Lee, J. C., & Hayes, B. K. (2020). Stimulus Discriminability and Induction as Independent Components of Generalization. *Journal of Experimental Psychology-Learning Memory and Cognition*, *46*(6), 1106-1120. https://doi.org/10.1037/xlm0000779

Luce, R. D. (2003). Whatever happened to information theory in psychology? *Review of General Psychology*, *7*(2), 183-188. https://doi.org/10.1037/1089-2680.7.2.183

Maes, E., De Filippo, G., Inkster, A. B., Lea, S. E. G., De Houwer, J., D'Hooge, R., Beckers, T., & Wills, A. J. (2015). Feature- versus rule-based generalization in rats, pigeons and humans. *Animal Cognition*, *18*(6), 1267-1284. https://doi.org/10.1007/s10071-015-0895-8

Masataka, N. (1983). Categorical responses to natural and synthesized alarm calls in Goeldi's monkeys (Callimico goeldii). *Primates*, *24*(1), 40-51. https://doi.org/10.1007/bf02381452

Mason, J. W. D. (2019). From Learning to Consciousness: An Example Using Expected Float Entropy Minimisation. *Entropy*, *21*(1), Article 60. https://doi.org/10.3390/e21010060

McDowell, J. J. (2004). Computational model of selection by consequences. *Journal of the Experimental Analysis of Behavior*, *81*(3), 297-317. https://doi.org/10.1901/jeab.2004.81-297

McDowell, J. J. (2013). A Quantitative Evolutionary Theory of Adaptive Behavior Dynamics [Article]. *Psychological Review*, *120*(4), 731-750. https://doi.org/10.1037/a0034244

McDowell, J. J. (2019). On the current status of the evolutionary theory of behavior dynamics. *Journal of the Experimental Analysis of Behavior*, *111*(1), 130-145. https://doi.org/10.1002/jeab.495

McDowell, J. J., Calvin, O. L., & Klapes, B. (2016). A survey of residual analysis and a new test of residual trend. *Journal of the Experimental Analysis of Behavior*, *105*(3), 445-458. https://doi.org/10.1002/jeab.208

McDowell, J. J., Caron, M. L., Kulubekova, S., & Berg, J. P. (2008). A Computational Theory of Selection by Consequences Applied to Concurrent Schedules. *Journal of the Experimental Analysis of Behavior*, *90*(3), 387-403. https://doi.org/10.1901/jeab.2008.90-387

McDowell, J. J., & Klapes, B. (2018). An evolutionary theory of behavior dynamics applied to concurrent ratio schedules [Article]. *Journal of the Experimental Analysis of Behavior*, *110*(3), 323-335. https://doi.org/10.1002/jeab.468

McDowell, J. J., & Klapes, B. (2019). An implementation of punishment in the evolutionary theory of behavior dynamics [Article]. *Journal of the Experimental Analysis of Behavior*, *112*(2), 128-143. https://doi.org/10.1002/jeab.543

McDowell, J. J., & Klapes, B. (2020). All behavior is choice: Revisiting an evolutionary theory's account of behavior on single schedules. *Journal of the Experimental Analysis of Behavior*, *114*(3), 430-446. https://doi.org/10.1002/jeab.630

McDowell, J. J., & Popa, A. (2009). Beyond continuous mathematics and traditional scientific analysis: Understanding and mining Wolfram's A New Kind of Science. *Behavioural Processes*, *81*(2), 343-352. https://doi.org/10.1016/j.beproc.2009.01.012

McDowell, J. J., Popa, A., & Calvin, N. T. (2012). Selection Dynamics in Joint Matching to Rate and Magnitude of Reinforcement [Article]. *Journal of the Experimental Analysis of Behavior*, *98*(2), 199-212. https://doi.org/10.1901/jeab.2012.98-199

McDowell, J. J., Soto, P. L., Dallery, J., & Kulubekova, S. (2006). A computational theory of adaptive behavior based on an evolutionary reinforcement mechanism. *Gecco 2006: Genetic and Evolutionary Computation Conference, Vol 1 and 2*, 175-+. https://doi.org/10.1145/1143997.1144028

Miller, G. A. (1956). The magical number 7, plus or minus 2 - some limits on our capacity for processing information [Article]. *Psychological Review*, *63*(2), 81-97. https://doi.org/10.1037/h0043158

Ming, X., Brimacombe, M., & Wagner, G. C. (2007). Prevalence of motor impairment in autism spectrum disorders. *Brain & Development*, *29*(9), 565-570. https://doi.org/10.1016/j.braindev.2007.03.002

Nava, M. J., Vargo, K. K., & Babino, M. M. (2016). An evaluation of a three-component multiple schedule to indicate attention availability. *Journal of Applied Behavior Analysis*, *49*(3), 674-679. https://doi.org/10.1002/jaba.297

Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, *57*(3), 301-316. https://doi.org/10.1901/jeab.1992.57-301

Niedfeld, A., Rapp, J. T., Coon, J. C., & Cook, J. L. (2020). Using a Multiple-Schedule Procedure to Signal the Availability of Attention: Three Demonstrations. *Behavior Modification*, *44*(4), 496-517. https://doi.org/10.1177/0145445519834640

Onat, S., & Büchel, C. (2015). The neuronal basis of fear generalization in humans. *Nature Neuroscience*, *18*(12), 1811-1818. https://doi.org/10.1038/nn.4166

Pavlov, I. P. (1927). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. In: Oxford University Press.

Pearce, J. M., Esber, G. P., George, D. N., & Haselgrove, M. (2008). The nature of discrimination learning in pigeons. *Learning & Behavior*, *36*(3), 188-199. https://doi.org/10.3758/lb.36.3.188

Pierce, W. D., & Cheney, C. D. (2013). *Behavior analysis and learning* (6th ed.). Psychology Press.

Popa, A., & McDowell, J. J. (2010). The effect of Hamming distances in a computational model of selection by consequences [Article; Proceedings Paper]. *Behavioural Processes*, *84*(1), 428-434. https://doi.org/10.1016/j.beproc.2010.02.002

Ramos, J. M. J. (2014). Perirhinal cortex lesions attenuate stimulus generalization in a tactual discrimination task in rats. *Acta Neurobiologiae Experimentalis*, *74*(1), 15-25. https://doi.org/10.55782/ane-2014-1968

Razran, G. (1949). Stimulus generalization of conditioned responses. *Psychological Bulletin*, *46*(5), 337-365. https://doi.org/10.1037/h0060507

Reynolds, G. S. (1961). BEHAVIORAL-CONTRAST. *Journal of the Experimental Analysis of Behavior*, *4*(1), 57-71. https://doi.org/10.1901/jeab.1961.4-57

Riley, S. (2022). *Approximating McDowell's Evolutionary Theory of Behavior Dynamics with Stochastic Neural Networks* Emory Theses and Dissertations]. Emory University. https://etd.library.emory.edu/concern/etds/ks65hd54v?locale=en

Rilling, M. (1977). Stimulus control and inhibitory processes. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of Operant Behavior* (pp. 432–480). Prentice Hall.

Schechtman, E., Laufer, O., & Paz, R. (2010). Negative Valence Widens Generalization of Learning. *Journal of Neuroscience*, *30*(31), 10460-10464. https://doi.org/10.1523/jneurosci.2377-10.2010

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599. https://doi.org/10.1126/science.275.5306.1593

Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience and Biobehavioral Reviews*, *32*(2), 265-278. https://doi.org/10.1016/j.neubiorev.2007.07.010

Shamlian, K. D., Fisher, W. W., Steege, M. W., Cavanaugh, B. M., Samour, K., & Querim, A. C. (2016). Evaluation of multiple schedules with naturally occurring and therapist-arranged discriminative stimuli following functional communication training. *Journal of Applied Behavior Analysis*, *49*(2), 228-250. https://doi.org/10.1002/jaba.293

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science [Article]. *Science*, *237*(4820), 1317-1323. https://doi.org/10.1126/science.3629243

Skinner, B. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.

Skinner, B. F. (1950). ARE THEORIES OF LEARNING NECESSARY. *Psychological Review*, *57*(4), 193-216. https://doi.org/10.1037/h0054367

Skinner, B. F. (1981). Selection by Consequences [Article]. *Science*, *213*(4507), 501-504. https://doi.org/10.1126/science.7244649

Somwanshi, D., Kumar, A., Sharma, P., Joshi, D., & Ieee. (2016). An efficient Brain Tumor Detection from MRI Images using Entropy Measures. *2016 International Conference on Recent Advances and Innovations in Engineering (Icraie)*. https://doi.org/10.1109/ICRAIE.2016.7939554

Song, Y. R., Zhou, D. L., & Li, S. T. (2021). Maximum Entropy Principle Underlies Wiring Length Distribution in Brain Networks. *Cerebral Cortex*, *31*(10), 4628-4641. https://doi.org/10.1093/cercor/bhab110

Soto, F. A., & Wasserman, E. A. (2010). Integrality/Separability of Stimulus Dimensions and Multidimensional Generalization in Pigeons. *Journal of Experimental Psychology-Animal Behavior Processes*, *36*(2), 194-205. https://doi.org/10.1037/a0016560

Stach, S., & Giurfa, M. (2005). The influence of training length on generalization of visual feature assemblies in honeybees. *Behavioural Brain Research*, *161*(1), 8-17. https://doi.org/10.1016/j.bbr.2005.02.008

Staddon, J. E. R., & Reid, A. K. (1990). On the dynamics of generalization. *Psychological Review*, *97*(4), 576-578. https://doi.org/10.1037/0033-295x.97.4.576

Thomas, D. R., & King, R. A. (1959). Stimulus-generalization as a function of level of motivation. *Journal of Experimental Psychology*, *57*(5), 323-328. https://doi.org/10.1037/h0042183

Tinbergen, N. (1951). *The Study Of Instinct*. Oxford University Press.

Vargo, K., Heal, N., Epperley, K., & Kooistra, E. (2014). The Effects of a Multiple Schedule Plus Rules on Hand Raising During Circle Time in Preschool Classrooms [Article]. *Journal of Behavioral Education*, *23*(3), 326-343. https://doi.org/10.1007/s10864-014-9199-3

Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010). Fear generalization in humans: Impact of prior non-fearful experiences. *Behaviour Research and Therapy*, *48*(11), 1078-1084. https://doi.org/10.1016/j.brat.2010.07.002

Viol, A., Palhano-Fontes, F., Onias, H., de Araujo, D. B., Hovel, P., & Viswanathan, G. M. (2019). Characterizing Complex Networks Using Entropy-Degree Diagrams: Unveiling Changes in Functional Brain Connectivity Induced by Ayahuasca. *Entropy*, *21*(2), Article 128. https://doi.org/10.3390/e21020128

Weaver, W. (2017). Recent contributions to the mathematical theory of communication [Book Review]. *ETC: A Review of General Semantics*, *74*(1/2), 136-157. https://login.proxy.library.emory.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=tfh&AN=132792721&site=ehost-live&scope=site

Wong, A. H. K., Glück, V. M., Boschet, J. M., & Engelke, P. (2020). Generalization of extinction with a generalization stimulus is determined by learnt threat beliefs. *Behaviour Research and Therapy*, *135*, Article 103755. https://doi.org/10.1016/j.brat.2020.103755

Wong, A. H. K., & Lovibond, P. F. (2020). Generalization of extinction of a generalization stimulus in fear learning. *Behaviour Research and Therapy*, *125*, Article 103535. https://doi.org/10.1016/j.brat.2019.103535

Wright, A. A. (1972). Psychometric and psychophysical hue discrimination functions for pigeon. *Vision Research*, *12*(9), 1447-&. https://doi.org/10.1016/0042-6989(72)90171-x

Wright, A. A. (1974). Psychometric and psychophysical theory within a framework of response bias. *Psychological Review*, *81*(4), 322-347. https://doi.org/10.1037/h0036538

Zaman, J., Ceulemans, E., Hermans, D., & Beckers, T. (2019). Direct and indirect effects of perception on generalization gradients. *Behaviour Research and Therapy*, *114*, 44-50. https://doi.org/10.1016/j.brat.2019.01.006

Zaman, J., Yu, K., & Lee, J. C. (2022). Individual Differences in Stimulus Identification, Rule Induction, and Generalization of Learning. *Journal of Experimental Psychology-Learning Memory and Cognition*. https://doi.org/10.1037/xlm0001153

Zentall, T. R., & Singer, R. A. (2007). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than a less aversive event. *Journal of the Experimental Analysis of Behavior*, *88*(1), 131-149. https://doi.org/10.1901/jeab.2007.27-06

Zidar, J., Balogh, A. C. V., Leimar, O., & Lovlie, H. (2019). Generalization of learned preferences covaries with behavioral flexibility in red junglefowl chicks. *Behavioral Ecology*, *30*(5), 1375-1381. https://doi.org/10.1093/beheco/arz088

Zielinski, K., & Jakubowska, E. (1977). Auditory intensity generalization after cer differentiation training. *Acta Neurobiologiae Experimentalis*, *37*(3), 191-205.

Table 1

Details of Experimental Procedures, Experiment One

Experiment	Source	Schedules	Stimulus Elements	AO Modification	AO Quantity
Exp One, Phase One Direct ETBD Replication	(McDowell et al., 2008)	(Experiment One Schedule) Schedule 1-11: Concurrent schedules (20k gen) TC1/TC2 RIs: 20/120, 30/110, 40/100, 50/90, 60/80, 70/70, 80/60, 90/50, 100/40, 110/30, 120/20 TC1/TC2 RMs: 40/40	One Wall	None	30
Exp One, Phase Two Replication With stimulus elements			One, Two, … 15	None	30
Exp One, Phase Two Modification One - Observation				Entropy Observation 5%	30
				Entropy Observation 2%	30
Exp One, Phase Two Modification Two - Bx Population				Entropy Observation 2% Bx Pop (50, 100 … 500)	300 (30 per condition)

Table 2

Details of Experimental Procedures, Experiment Series Two

Experiment	Source	Schedules	Stimulus Elements	AO/Procedure Modifications	AO Quantity
Exp Two ETBD w/ Distinct stimulus element	(Based on example in p.262 Pierce & Cheney 2013)	(Experiment Two Schedule) Schedule 1: Wall stimulus element only Schedules 2-20 (even): Wall & Red stimulus element Target RI = 70 Schedules 3-21 (odd): Wall & Green stimulus element Target RI = 0 (extinction)	5 Wall, 5 Green, 5 Red	Bx Pop 200 Entropy Observation 2%	30
Exp Two Parameter Baselines				Extinction Only Bx Pop 200 Entropy Observation 2%	30
				Bx Pop 200 Entropy Observation 2% BKGD 200 Targets (RI 20, RM 40)	30
Exp Two Modification Three - Background R+				Bx Pop 200 Entropy Observation 2% BKGD 200 Targets, (RI = various, RM 40)	300 (30 per condition)
Exp Two Modification Four - stimulus element Quantity		Same schedules as Experiment Two	X Green, X Red, (X = Various)	Bx Pop 200 Entropy Observation 2% BKGD 200 Targets (RI 20, RM 40)	330 (30 per condition)
			X Wall, X Green, X Red, (X = Various)	Bx Pop 200 Entropy Observation 2% BKGD 200 Targets (RI 20, RM 40)	150 (30 per condition)
Exp Two, Concurrent Schedules Review	(McDowell et al., 2008)	Same schedules as Experiment One	10 Wall	Bx Pop 200 Entropy Observation 2% BKGD 200 Targets (RI = Various, RM = Various)	1980 (30 per condition)

Table 3 (Part 1 of 3)

Details of Experimental Procedures, Experiment Three

Experiment	Source	Schedules/Targets	Stimulus Elements	AO Modification	AO Quantity
Exp 3, Phase 1 Modification Five - Selection Modifier	Based on Guttman & Kalish, 1956)	Experiment Three Schedules Schedule 1: Wall only (20k gen) Schedule 2: R+ Training (20k gen) Schedule 3-12: Testing (200 gen) Target Class (RI 10, RM 05) BKGD 200 Targets (RI 10, RM 40)	5 Wall stimulus element, 5 Training Wall stimulus element, 5 Red stimulus element, 10 Red adjacent stimulus element	Bx Pop 200 Entropy Observation 2% Selection Modifier (Concave) Minimum Mutation Every X Gen (x = Various)	480 (30 per condition)
Exp 3, Phase 1 Training Time	Based on Guttman & Kalish, 1956)	Schedule 1: Wall only (20k gen) Schedule 2: R+ Training (Various gen) Schedule 3-12: Testing (200 gen) Target Class (RI = Various, RM = Various) BKGD 200 Targets (RI 10, RM 40)	5 wall 5 training wall 5 target stimulus element 5 stimulus element below 5 stimulus element above	Bx Pop 200 Entropy Observation 2% Selection Modifier (Concave) Min Mutation Every X Gen (X = 100)	150 (30 per condition)

Table 3 (Part 2 of 3)

Details of Experimental Procedures, Experiment Three

Experiment	Source	Schedules/Targets	Stimulus Elements	AO Modification	AO Quantity
Exp Three, Phase 2 Modification Six - Variation Reduction	Based on Guttman & Kalish, 1956)	Same Schedules as Experiment Three Target Class (RI 10, RM 05) BKGD 200 Targets (RI 10, RM 40)	5 wall 5 training wall 5 target stimulus element 5 stimulus element below 5 stimulus element above	Bx Pop 200 Entropy Observation 2% Selection Modifier (Concave) Min Mutation Every X Gen (X = Various) 80% Training cut off/ No Training cut off	660 (30 per condition)
		Same Schedules as Experiment Three Target Class (RI 10, RM 05) BKGD 200 Targets (No R+)			660 (30 per condition)
		Same Schedules as Experiment Three Target Class (RI 10, RM 05) BKGD 200 Targets (RI 10, RM 40)	10 target stimulus element 10 stimulus element below 10 stimulus element above		660 (30 per condition)
		Same Schedules as Experiment Three Target Class (RI 10, RM 05) BKGD 200 Targets (No R+)			660 (30 per condition)

Table 3 (Part 3 of 3)

Details of Experimental Procedures, Experiment Three

Experiment	Source	Schedules/Targets	Stimulus Elements	AO Modification	AO Quantity
Exp 3, Concurrent Schedules Review (Selection and Mutation Modifer)	(McDowell et al., 2008)	Same Schedules as Experiment One BKGD 200 Targets (RI = Various, RM = Various)	10 ıs element	Bx Pop 200 Entropy Observation 2% Selection Modifier (Concave) Minimum Mutation Every X Gen (x = None)	300 (10 per condition)
		Same Schedules as Experiment One BKGD 200 Targets (RI 10, RM 40)		Bx Pop 200 Entropy Observation 2% Selection Modifier (Concave) Minimum Mutation Every X Gen (x = Various)	180 (30 per condition)
Exp 3, Concurrent Schedules Review (Behavior Population Siz)		Same Schedules as Experiment One BKGD 200 Targets (RI 10, RM 40)		Bx Pop (Various) Entropy Observation 2% Selection Modifier (Concave) Min Mutation Every X Gen (x = Various)	210 (30 per condition)
Exp 3, Concurrent Schedules Review (Selection Modifer Curve)		Same Schedules as Experiment One BKGD 200 Targets (RI 10, RM 40)	10 stimulus element	Bx Pop 200 Entropy Observation 2% Selection Modifier (Concave, Linear, Convex) Min Mutation Every X Gen (x = Various)	1890 (30 per condition)

Table 4

Experiment 1, Phase One data

No.	*a*	*b*	PVAF	RCTT r^2
1	0.89	0.94	0.99	0.01
2	0.85	1.04	0.97	0.02
3	0.87	1.04	0.98	0.02
4	0.90	0.99	0.98	0.00
5	0.89	1.02	0.96	0.02
6	0.95	0.96	0.99	0.00
7	0.80	0.95	0.99	0.01
8	0.95	0.95	0.99	0.02
9	0.88	0.96	0.98	0.01
10	0.93	0.97	0.99	0.00
11	0.95	0.98	0.99	0.00
12	0.84	1.02	0.98	0.01
13	0.83	0.96	0.98	0.01
14	0.82	0.95	0.99	0.01
15	0.92	1.04	0.98	0.01
16	0.89	1.00	0.98	0.00
17	0.87	0.98	0.98	0.01
18	0.81	1.03	0.99	0.00
19	0.89	0.99	0.97	0.00
20	0.87	1.00	0.98	0.01
21	0.99	1.00	0.99	0.01
22	0.92	0.95	0.99	0.02
23	0.84	0.96	0.98	0.01
24	0.85	1.03	0.96	0.03
25	0.86	1.05	0.98	0.00
26	0.90	0.91	0.99	0.01
27	0.88	1.05	0.98	0.01
28	0.84	1.01	0.99	0.00
29	0.91	1.02	0.99	0.01
30	0.85	1.05	0.98	0.02
Mean	0.88	0.99	0.98	0.01

Note for Table 4. The RCTT r^2 in the table is the proportion of the variance in the residuals explained by the best fitting cubic polynomial.

Table 5

Random Five Observation Style (Experiment 1 Phase 2, Dev. 1)

stimulus element Quantity	Measure	Mean	SEM	CI +90%	CI -90%	Standard Modification	Max	Min
1	a	0.881	0.008	0.895	0.866	0.046	0.985	0.799
	b	0.994	0.007	1.006	0.982	0.039	1.053	0.909
	PVAF	0.982	0.002	0.985	0.98	0.008	0.994	0.956
2	a	0.578	0.008	0.593	0.564	0.046	0.656	0.494
	b	0.995	0.007	1.007	0.983	0.039	1.047	0.915
	PVAF	0.964	0.003	0.969	0.959	0.016	0.988	0.913
3	a	0.539	0.007	0.551	0.527	0.038	0.608	0.464
	b	1.006	0.006	1.017	0.996	0.033	1.051	0.938
	PVAF	0.956	0.004	0.963	0.95	0.021	0.988	0.904
4	a	0.499	0.007	0.511	0.487	0.039	0.555	0.413
	b	0.988	0.008	1.001	0.975	0.042	1.091	0.922
	PVAF	0.958	0.003	0.963	0.954	0.015	0.985	0.933
5	a	0.51	0.008	0.524	0.497	0.043	0.595	0.419
	b	1.011	0.005	1.02	1.002	0.028	1.061	0.945
	PVAF	0.961	0.003	0.967	0.956	0.017	0.991	0.927
6	a	0.502	0.007	0.514	0.491	0.037	0.57	0.421
	b	0.999	0.006	1.009	0.99	0.031	1.05	0.923
	PVAF	0.963	0.003	0.968	0.959	0.014	0.987	0.932
7	a	0.51	0.006	0.52	0.5	0.032	0.592	0.44
	b	0.988	0.006	0.998	0.979	0.032	1.063	0.929
	PVAF	0.967	0.003	0.972	0.963	0.014	0.994	0.928
8	a	0.514	0.006	0.525	0.503	0.035	0.602	0.454
	b	1	0.006	1.009	0.99	0.03	1.08	0.946
	PVAF	0.964	0.003	0.969	0.958	0.018	0.995	0.901
9	a	0.511	0.006	0.521	0.501	0.031	0.579	0.447
	b	0.994	0.006	1.004	0.985	0.031	1.069	0.942
	PVAF	0.961	0.003	0.965	0.957	0.014	0.988	0.93
10	a	0.519	0.007	0.53	0.508	0.036	0.578	0.452
	b	0.993	0.007	1.004	0.982	0.036	1.083	0.923
	PVAF	0.962	0.002	0.966	0.958	0.012	0.98	0.932
11	a	0.523	0.006	0.532	0.513	0.031	0.566	0.433
	b	1	0.005	1.008	0.991	0.029	1.061	0.955
	PVAF	0.962	0.003	0.968	0.957	0.019	0.989	0.904
12	a	0.508	0.006	0.518	0.497	0.033	0.563	0.437
	b	1.001	0.006	1.011	0.99	0.033	1.059	0.928

	PVAF	0.968	0.002	0.972	0.963	0.013	0.99	0.929
13	a	0.499	0.007	0.51	0.487	0.037	0.59	0.407
	b	1.001	0.006	1.011	0.991	0.033	1.063	0.938
	PVAF	0.967	0.002	0.97	0.963	0.012	0.994	0.938
14	a	0.516	0.007	0.529	0.504	0.041	0.596	0.436
	b	0.988	0.005	0.997	0.979	0.029	1.033	0.911
	PVAF	0.963	0.002	0.967	0.958	0.013	0.98	0.93
15	a	0.505	0.005	0.514	0.497	0.027	0.567	0.441
	b	1.002	0.007	1.014	0.99	0.039	1.1	0.91
	PVAF	0.962	0.004	0.969	0.956	0.02	0.989	0.911

Table 6

Entropy 5% Observation Style (Experiment 1, Dev. 1)

stimulus element Quantity	Measure	Mean	SEM	CI +90%	CI -90%	Standard Modification	Max	Min
1	a	0.881	0.009	0.896	0.866	0.049	1.019	0.81
	b	0.997	0.008	1.01	0.984	0.042	1.081	0.911
	PVAF	0.983	0.002	0.986	0.98	0.011	0.995	0.949
2	a	0.812	0.006	0.822	0.802	0.034	0.896	0.766
	b	0.998	0.006	1.009	0.987	0.035	1.056	0.938
	PVAF	0.983	0.001	0.985	0.981	0.007	0.995	0.969
3	a	0.775	0.007	0.787	0.763	0.04	0.856	0.697
	b	1.005	0.006	1.016	0.995	0.035	1.073	0.924
	PVAF	0.982	0.002	0.984	0.979	0.009	0.997	0.959
4	a	0.756	0.006	0.766	0.746	0.033	0.823	0.675
	b	1.001	0.008	1.014	0.987	0.044	1.099	0.916
	PVAF	0.981	0.002	0.984	0.978	0.009	0.993	0.962
5	a	0.764	0.006	0.774	0.754	0.032	0.809	0.694
	b	1.007	0.007	1.019	0.995	0.039	1.086	0.923
	PVAF	0.979	0.002	0.982	0.977	0.009	0.992	0.958
6	a	0.731	0.008	0.744	0.718	0.041	0.832	0.614
	b	0.998	0.006	1.008	0.987	0.033	1.084	0.945
	PVAF	0.981	0.002	0.984	0.979	0.008	0.993	0.965
7	a	0.721	0.006	0.732	0.711	0.033	0.782	0.65
	b	0.999	0.005	1.006	0.991	0.025	1.056	0.956
	PVAF	0.982	0.002	0.985	0.979	0.01	0.994	0.949
8	a	0.703	0.006	0.713	0.693	0.033	0.762	0.633
	b	1.007	0.005	1.016	0.997	0.03	1.076	0.957
	PVAF	0.979	0.002	0.983	0.976	0.01	0.996	0.95
9	a	0.695	0.006	0.705	0.686	0.032	0.768	0.642
	b	0.998	0.007	1.009	0.987	0.036	1.057	0.901
	PVAF	0.981	0.002	0.984	0.979	0.009	0.994	0.962
10	a	0.688	0.005	0.696	0.68	0.025	0.737	0.635
	b	0.997	0.005	1.005	0.988	0.028	1.07	0.944
	PVAF	0.983	0.001	0.985	0.981	0.007	0.993	0.969
11	a	0.691	0.005	0.7	0.681	0.03	0.754	0.635
	b	1	0.007	1.011	0.989	0.036	1.065	0.942
	PVAF	0.987	0.001	0.989	0.984	0.007	0.997	0.969
12	a	0.695	0.005	0.704	0.687	0.027	0.762	0.636
	b	0.993	0.005	1.001	0.985	0.026	1.04	0.933

	PVAF	0.984	0.001	0.987	0.982	0.007	0.991	0.958
13	a	0.671	0.004	0.678	0.663	0.023	0.736	0.628
	b	1.002	0.006	1.012	0.992	0.032	1.071	0.938
	PVAF	0.984	0.001	0.986	0.981	0.008	0.996	0.967
14	a	0.665	0.006	0.674	0.655	0.031	0.717	0.611
	b	0.996	0.006	1.005	0.986	0.03	1.068	0.941
	PVAF	0.982	0.002	0.985	0.979	0.01	0.998	0.949
15	a	0.666	0.005	0.674	0.658	0.026	0.718	0.598
	b	0.993	0.006	1.003	0.984	0.03	1.047	0.932
	PVAF	0.983	0.001	0.986	0.981	0.008	0.994	0.968

Table 7

Entropy 2% Observation Style (Experiment 1, Dev. 1)

stimulus element Quantity	Measure	Mean	SEM	CI +90%	CI -90%	Standard Modification	Max	Min
1	a	0.868	0.008	0.881	0.855	0.041	0.942	0.762
	b	1.003	0.008	1.017	0.989	0.045	1.102	0.919
	PVAF	0.982	0.001	0.985	0.98	0.008	0.993	0.964
2	a	0.858	0.007	0.871	0.846	0.04	0.923	0.771
	b	1.013	0.006	1.023	1.003	0.032	1.106	0.962
	PVAF	0.981	0.002	0.984	0.978	0.01	0.993	0.946
3	a	0.866	0.008	0.88	0.852	0.045	0.927	0.746
	b	1.012	0.009	1.028	0.996	0.052	1.137	0.909
	PVAF	0.984	0.001	0.987	0.982	0.007	0.994	0.965
4	a	0.854	0.007	0.866	0.842	0.038	0.91	0.761
	b	0.999	0.005	1.008	0.989	0.03	1.069	0.929
	PVAF	0.984	0.001	0.986	0.982	0.007	0.995	0.964
5	a	0.845	0.007	0.856	0.834	0.036	0.931	0.775
	b	1.004	0.009	1.019	0.989	0.049	1.092	0.89
	PVAF	0.984	0.001	0.986	0.982	0.008	0.995	0.964
6	a	0.817	0.007	0.828	0.805	0.037	0.92	0.728
	b	0.991	0.008	1.003	0.978	0.041	1.064	0.906
	PVAF	0.987	0.001	0.989	0.985	0.006	0.997	0.969
7	a	0.781	0.007	0.792	0.77	0.036	0.864	0.713
	b	1.008	0.006	1.018	0.998	0.032	1.075	0.94
	PVAF	0.985	0.001	0.987	0.983	0.007	0.996	0.972
8	a	0.784	0.005	0.792	0.775	0.028	0.836	0.727
	b	1	0.005	1.008	0.992	0.027	1.044	0.943
	PVAF	0.986	0.001	0.987	0.984	0.006	0.994	0.971
9	a	0.766	0.006	0.776	0.756	0.032	0.826	0.699
	b	1.007	0.006	1.018	0.997	0.033	1.102	0.945
	PVAF	0.989	0.001	0.991	0.987	0.005	0.997	0.976
10	a	0.754	0.006	0.764	0.744	0.033	0.835	0.688
	b	0.999	0.006	1.009	0.99	0.03	1.055	0.911
	PVAF	0.987	0.001	0.989	0.985	0.006	0.996	0.972
11	a	0.742	0.005	0.75	0.733	0.028	0.791	0.696
	b	0.997	0.006	1.006	0.987	0.031	1.068	0.938
	PVAF	0.986	0.001	0.988	0.984	0.007	0.994	0.97
12	a	0.728	0.006	0.737	0.718	0.031	0.785	0.655
	b	0.998	0.006	1.008	0.989	0.031	1.081	0.951

	PVAF	0.985	0.001	0.987	0.982	0.008	0.993	0.961
13	a	0.724	0.006	0.735	0.713	0.035	0.803	0.675
	b	1.007	0.006	1.017	0.997	0.033	1.069	0.945
	PVAF	0.986	0.001	0.988	0.983	0.008	0.996	0.965
14	a	0.725	0.005	0.733	0.717	0.026	0.782	0.667
	b	1.003	0.006	1.014	0.993	0.034	1.08	0.946
	PVAF	0.989	0.001	0.991	0.987	0.006	0.995	0.972
15	a	0.713	0.004	0.721	0.705	0.025	0.766	0.665
	b	0.994	0.006	1.004	0.983	0.033	1.048	0.929
	PVAF	0.987	0.001	0.989	0.984	0.008	0.998	0.958

Table 8

Impact of population size on sensitivity data (Experiment 1, Dev. 2)

Population Size	Measure	Mean	SEM	CI +90%	CI -90%	Standard Modification	Max	Min
50	a	0.703	0.006	0.712	0.694	0.03	0.753	0.63
	b	1.01	0.005	1.019	1.002	0.027	1.068	0.962
	PVAF	0.985	0.001	0.987	0.982	0.007	0.996	0.969
100	a	0.754	0.006	0.764	0.744	0.033	0.835	0.688
	b	0.999	0.006	1.009	0.99	0.03	1.055	0.911
	PVAF	0.987	0.001	0.989	0.985	0.006	0.996	0.972
150	a	0.774	0.005	0.783	0.766	0.027	0.835	0.714
	b	1.003	0.005	1.012	0.994	0.03	1.056	0.934
	PVAF	0.987	0.001	0.989	0.985	0.006	0.994	0.969
200	a	0.796	0.006	0.807	0.785	0.035	0.864	0.715
	b	1.003	0.006	1.013	0.993	0.032	1.071	0.946
	PVAF	0.986	0.002	0.989	0.983	0.01	0.997	0.945
250	a	0.8	0.005	0.81	0.791	0.029	0.872	0.745
	b	1.001	0.005	1.01	0.993	0.027	1.052	0.934
	PVAF	0.987	0.001	0.989	0.985	0.006	0.994	0.97
300	a	0.8	0.005	0.809	0.792	0.028	0.849	0.734
	b	1.004	0.006	1.014	0.994	0.033	1.064	0.937
	PVAF	0.989	0.001	0.991	0.988	0.005	0.997	0.978
350	a	0.819	0.006	0.829	0.809	0.031	0.867	0.771
	b	1.006	0.006	1.016	0.996	0.032	1.087	0.947
	PVAF	0.988	0.001	0.99	0.986	0.007	0.997	0.973
400	a	0.817	0.006	0.827	0.807	0.033	0.906	0.736
	b	1.003	0.007	1.015	0.991	0.038	1.118	0.95
	PVAF	0.986	0.001	0.988	0.985	0.006	0.995	0.974
450	a	0.82	0.006	0.831	0.81	0.035	0.88	0.735
	b	1	0.007	1.012	0.989	0.038	1.077	0.926
	PVAF	0.988	0.001	0.99	0.985	0.007	0.997	0.966
500	a	0.818	0.007	0.829	0.807	0.036	0.89	0.761
	b	1.001	0.007	1.012	0.99	0.036	1.099	0.951
	PVAF	0.987	0.001	0.989	0.984	0.007	0.998	0.97

Table 9

Experiment 1, Conditions: stimulus element 10, EN02, and Behavior Population 200

No.	***a***	***b***	**PVAF**	**RCTT r^2**
1	0.79	1.00	0.99	0.00
2	0.84	0.96	0.99	0.01
3	0.81	1.04	0.99	0.01
4	0.78	0.97	0.99	0.01
5	0.78	1.00	0.99	0.01
6	0.86	1.00	0.99	0.00
7	0.82	1.02	0.99	0.00
8	0.79	0.95	0.94	0.01
9	0.80	0.99	1.00	0.00
10	0.83	1.00	0.99	0.00
11	0.76	1.00	0.98	0.00
12	0.71	1.02	0.97	0.01
13	0.76	0.97	0.98	0.01
14	0.83	1.05	0.98	0.00
15	0.78	1.06	0.98	0.01
16	0.80	0.99	0.99	0.00
17	0.79	0.95	0.99	0.00
18	0.75	0.99	0.98	0.01
19	0.80	0.98	0.99	0.01
20	0.75	0.99	0.99	0.00
21	0.79	1.05	0.98	0.01
22	0.80	1.00	0.99	0.00
23	0.78	1.05	0.98	0.00
24	0.81	1.03	0.99	0.00
25	0.84	1.02	1.00	0.00
26	0.86	1.00	0.99	0.01
27	0.76	1.00	0.99	0.00
28	0.82	0.97	0.99	0.00
29	0.76	0.99	0.98	0.00
30	0.83	1.07	0.99	0.01
Mean	0.80	1.00	0.99	0.01

Table 10

Matching law parameters (Experiment Series Two, Modification Three)

Reinforcement Magnitude	a	b	PVAF
20	0.76 [0.74, 0.79]	1 [0.98, 1.02]	0.95 [0.95, 0.96]
40	0.75 [0.73, 0.77]	1 [0.98, 1.01]	0.96 [0.95, 0.96]
60	0.76 [0.73, 0.78]	1 [0.98, 1.02]	0.96 [0.95, 0.96]

Table 10 Note. Parameters presented here are averages of data across 11 background reinforcement levels, ranging from RI 20 to RI 120. Data in brackets represent 90% confidence intervals.

Table 11

ANOVA Table - 80% Training Cut off (Set 1, Experiment 3, Modification Six)

PMR	*df*	*F*	ω^2	*p*
Wall stimulus element and Background R+				
10	10	14.235	0.284	0.000
15	10	20.286	0.367	0.000
20	10	14.873	0.294	0.000
30	10	17.803	0.335	0.000
40	10	19.446	0.356	0.000
50	10	22.381	0.391	0.000
75	10	13.603	0.274	0.000
100	10	21.771	0.384	0.000
500	10	22.137	0.388	0.000
1000	10	25.748	0.426	0.000
2000	10	25.988	0.428	0.000
5000	10	36.310	0.514	0.000
10000	10	47.519	0.583	0.000
Null	10	40.928	0.545	0.000
No Wall stimulus element and Background R+				
10	10	28.672	0.454	0.000
15	10	51.135	0.601	0.000
20	10	62.895	0.650	0.000
30	10	50.391	0.597	0.000
40	10	74.241	0.687	0.000
50	10	88.601	0.724	0.000
75	10	66.233	0.662	0.000
100	10	71.584	0.679	0.000
500	10	65.668	0.660	0.000
1000	10	93.063	0.734	0.000
2000	10	112.235	0.769	0.000
5000	10	121.524	0.783	0.000
10000	10	103.234	0.754	0.000
Null	10	108.914	0.764	0.000

Table 12

ANOVA Table - 80% Training Cut off (Set 2, Experiment 3, Modification Six)

PMR	*df*	*F*	ω^2	*p*
Wall stimulus element and No Background R+				
10	10	19.636	0.359	0.000
15	10	23.479	0.403	0.000
20	10	34.116	0.498	0.000
30	10	32.094	0.483	0.000
40	10	31.135	0.475	0.000
50	10	21.003	0.375	0.000
75	10	25.099	0.420	0.000
100	10	23.511	0.403	0.000
500	10	35.068	0.505	0.000
1000	10	44.059	0.564	0.000
2000	10	50.475	0.597	0.000
5000	10	54.700	0.617	0.000
10000	10	49.245	0.591	0.000
Null	10	36.695	0.517	0.000
No Wall stimulus element and No Background R+				
10	10	36.943	0.519	0.000
15	10	64.877	0.657	0.000
20	10	56.104	0.623	0.000
30	10	87.314	0.721	0.000
40	10	61.064	0.643	0.000
50	10	95.341	0.739	0.000
75	10	82.263	0.709	0.000
100	10	63.148	0.651	0.000
500	10	106.307	0.760	0.000
1000	10	140.514	0.807	0.000
2000	10	133.665	0.799	0.000
5000	10	135.955	0.802	0.000
10000	10	103.775	0.755	0.000
Null	10	83.999	0.713	0.000

Table 13

ANOVA Table – Primary Contrast (Set 3, Experiment 3, Modification Six)

PMR	*df*	*F*	ω^2	*p*
Wall stimulus element and Background R+				
10	1	4.935	0.009	0.034
15	1	5.529	0.010	0.025
20	1	6.290	0.012	0.018
30	1	5.531	0.010	0.025
40	1	6.093	0.011	0.019
50	1	6.214	0.011	0.018
75	1	7.180	0.015	0.012
100	1	5.591	0.010	0.025
500	1	5.862	0.011	0.022
1000	1	6.214	0.011	0.018
2000	1	6.130	0.011	0.019
5000	1	5.453	0.008	0.026
10000	1	5.049	0.006	0.032
Null	1	5.068	0.007	0.032
No Wall stimulus element and Background R+				
10	1	5.995	0.009	0.020
15	1	5.563	0.007	0.025
20	1	4.921	0.005	0.034
30	1	5.671	0.006	0.024
40	1	4.805	0.004	0.036
50	1	4.853	0.004	0.035
75	1	5.239	0.005	0.029
100	1	5.748	0.005	0.023
500	1	5.513	0.005	0.026
1000	1	5.430	0.004	0.027
2000	1	5.089	0.003	0.032
5000	1	4.868	0.003	0.035
10000	1	5.068	0.003	0.032
Null	1	5.005	0.003	0.033

Table 14

ANOVA Table – Primary Contrast (Set 4, Experiment 3, Modification Six)

PMR	*df*	*F*	ω^2	*p*
Wall stimulus element and No Background R+				
10	1	5.822	0.010	0.022
15	1	5.373	0.009	0.027
20	1	4.707	0.007	0.038
30	1	4.804	0.007	0.036
40	1	5.205	0.008	0.030
50	1	5.958	0.011	0.021
75	1	5.992	0.010	0.020
100	1	5.643	0.010	0.024
500	1	5.728	0.008	0.023
1000	1	5.217	0.007	0.030
2000	1	5.232	0.006	0.029
5000	1	5.115	0.007	0.031
10000	1	5.143	0.007	0.031
Null	1	5.597	0.008	0.025
No Wall stimulus element and No Background R+				
10	1	6.006	0.008	0.020
15	1	4.958	0.005	0.034
20	1	5.707	0.006	0.023
30	1	5.116	0.004	0.031
40	1	5.151	0.005	0.031
50	1	5.243	0.004	0.029
75	1	5.142	0.004	0.031
100	1	5.341	0.005	0.028
500	1	5.073	0.003	0.032
1000	1	4.721	0.002	0.038
2000	1	4.916	0.003	0.034
5000	1	4.959	0.003	0.034
10000	1	5.079	0.003	0.032
Null	1	10.188	0.008	0.003

Figure 1

Example of Stimulus Discrimination in a Multiple Schedule

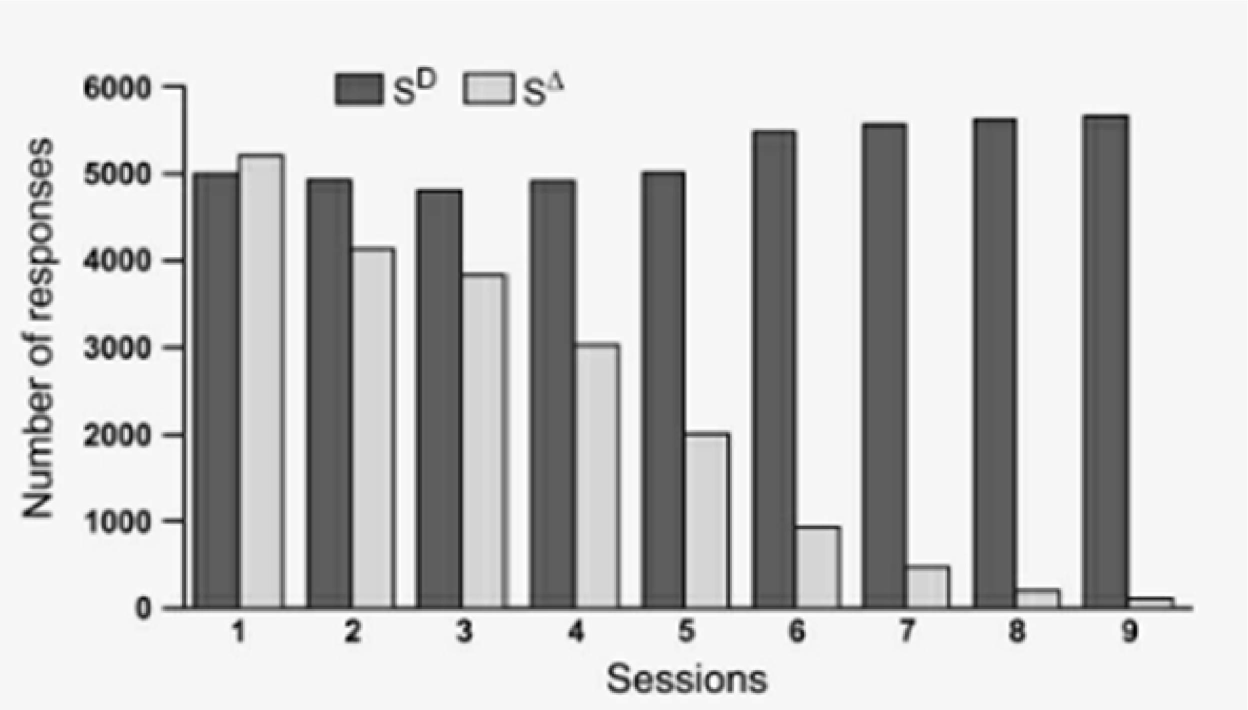


Figure 1 Note. Idealized experimental results are shown for a multiple schedule with VI 2-min EXT schedule of reinforcement. Relative to the VI 2-min (S^D), responses decline over sessions to almost zero responses per minute in the extinction (S^Δ) component. Reprinted from *Behavior analysis and learning* (6th ed.). p.262. by Pierce, W. D., & Cheney, C. D., 2013. Psychology Press. Copyright 2017 by Taylor & Francis

Figure 2

Figures from Guttman and Kalish 1956

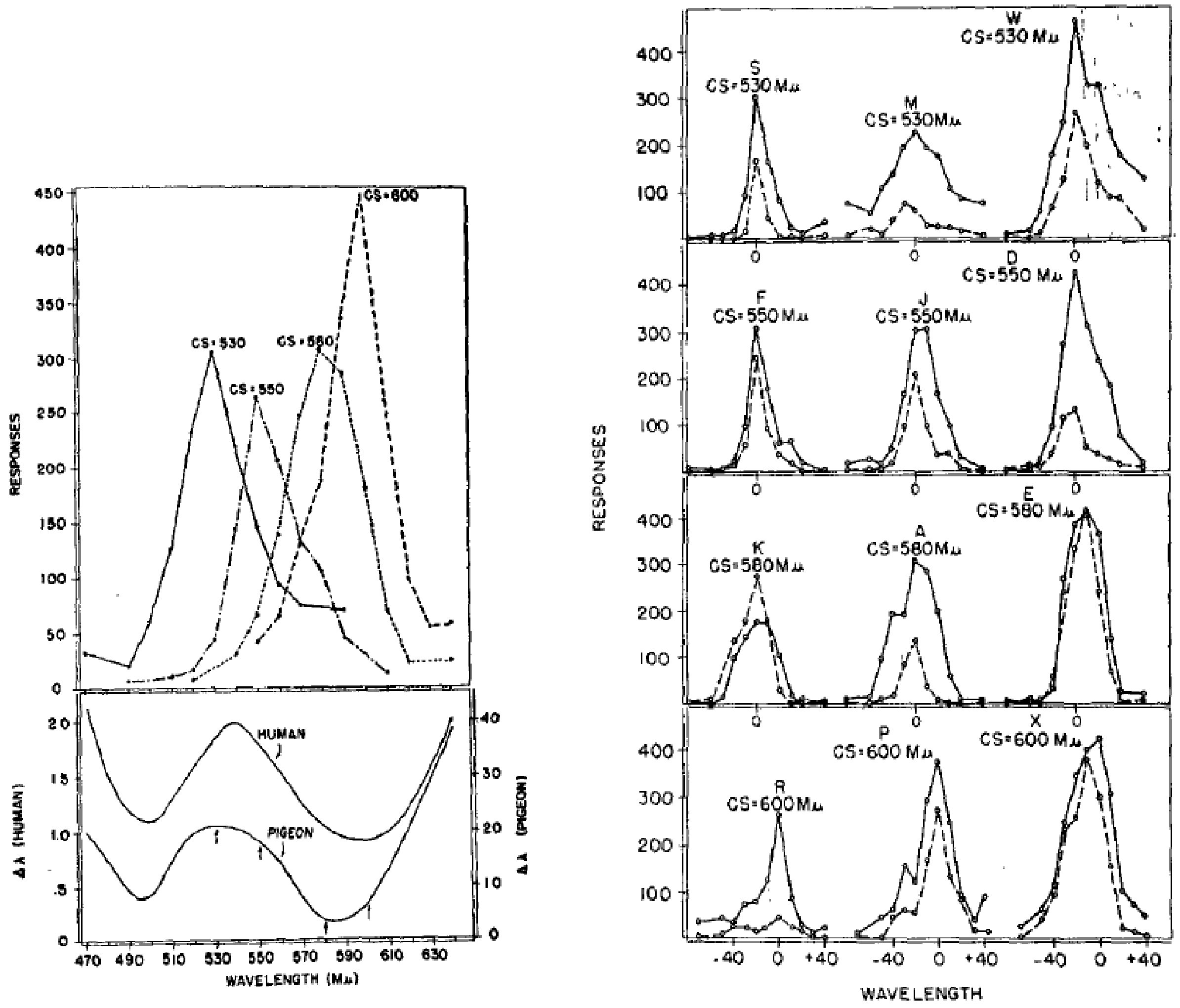


Figure 2 Note. From “Discriminability and Stimulus Gneralization,” Norman Guttman and Harry I. Kalish, 1956. *Journal of Experimental Psychology*, 50(1), p. 81, 83.

Figure 3

Figure from Blough, P. M. 1972

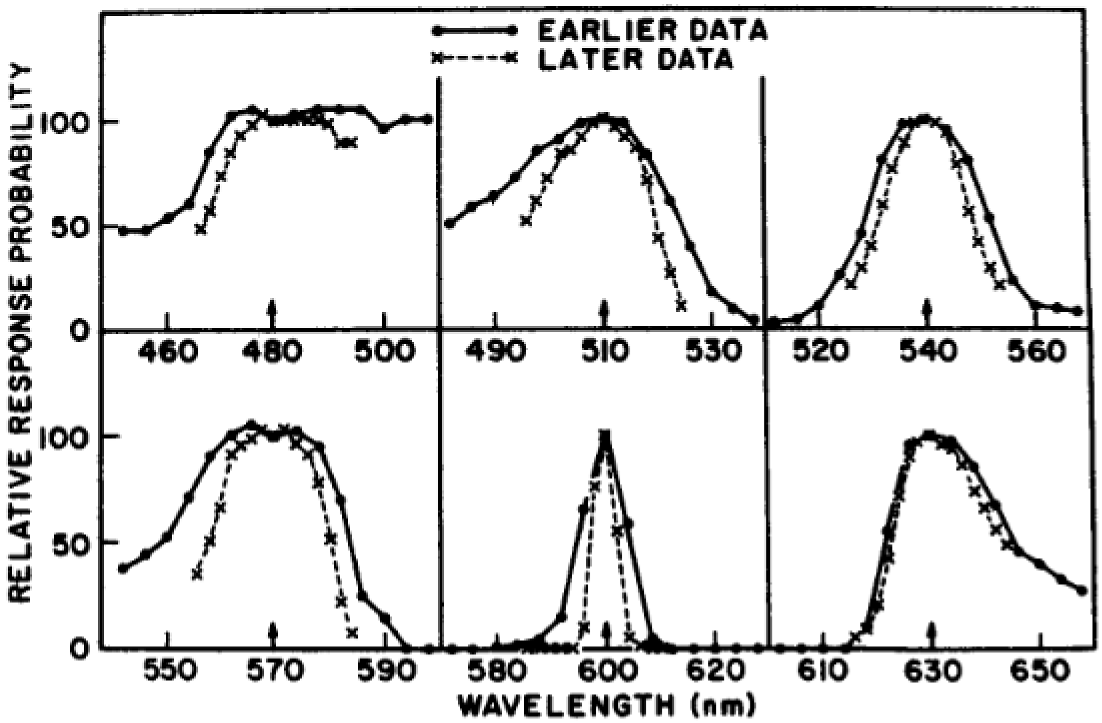


Figure 3 Note. From “Wavelength Generalization and Discrimination in Pigeon,” Blough, P. M., 1972. *Perception and Psychophysics*, 12(4), p. 346

Figure 4

Figure from Thomas, D. R., & King, R. A., 1959

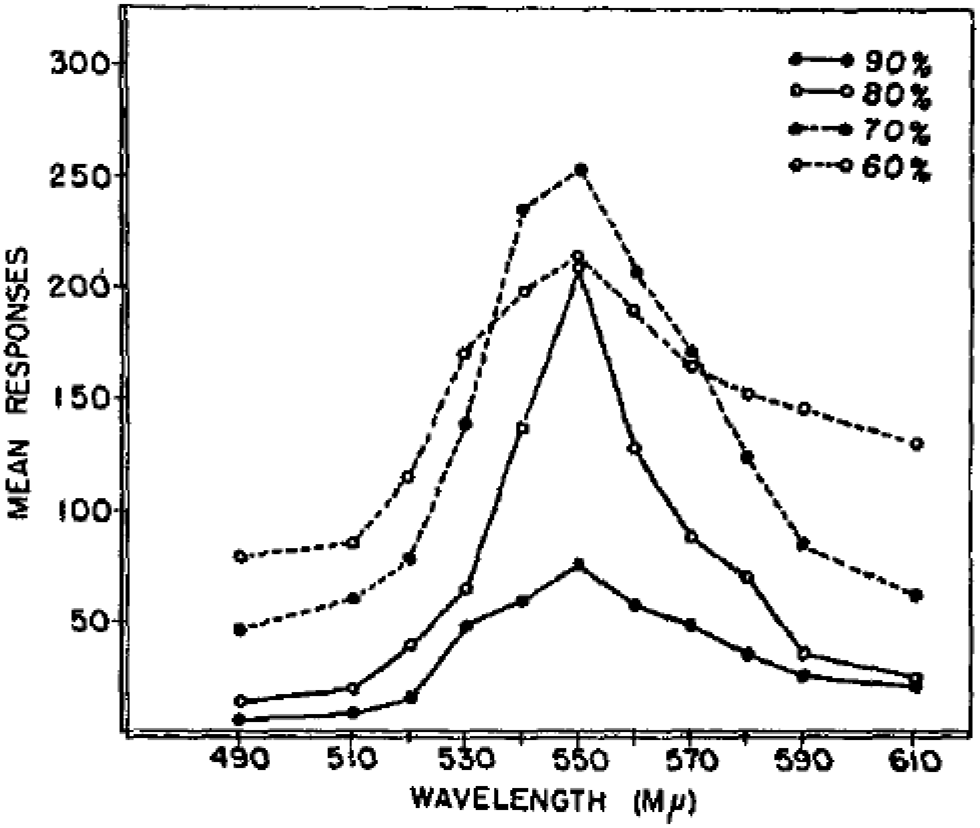


Figure 4 Note. From “Stimulus-Generalization as a Functions of Level of Motivation,” Thomas, D. R., & King, R. A., 1959. *Journal of Experimental Psychology,* 57(5), p. 324

Figure 5*Evolutionary Theory of Behavioral Dynamics - Flowchart*

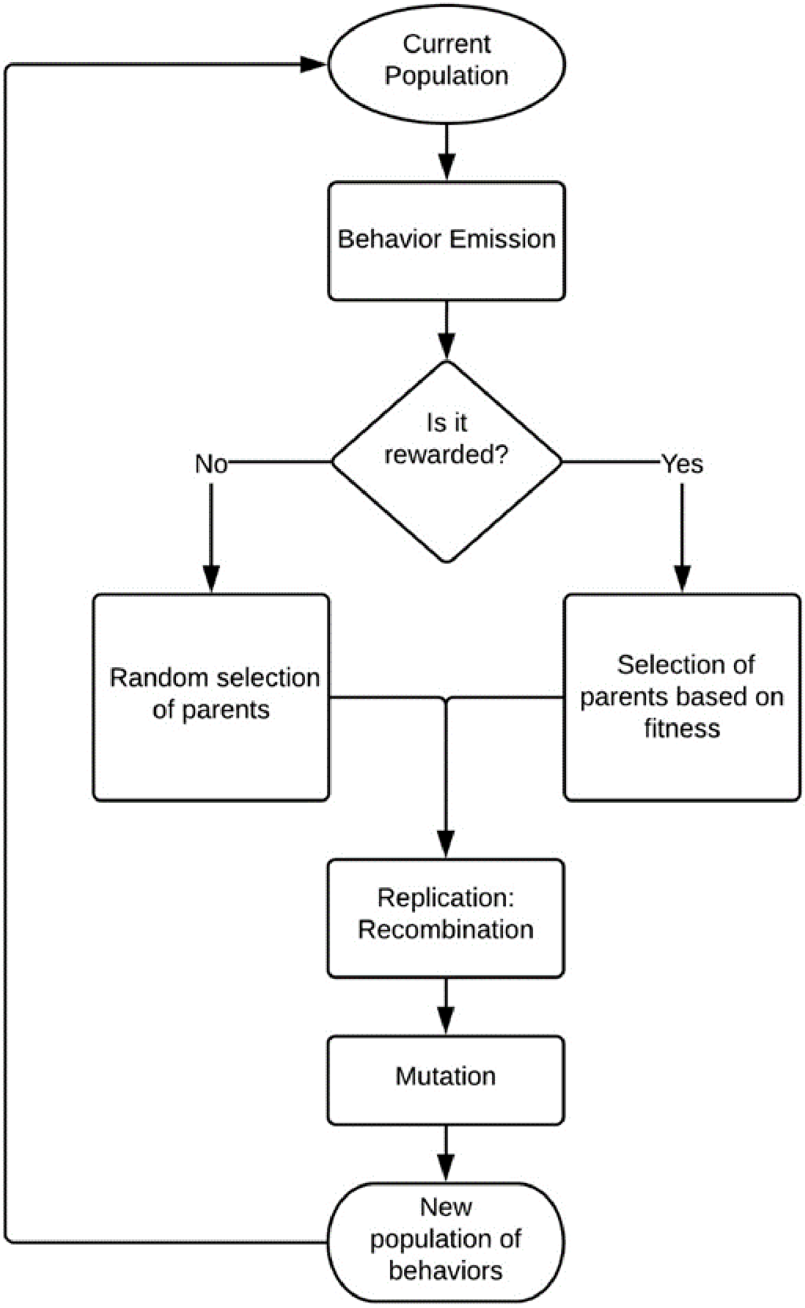


Figure 5 Note. The flowchart demonstrates one iteration of the ETBD.

Figure 6

SST Sampling Example

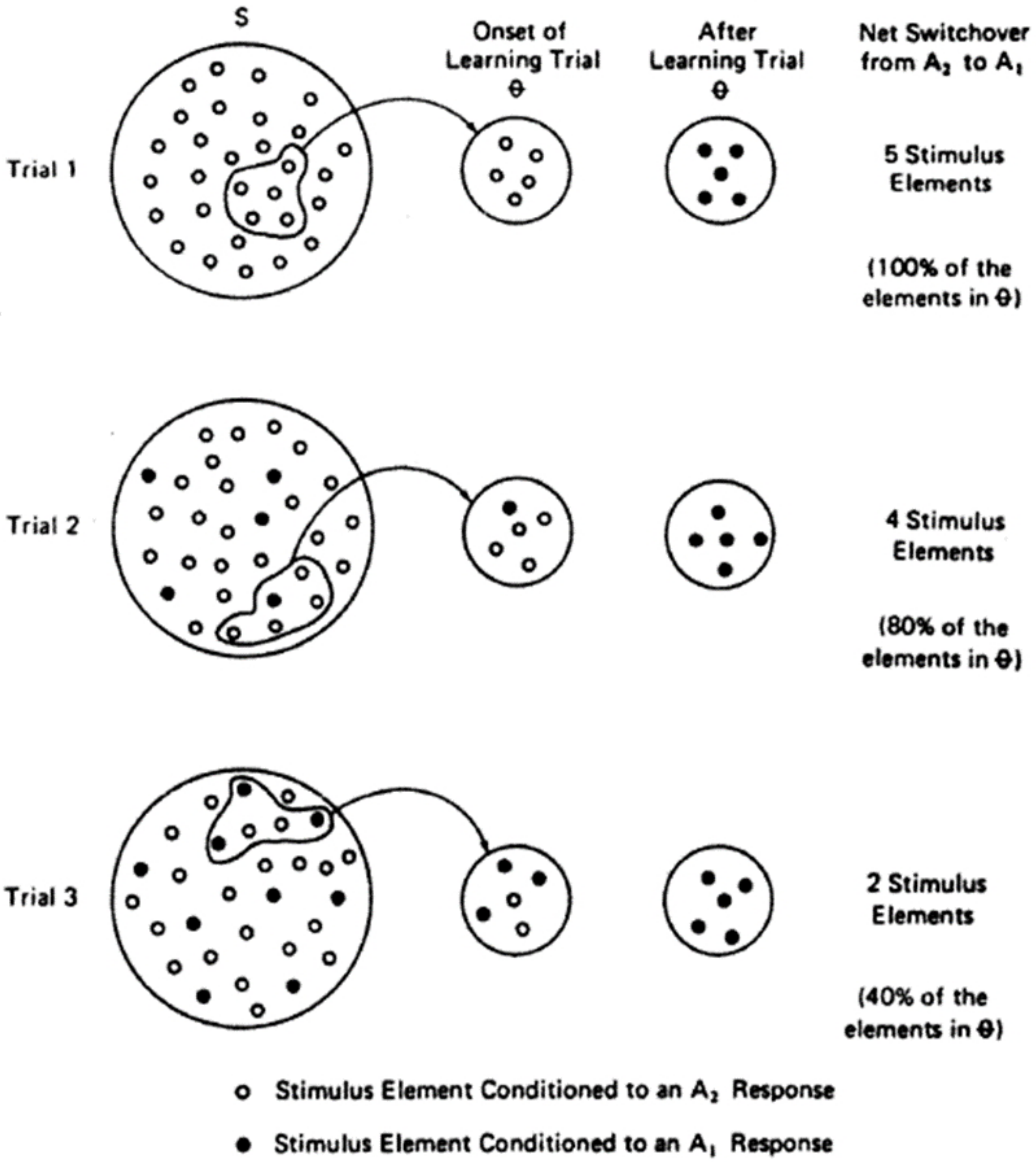


Figure 6 Note. Estes' model of how stimulus elements are sampled from the population of elements (Left circle). The figure shows three learning trials. The clear circles represent stimulus elements associated with response A_2, and the filled circles are stimulus elements associated with A_1. Initially, all of the stimulus elements are associated with the A_2 response class. In trial 1, five

stimulus elements are sampled from the environment (θ). They are then induced by the experimental set up (not shown) to be conditioned to response A_1 breaking their link to response A_2. This modifies the pool of available stimulus elements available during the next trial. During trial 2, five more stimulus elements are sampled at random from the pool. Of these stimulus elements, four of the five are associated with A_2, while one is associated with A_1 (due to the first learning trial). In this trial, four of the stimulus elements are conditioned to A_1. The third trial repeats the process. From *An Introduction to Theories of Learning* (4th ed., p. 230) by B. R. Hergenhahn and M. H. Olson, 1993, Englewood Cliffs, NJ: Prentice Hall.

Figure 7

The Modified ETBD Flow Chart

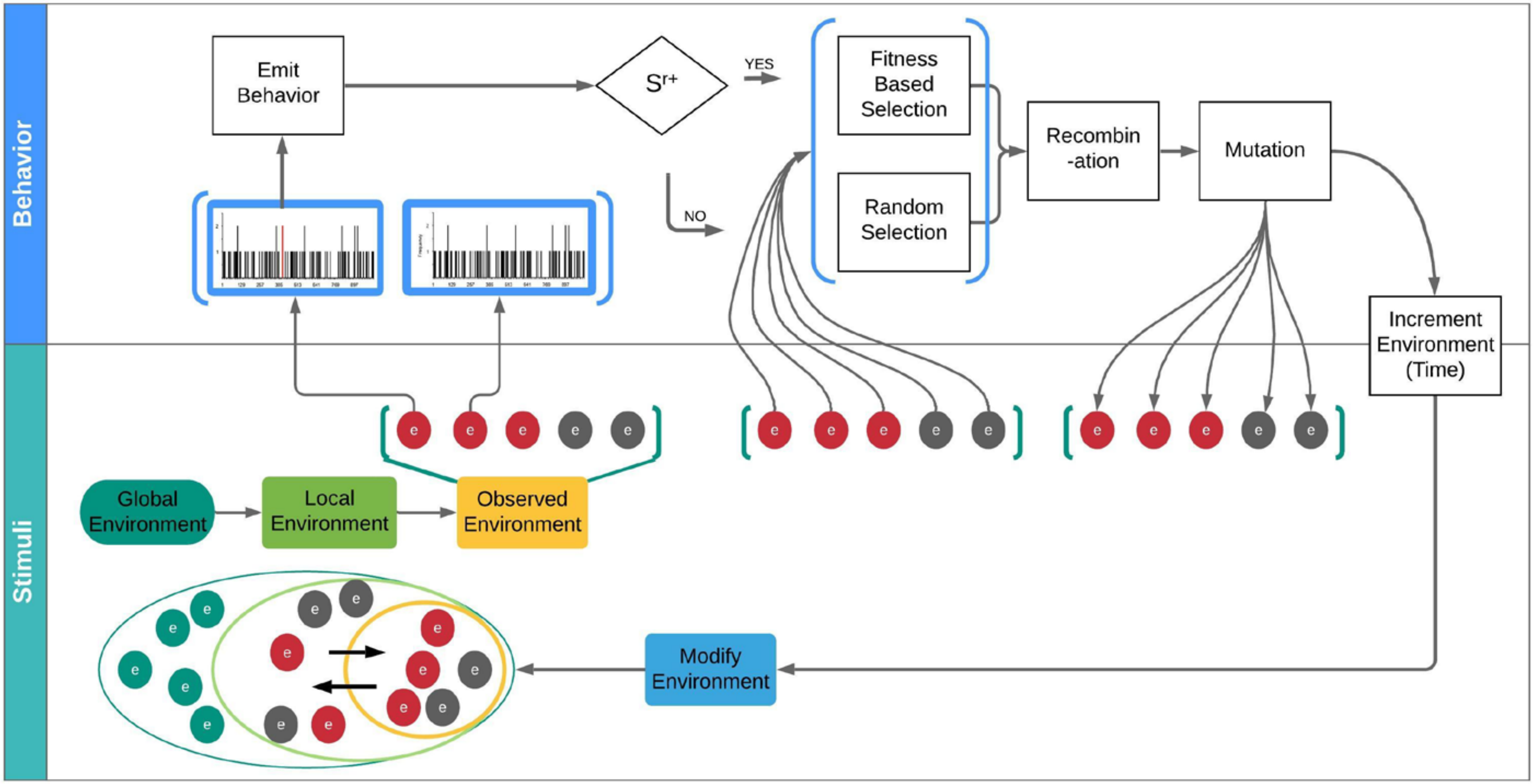


Figure 7 Note. The flow chart begins in the bottom left corner. The AO samples stimulus elements from the local environment and then one stimulus element from the sample. One behavior is emitted from the population of behaviors linked to that stimulus element. If the behavior allows the AO to obtain reinforcement (S^{r+}) then all populations of behaviors linked to the observed stimulus elements will undergo fitness-based selection. Otherwise, they will undergo random selection. Each observed population will individually undergo recombination and mutation, resulting in new populations. After the next time tick, the program will check whether the environment needs to be modified based on the reinforcement schedule being used. Afterward, the cycle repeats.

Figure 8

Histogram Comparison

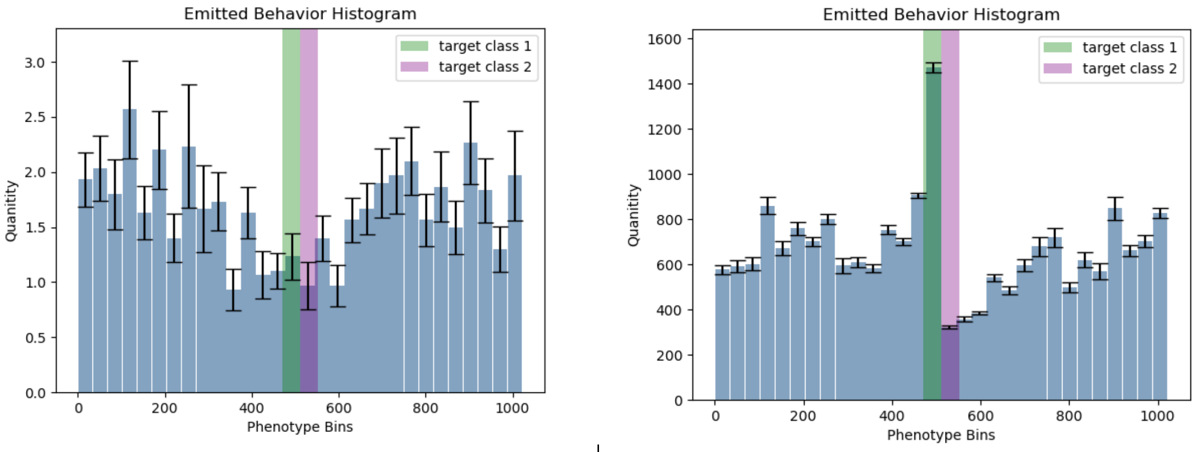


Figure 8 Note. The left figure is a histogram of 50 generations of behavior at the beginning of a schedule. The right figure is a histogram of all behavior within a schedule.

Figure 9

Average Matching Log Plot - ETBD Replication

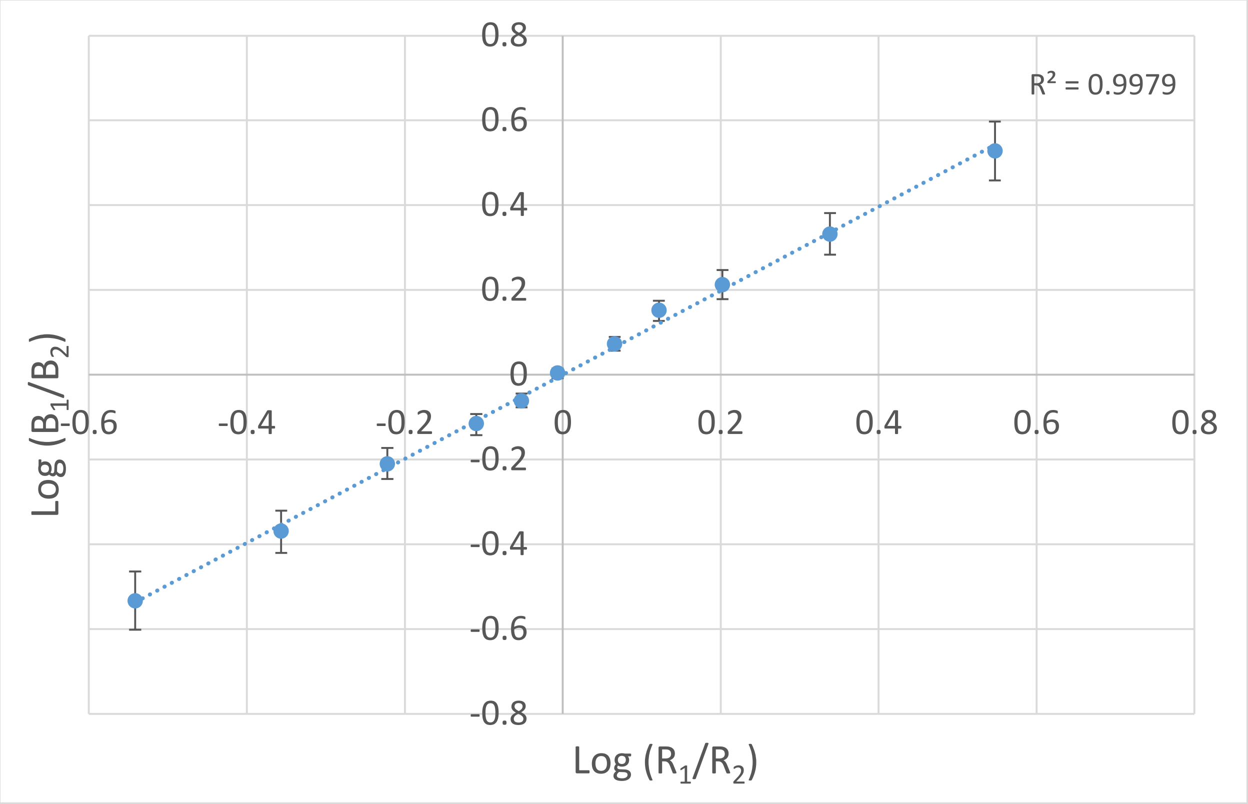


Figure 10

Matching sensitivity based on observation style and number of stimulus elements

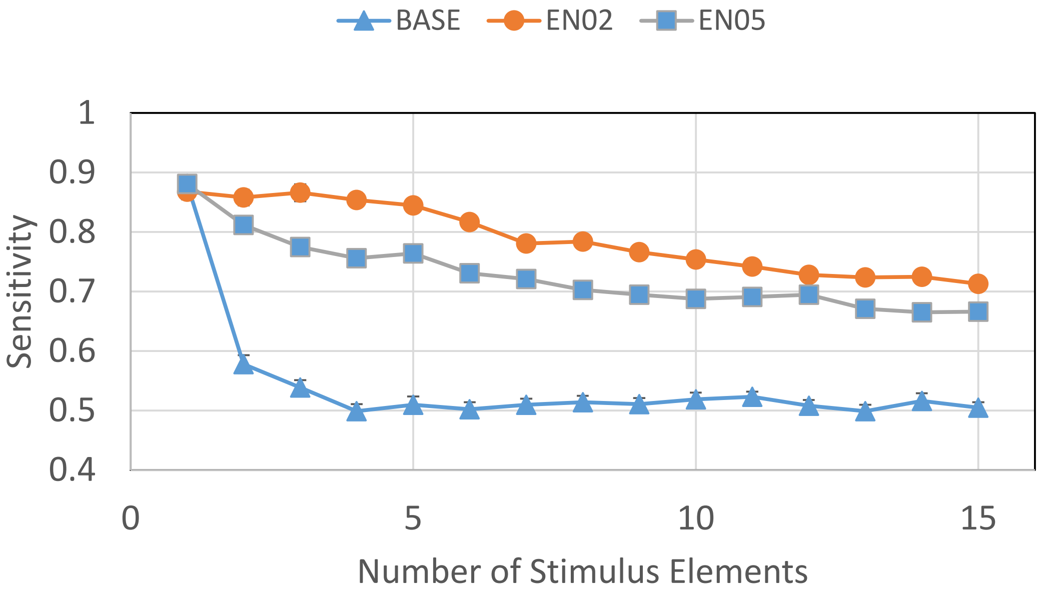


Figure 10 Note. Each unique marker indicates a specific observation style. Each point on the graph shows the average sensitivity of 30 AOs. 90% confidence intervals are marked on the graph, with the majority of the intervals being smaller than the marker itself.

Figure 11

Impact of Behavior Population Size on Matching Sensitivity

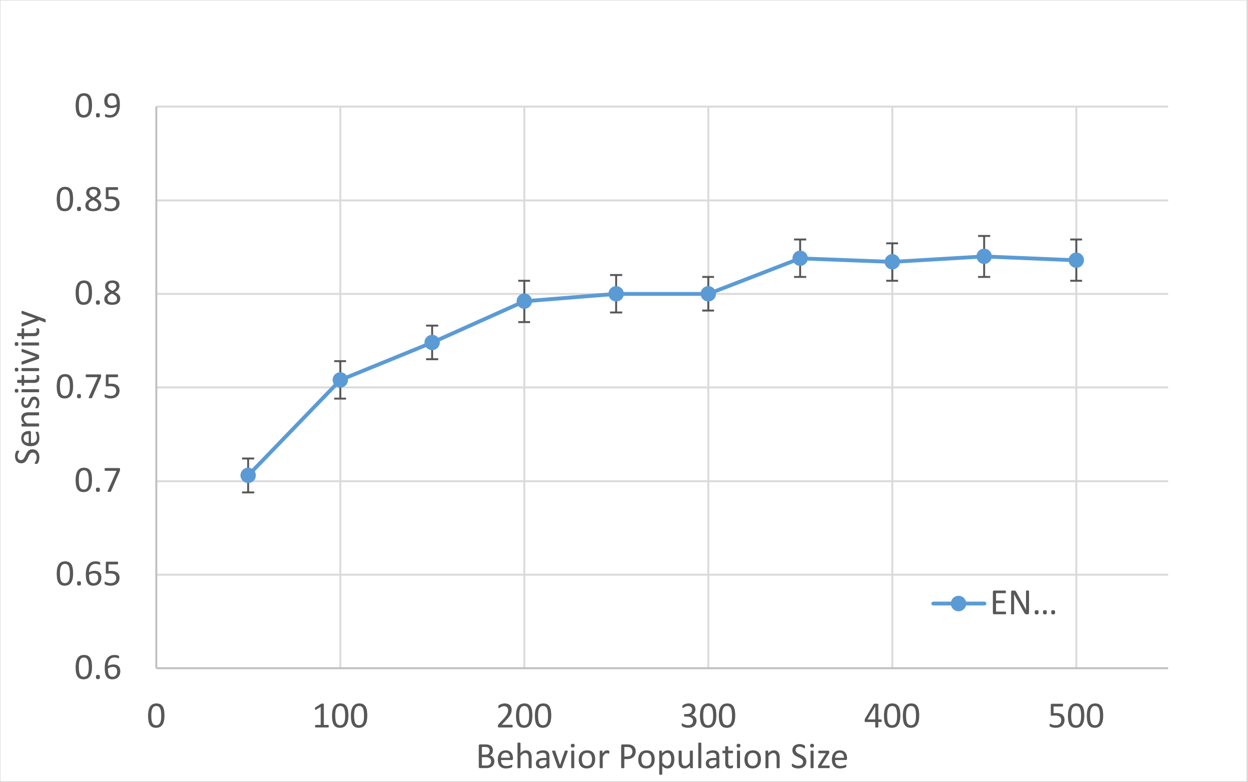


Figure 11 Note. Error bars represent 90% confidence intervals.

Figure 12

Average Matching Log Plot Experiment 1,

Conditions: stimulus element 10, EN02, and Behavior Population 200

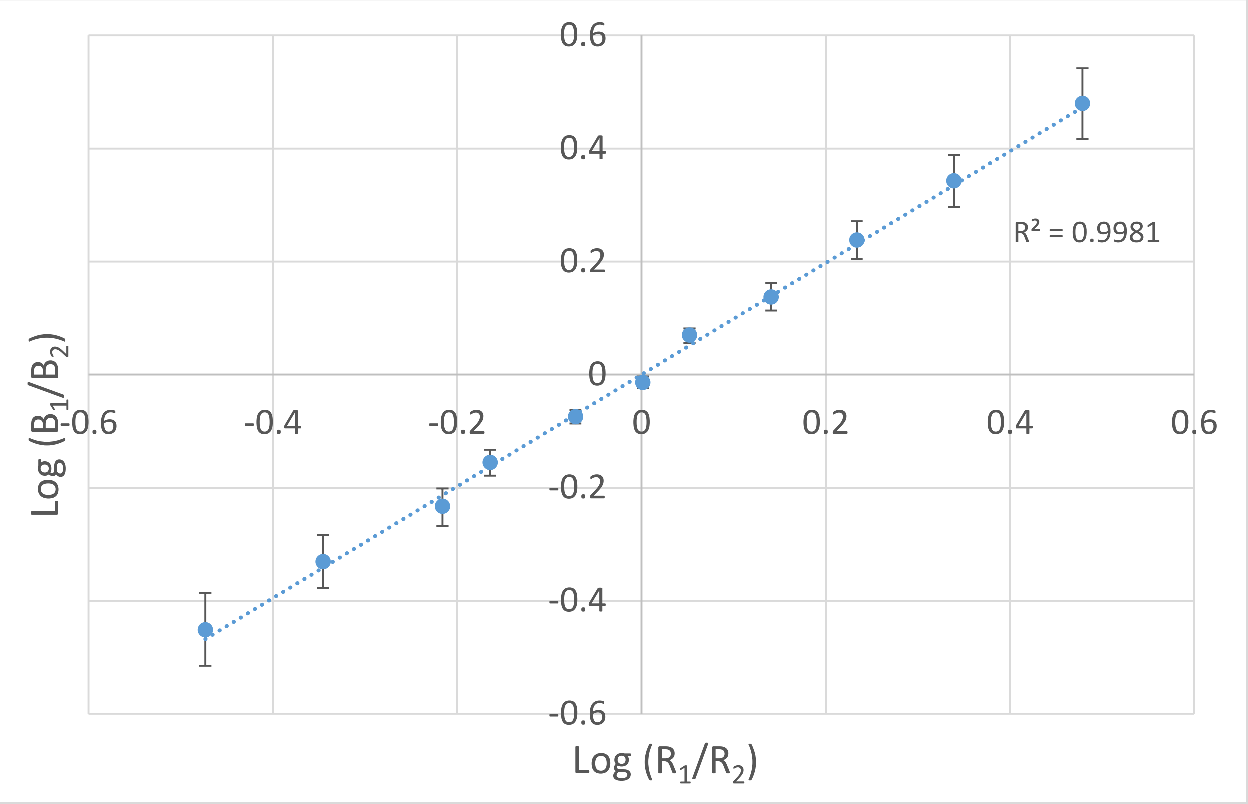


Figure 13

Behavior of AOs on Multiple Schedules

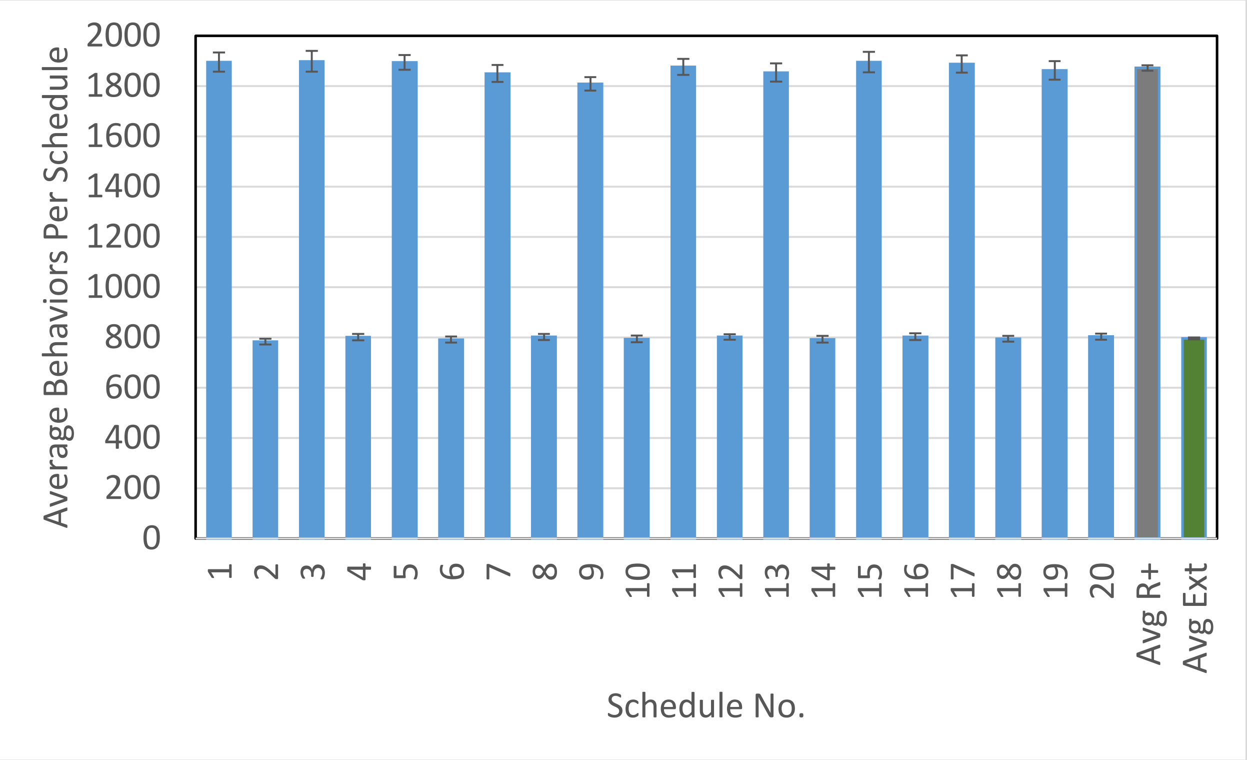


Figure 13 Note. Odd numbered schedules were reinforced with a RI 70 schedule, while even numbered schedules were on extinction. Error bars denote 90% confidence intervals. The average stimulus element+ shows the average amount of behavior across all schedules with reinforcement. The average stimulus element- shows the average amount of behavior across all extinction schedules.

Figure 14

Impact of background reinforcement on reinforced and extinction schedules

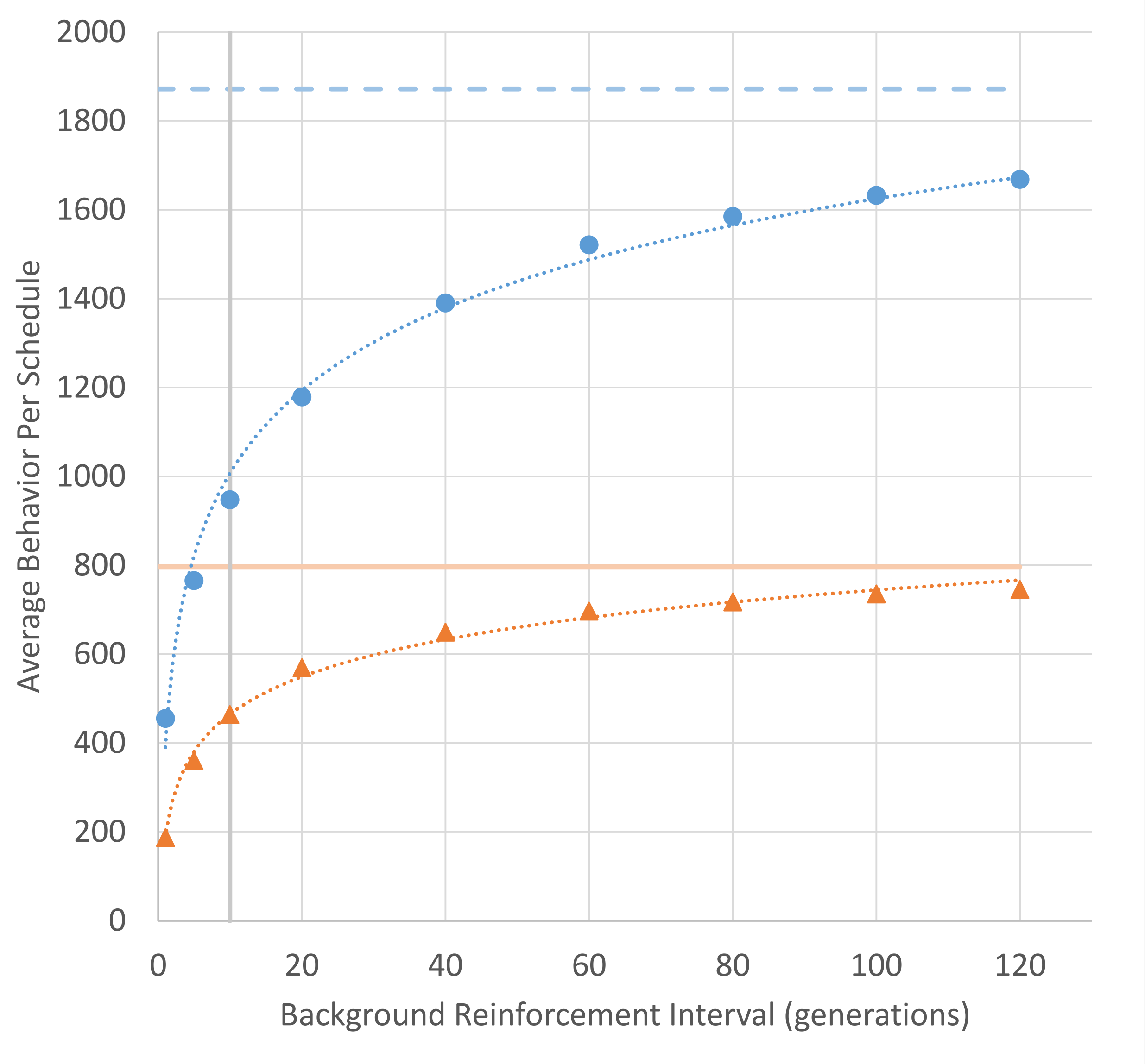


Figure 14 Note. Impact of background reinforcement rates on average rates of behavior during schedules with reinforcement (circle), and average rates of behavior during extinction schedules (triangle) is shown. The baseline rates of behavior for reinforced schedules (dashed line) and extinction schedules (solid horizontal line) are shown. Logarithmic trends for each group are shown by dotted lines.

Figure 15

Impact of background reinforcement rate on matching sensitivity

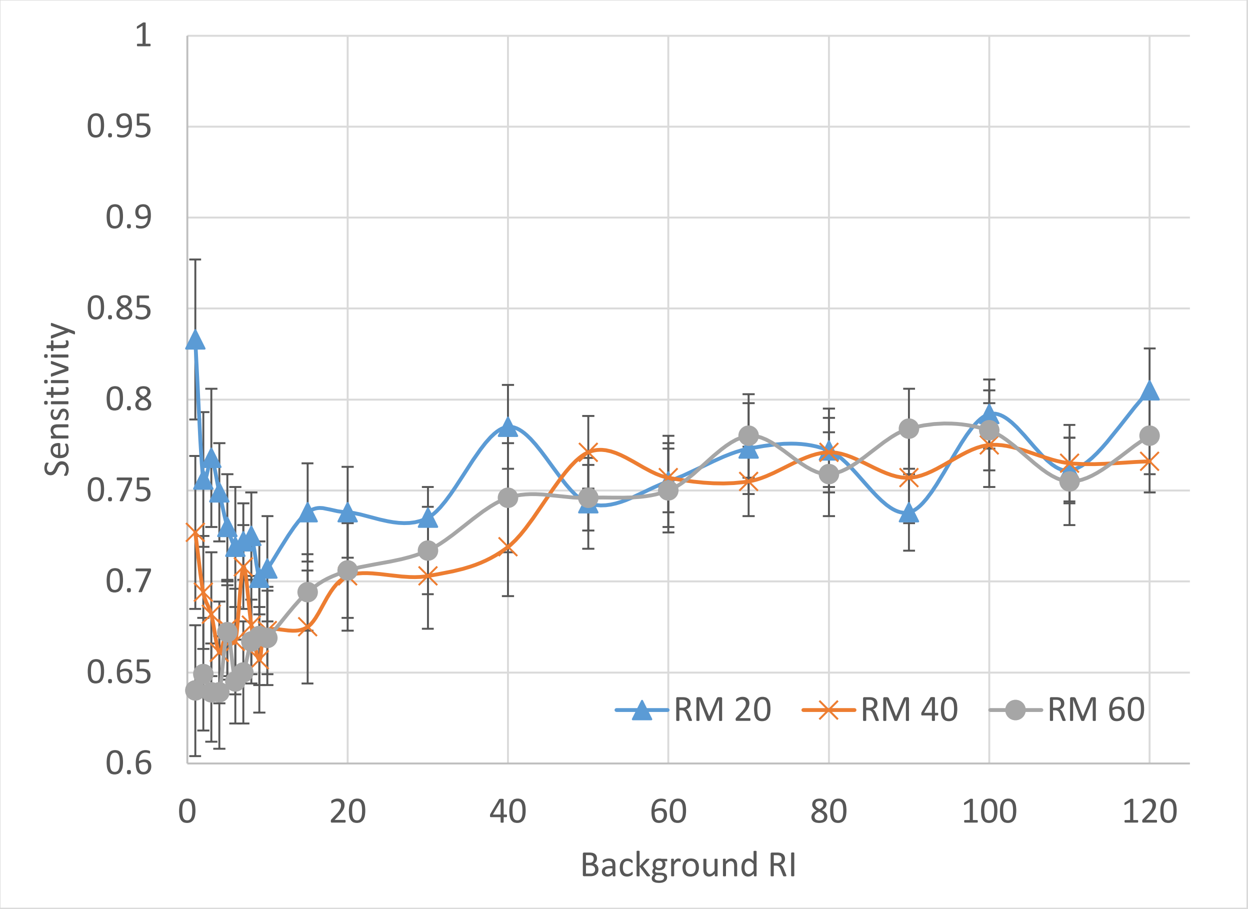


Figure 16

Absence of constant stimulus element has no impact on average rates of behavior (Exp 2, Dev 4)

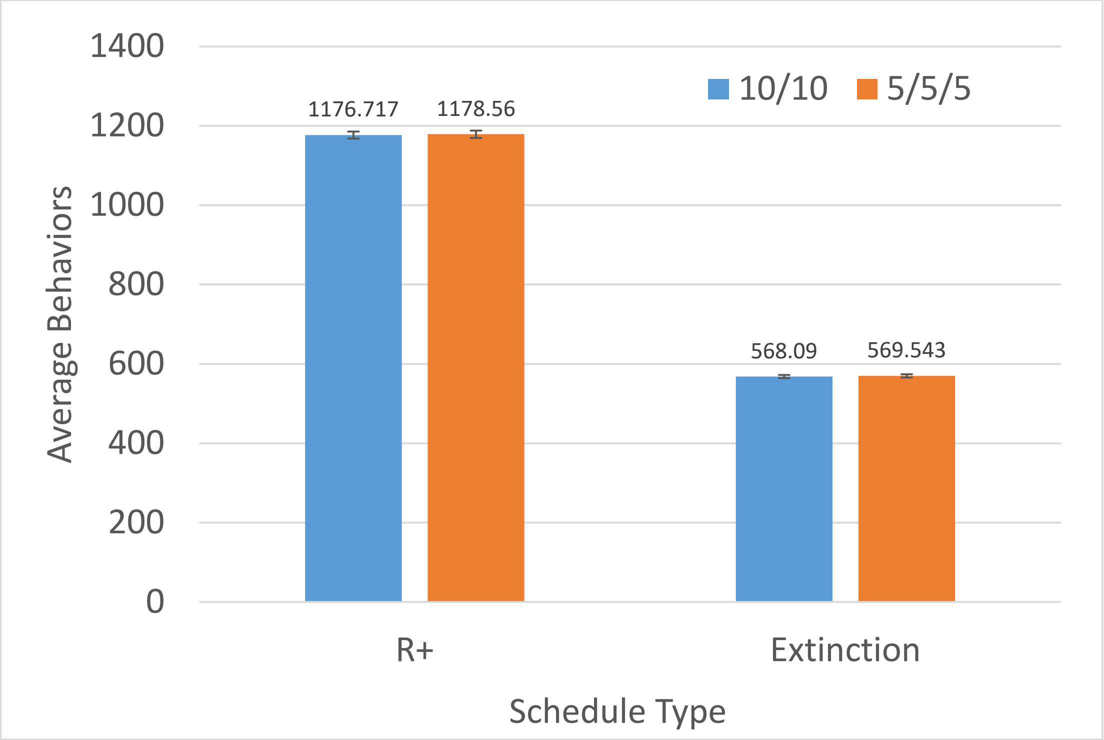


Figure 16 Note. The average number of behaviors across schedule types is show for two stimulus element presentation styles. While both styles have 10 stimulus elements present in the local environment at any one time, the ‘10/10’ switches completely between red stimuli and green stimuli, while the ‘5/5/5’ style has five green and five wall stimuli during the reinforced schedules, and five red stimuli and the same five wall stimuli present during extinction schedules.

Figure 17.

Initial learning (Experiment Series 2, Modification 4)

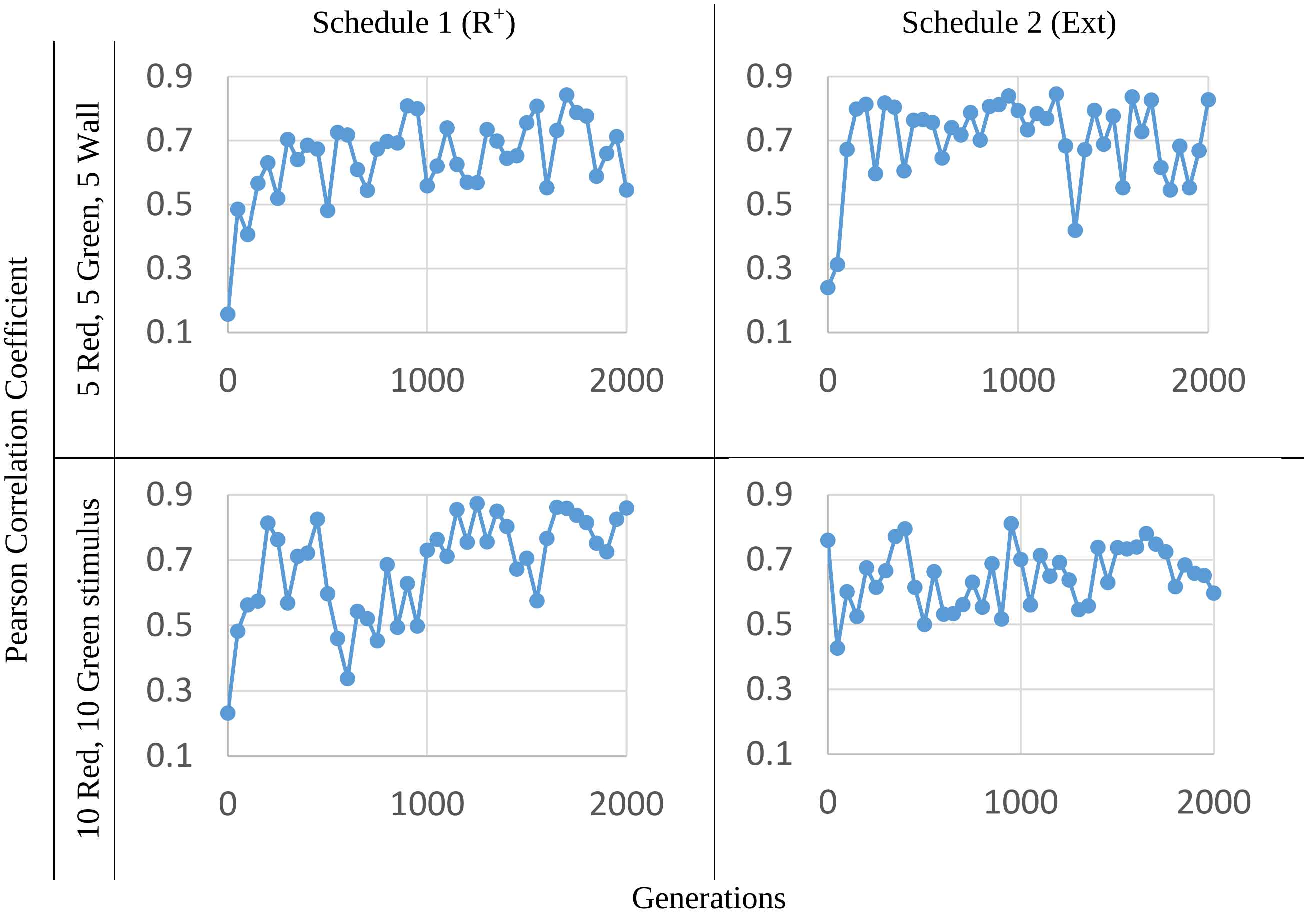


Figure 18

Stimulus Element quantity impact on behavior during reinforced schedules (Exp II, Dev 4)

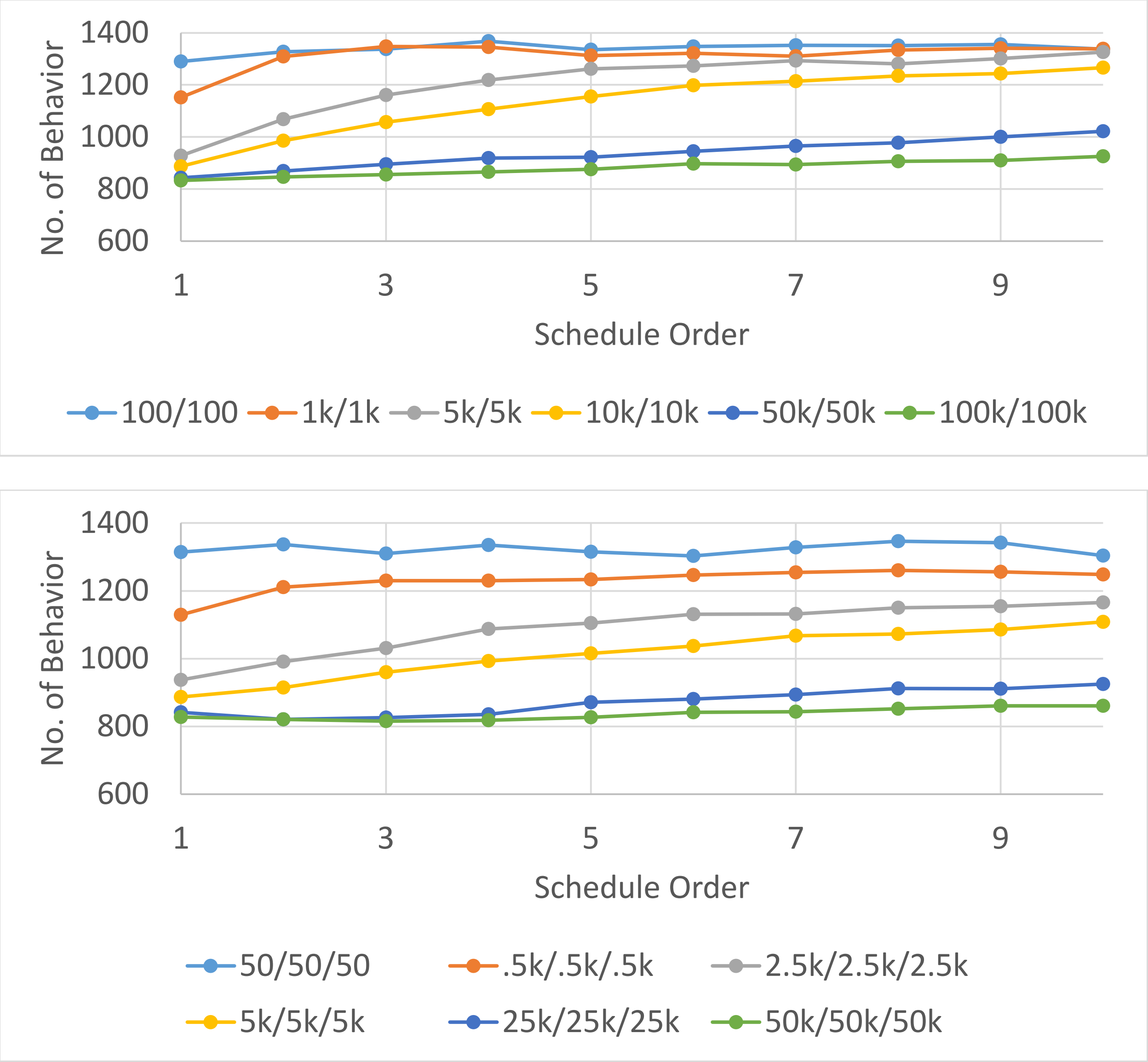


Figure 18 Note. Each point represents the average number of behaviors during one schedule. Only the schedules with reinforcement are shown here. The ratios in the legend represent the number of stimulus elements of each type. The ratios with two items are red and green stimulus elements and the ratios with three items represent red, green, and wall stimulus elements.

Figure 19

Stimulus Element quantity impact on behavior during extinction schedules (Exp II, Dev 4)

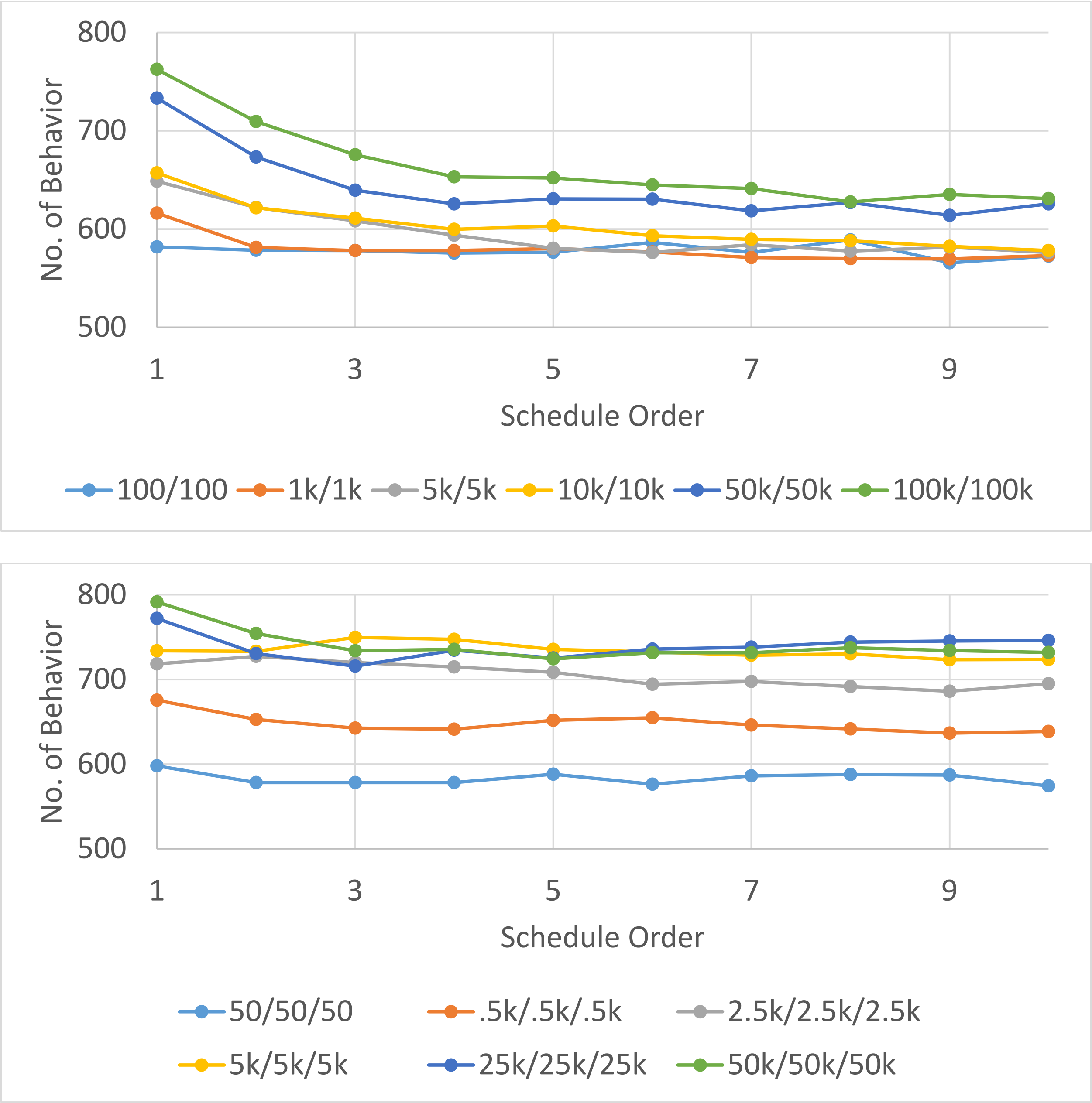
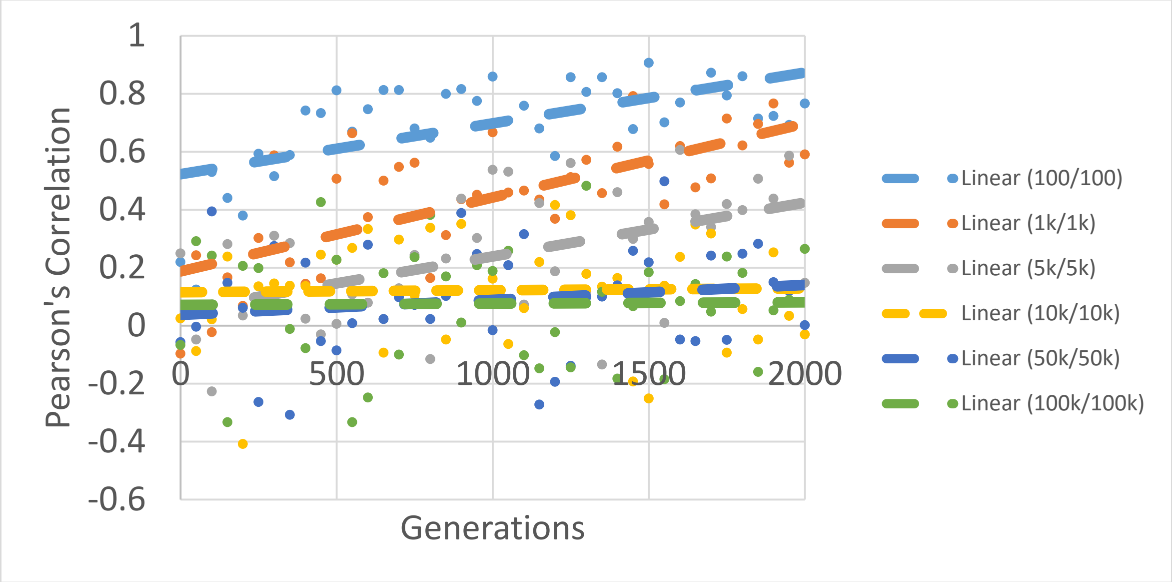


Figure 19 Note. Each point represents the average number of behaviors during one schedule. Only the schedules without reinforcement are shown here. The ratios in the legend represent the number of stimulus elements of each type. The ratios with two items are red and green stimulus elements and the ratios with three items represent red, green, and wall stimulus elements.

Figure 20.

Impact of stimulus element Quantity on learning during first R$^+$ Schedule (Exp 2, Dev 4)



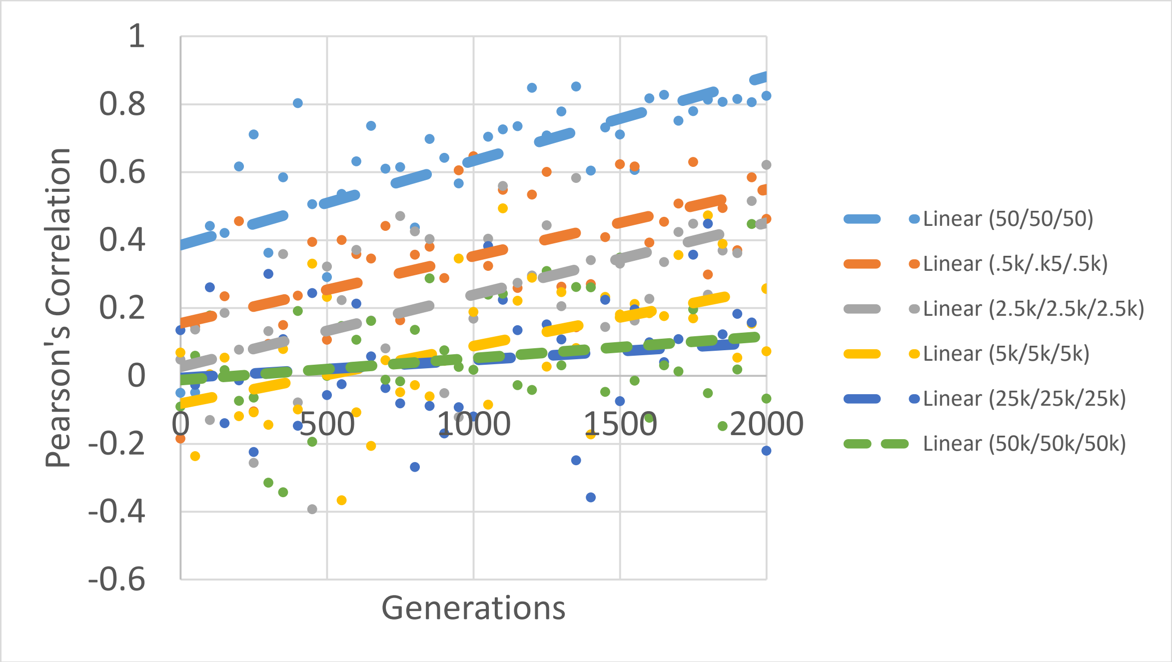
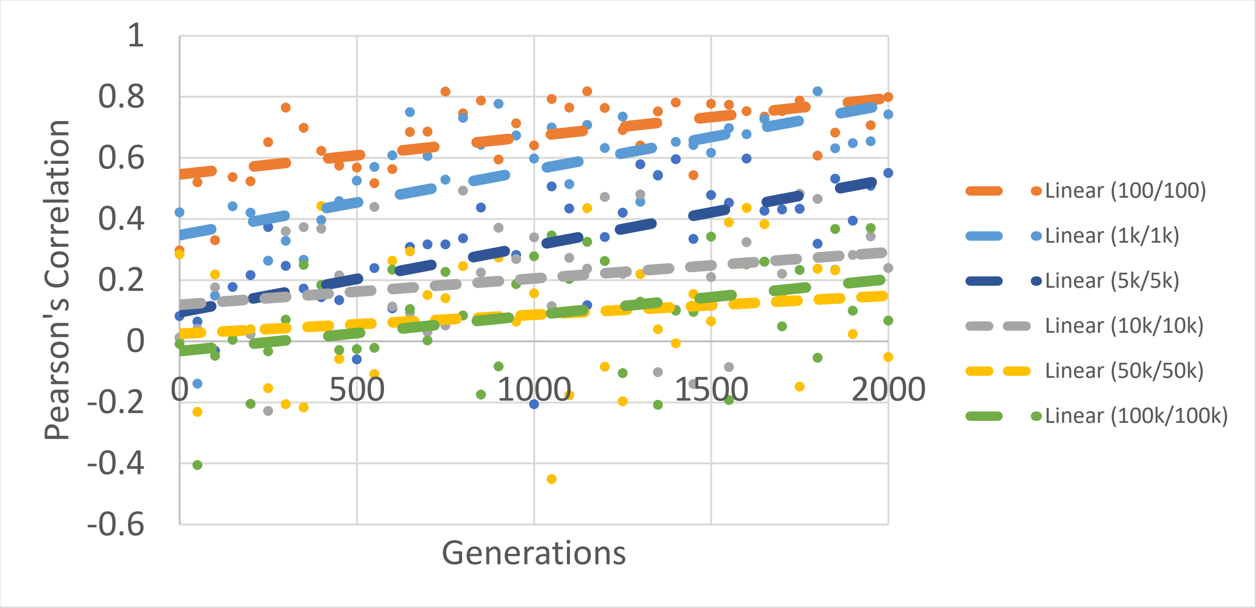


Figure 21

Impact of stimulus element quantity on learning during first Extinction Schedule (Exp 2, Dev 4)



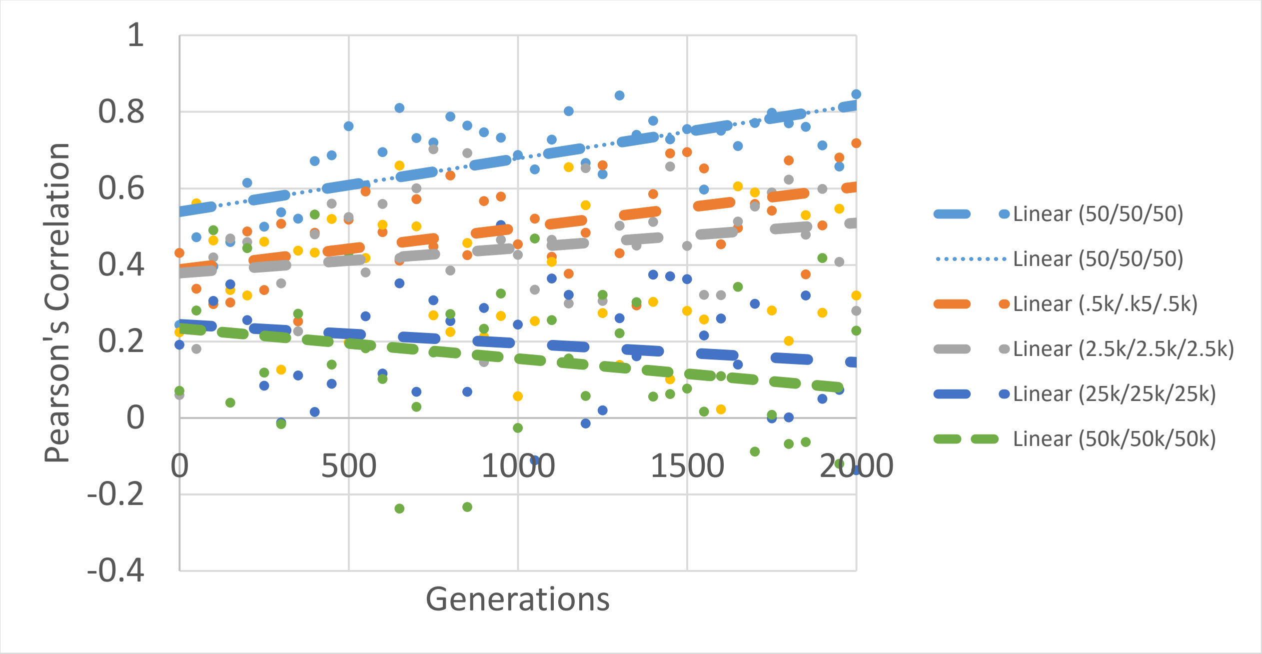


Figure 22

Behavior on the most optimal conditions

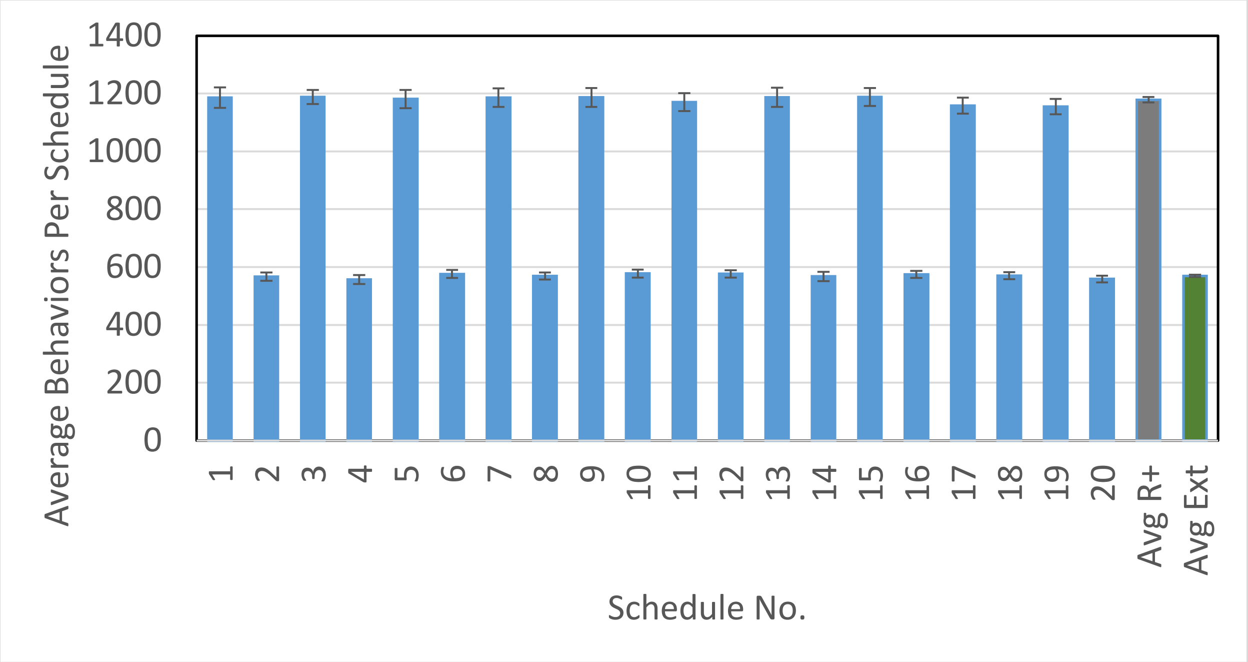


Figure 22 Note. Odd numbered schedules were reinforced with a RI 70 schedule, while even numbered schedules were on extinction. Error bars denote 90% confidence intervals. The average R+ shows the average amount of behavior across all schedules with reinforcement. The average Ext shows the average amount of behavior across all extinction schedules.

Figure 23

Reinforcement Context Kernel

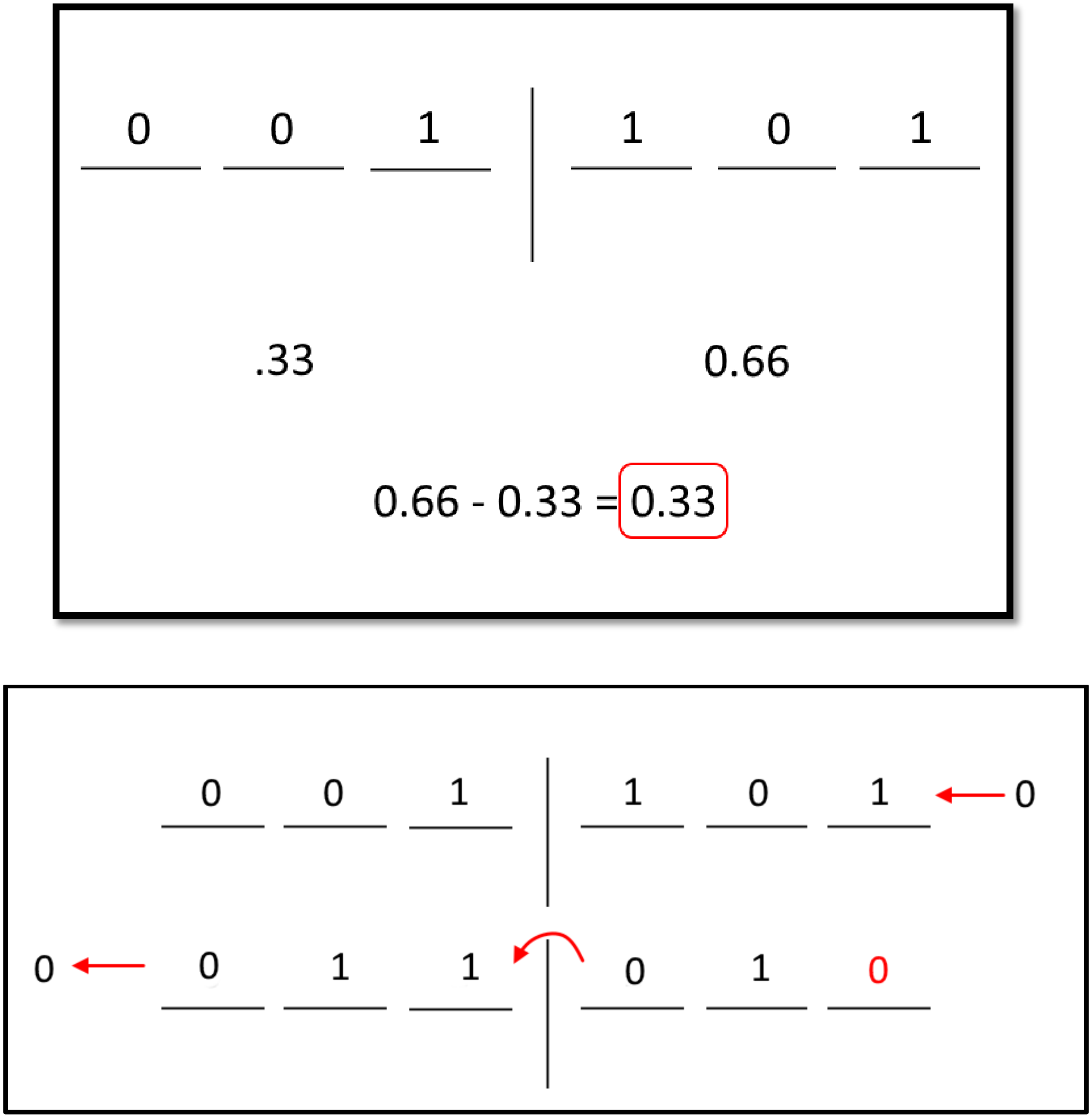


Figure 23 Note. The calculation for stability (top) and the adding of new information each generation (bottom)

Figure 24

Window Size Adjustment

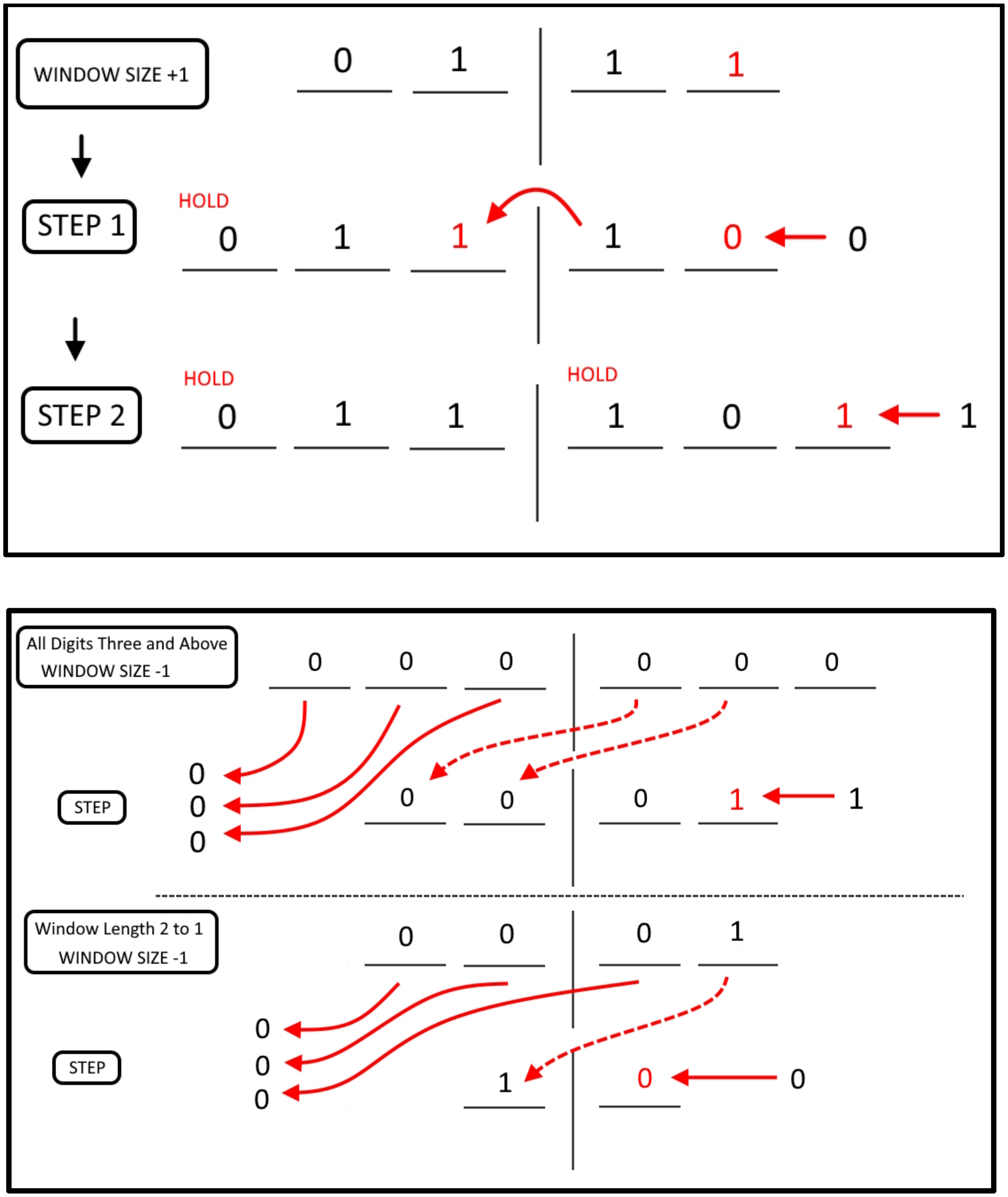


Figure 24 Note. The two steps needed for increasing the window size (top) and the single steps required to decrease the window size.

Figure 25

Stimulus Generalization Gradient using a Reinforcement Context Kernel

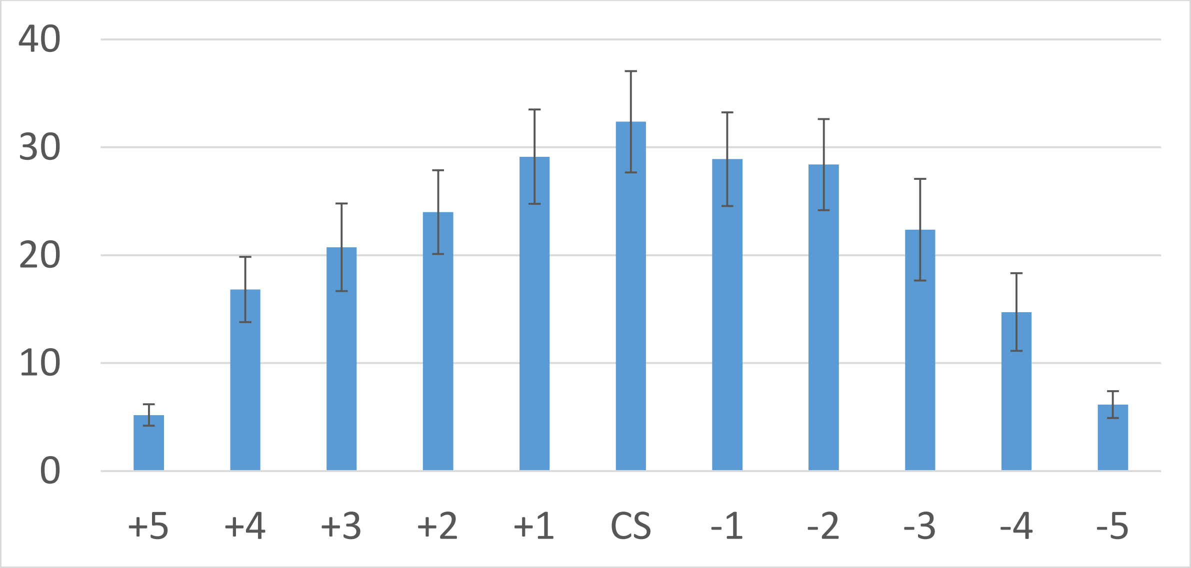


Figure 25 Note. Error bars represent 90% confidence intervals.

Figure 26

Discriminative Stimuli Conditions

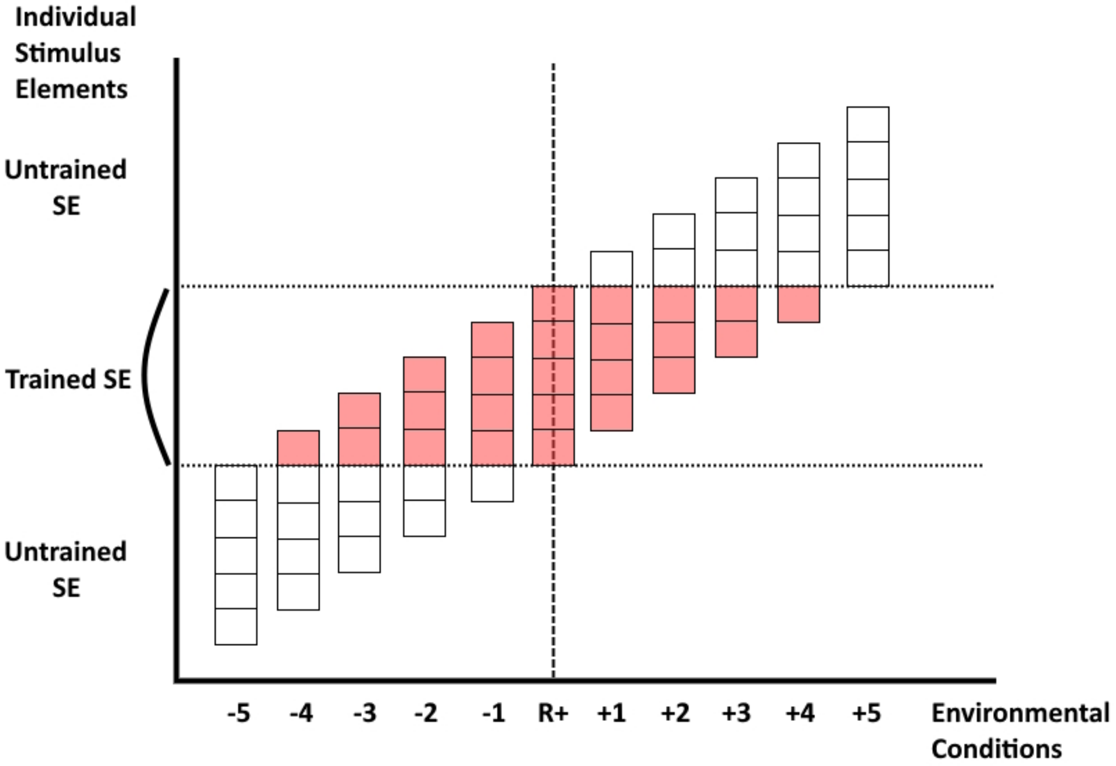


Figure 26 note. A visual representation of the mix of trained (pink shaded) and untrained stimuli per environmental condition. Each column represents a different environmental condition, with R+ signifying the reinforced condition. Each row represents a specific stimulus element.

Figure 27

Entropy to Selection Percentage Function Forms

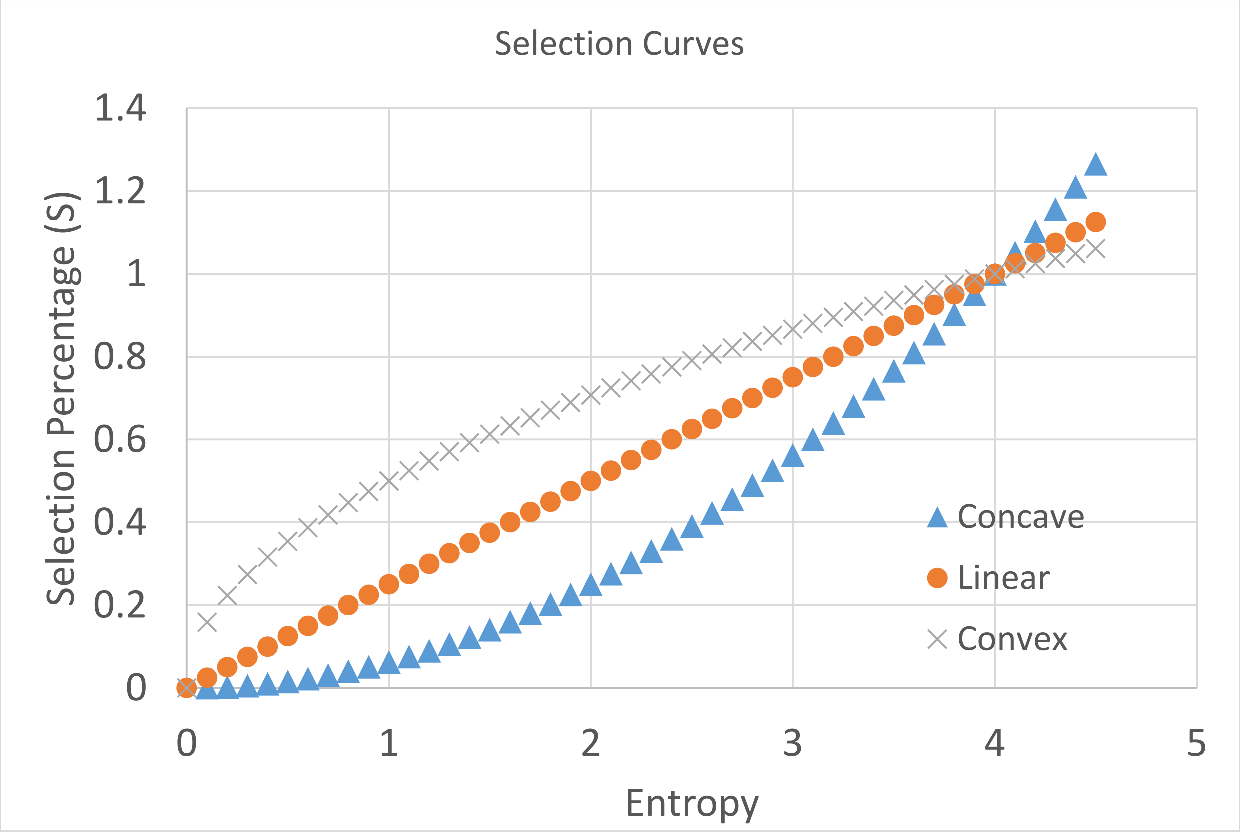


Figure 27 Note. All functions were generated by using the equation, $S = ax^b$. For the concave function (triangle), a = 0.0625 and b = 2. For the linear function (circle), a = 0.25 and b = 1. For the convex function (square), a = 0.5 and b = 0.5.

Figure 28

Stimulus generalization gradients (Experiment Three, Modification Five)

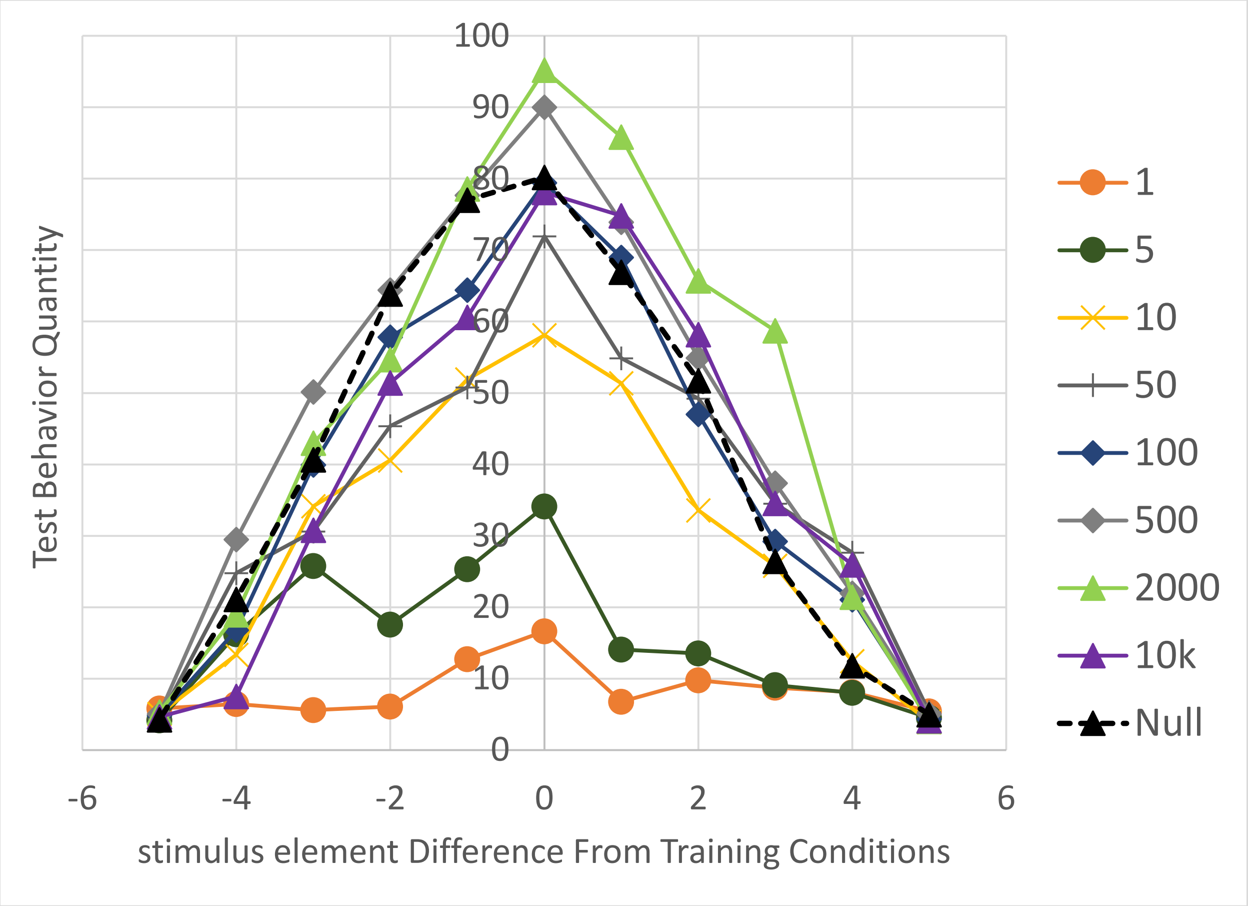


Figure 28 Note. The quantity in the legend denotes the frequency of mutation, if no child behaviors are created. In the condition, ‘1,’ every generation that no child behaviors are created there is also a single mutation event on the active behavior population In the “Null” condition, no general population mutations occur.

Figure 29

Stimulus generalization gradients behavior totals (Experiment Three, Modification Five)

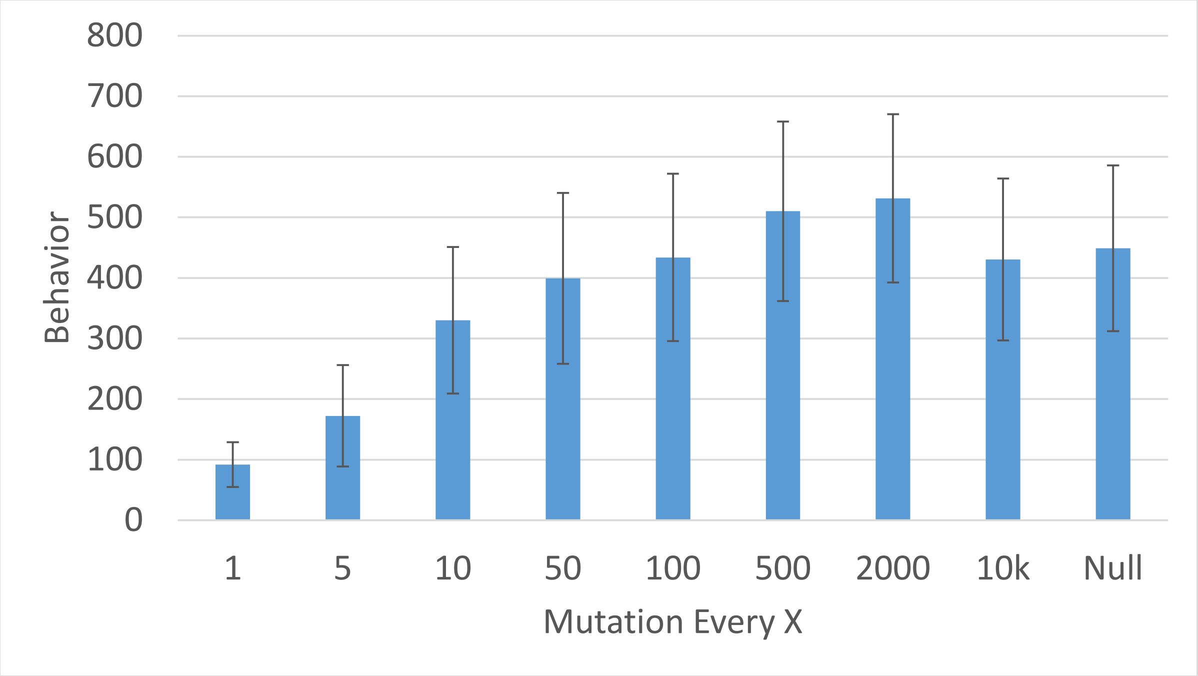


Figure 29 Note. The x-axis denotes the frequency of mutation, if no child behaviors are created. In the first column, ‘1,’ every generation that no child behaviors are created there is also a single mutation event on the active behavior population. In the “Null” column, no general population mutations occur. Error bars are the sum of the 90% confidence intervals for each condition, as seen in Figure 28.

Figure 30

AOs with under 10% behaviors on target during testing (Experiment Three, Modification Five)

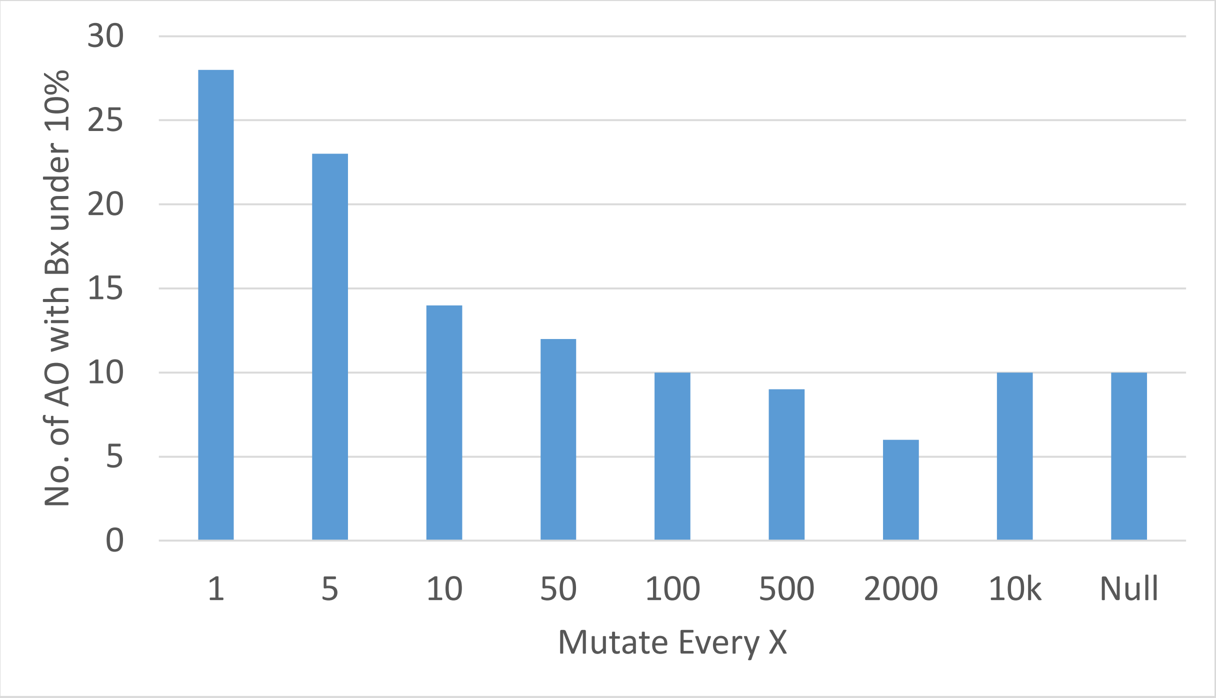


Figure 30 Note.

Figure 31

Impact of Additional Training Generations (Experiment Three, Modification Five)

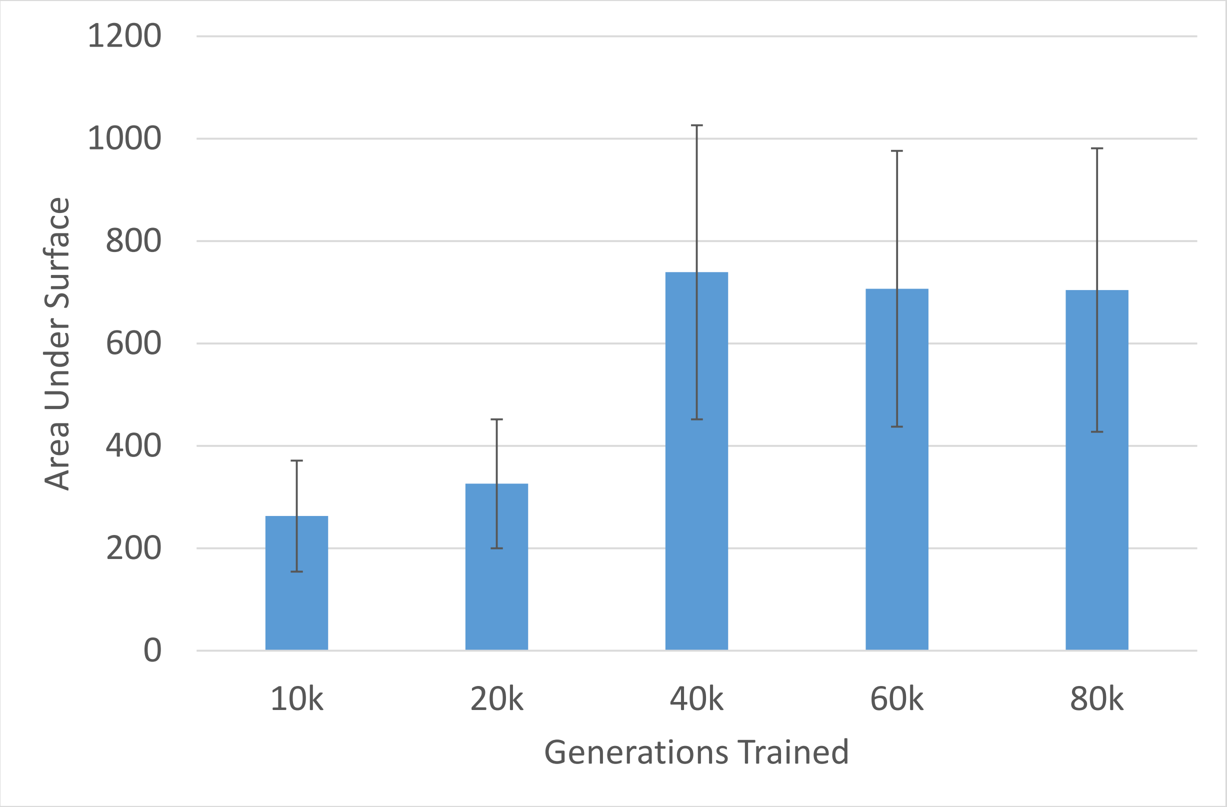


Figure 31 Note. Error bars are the sum of the 90% confidence intervals for each condition within the training set. The area under the surface is calculated by summing together the peak behavior of each condition within the training set.

Figure 32

Stimulus generalization gradients, Condition: Yes Wall, Yes Background

(Experiment Three, Modification Six)

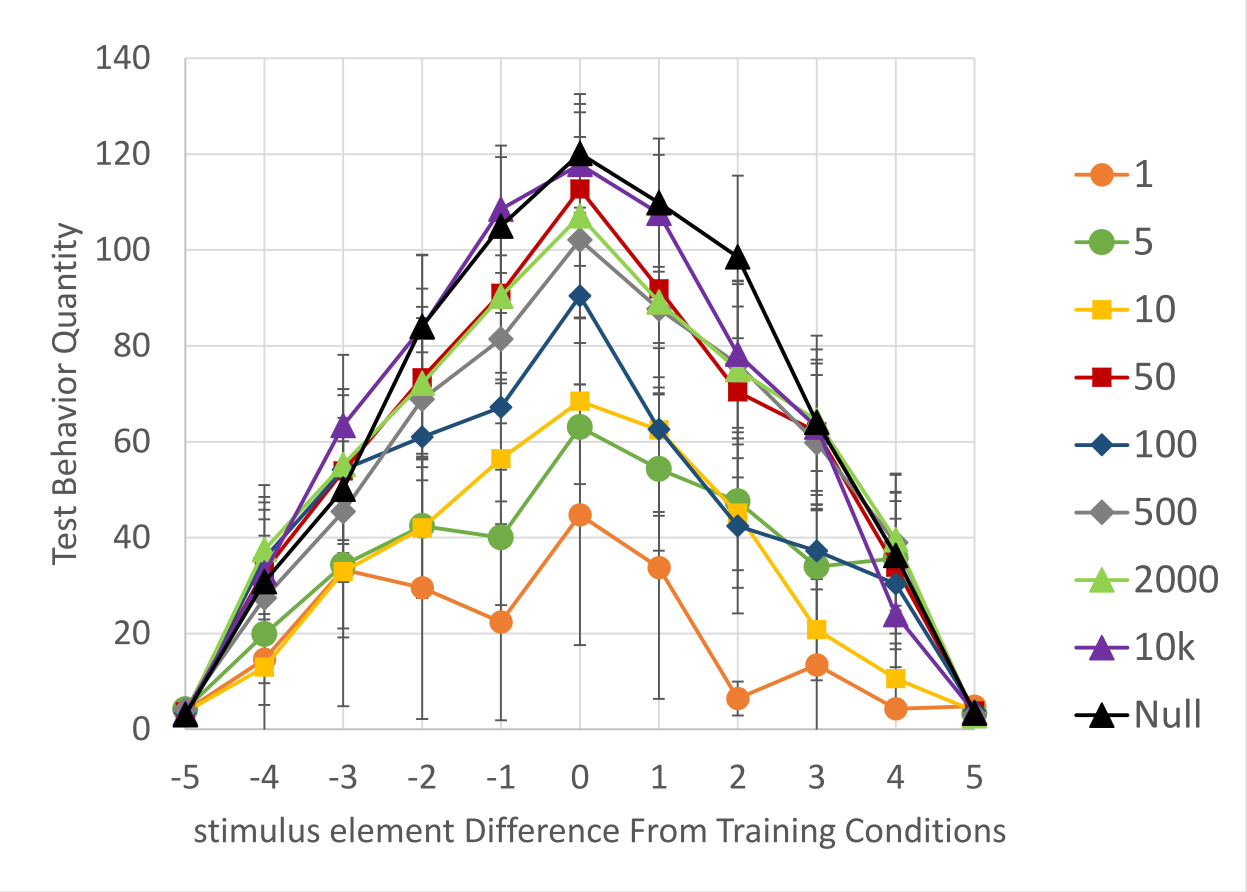


Figure 33

Stimulus generalization gradients, Condition: Yes Wall, No Background

(Experiment Three, Modification Six)

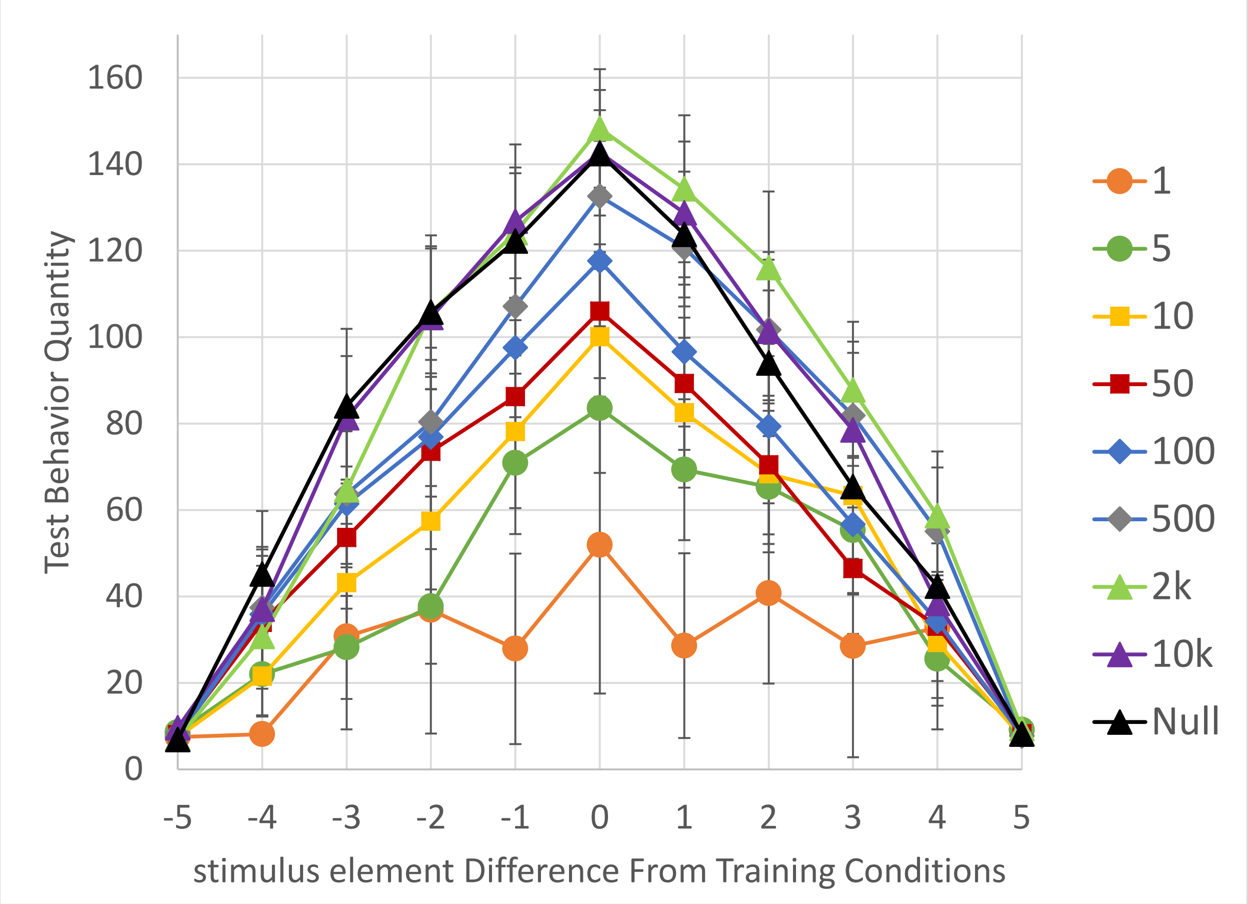


Figure 34

Stimulus generalization gradients, Condition: No Wall, Yes Background

(Experiment Three, Modification Six)

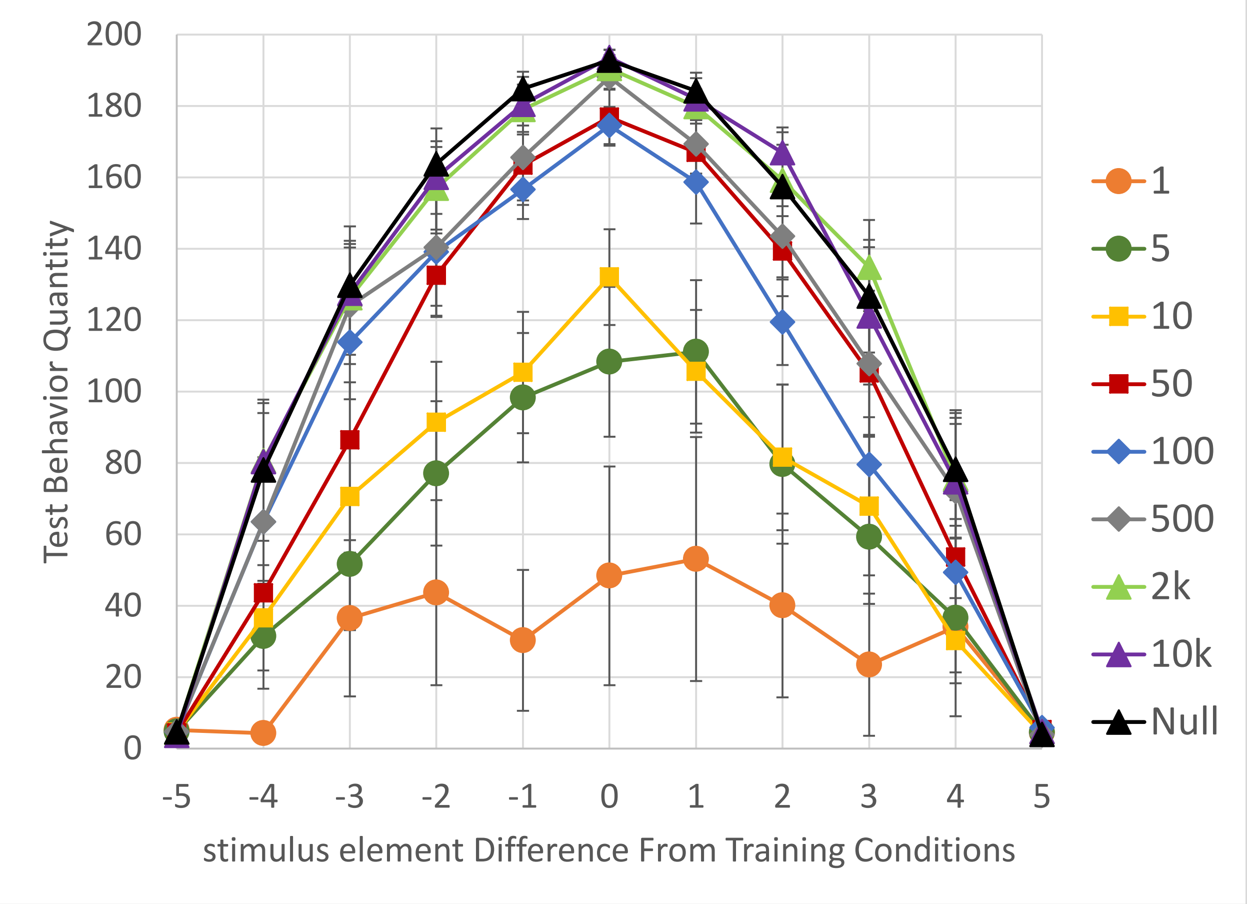


Figure 35

Stimulus generalization gradients, Condition: No Wall, No Background

(Experiment Three, Modification Six)

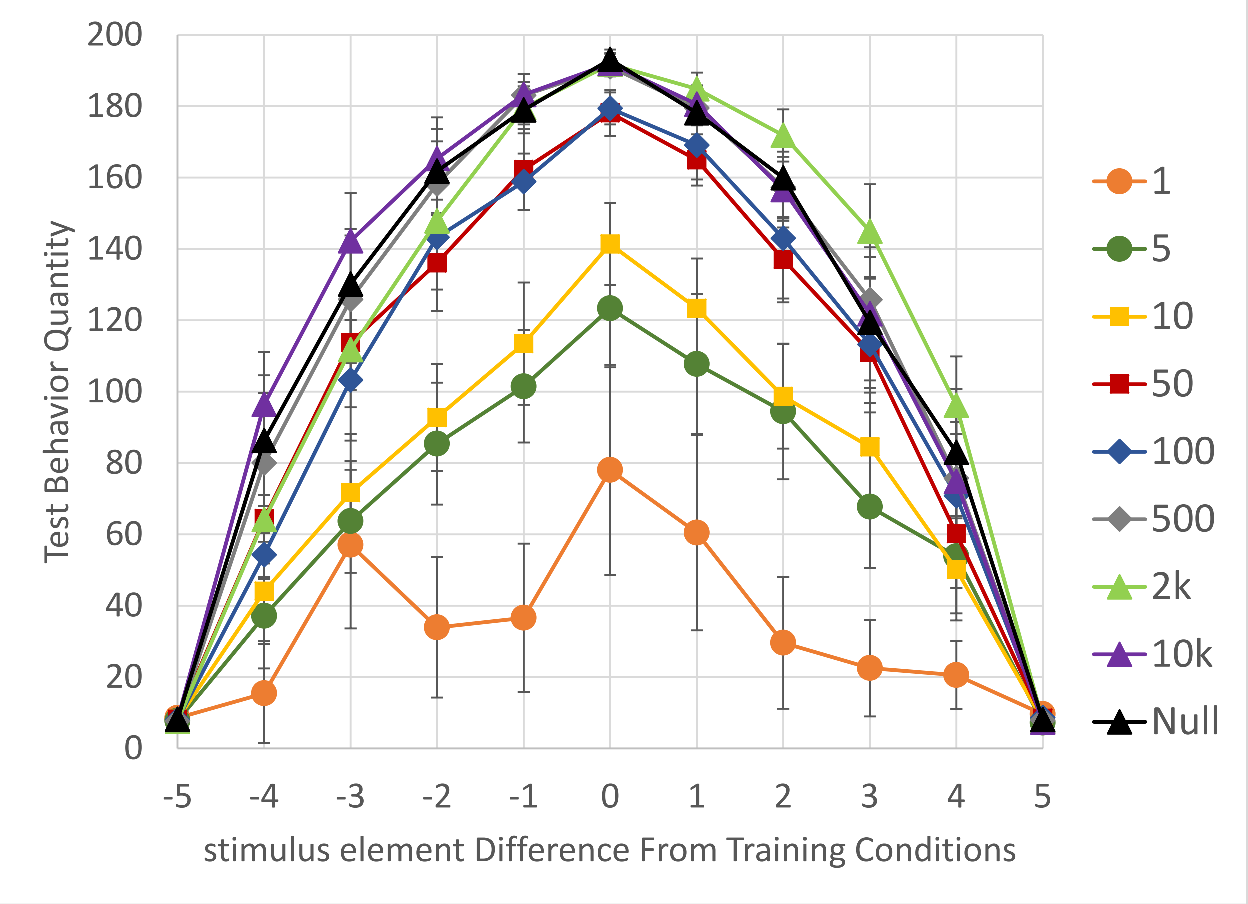


Figure 36

Adjusted Confidence Intervals (Experiment Three, Modification Six)

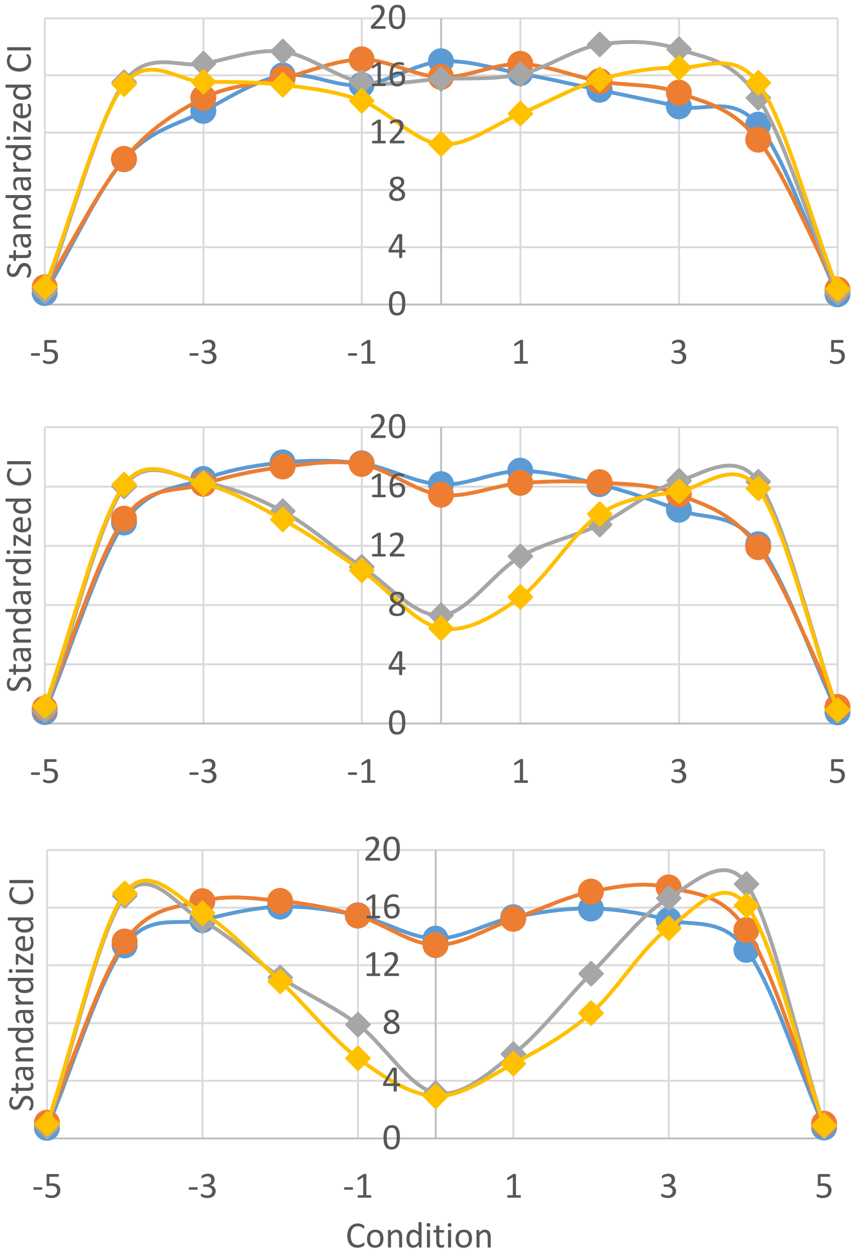


Figure 36 Note. The four groups represented here are the wall and background group (blue circle), the wall only group (orange circle), the background only group (grey diamond), and the no wall and no background group (yellow diamond).

Figure 37

Gaussian and Exponential function fits

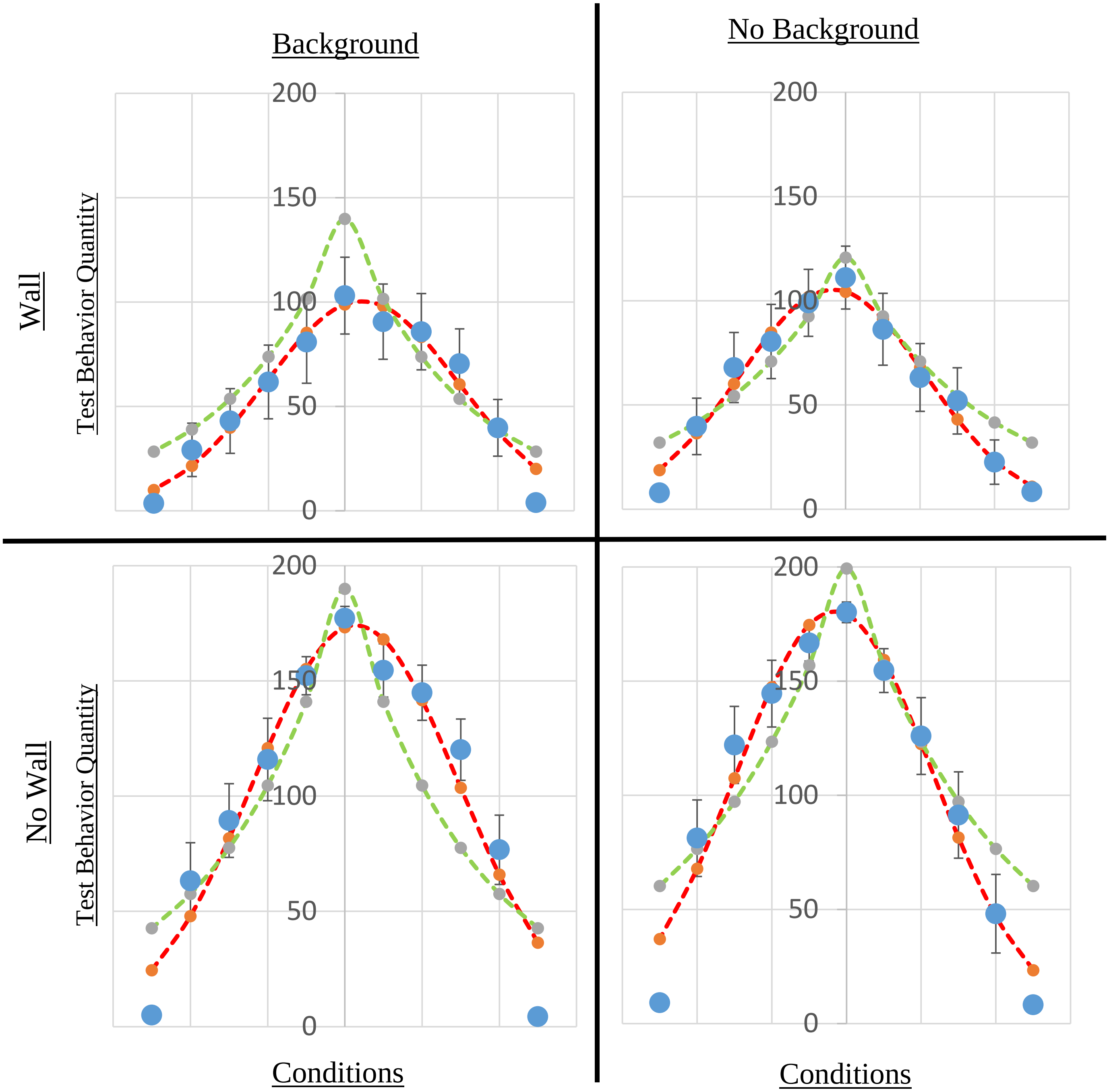


Figure 37 Note. An exponential function (green dashed line) and a Gaussian function (red dashed line) were both fitted to the data (blue)

Figure 38

Background Reinforcement Impact on Concurrent Schedule Sensitivity

(Experiment 3 Behavior on Concurrent Schedules Review)

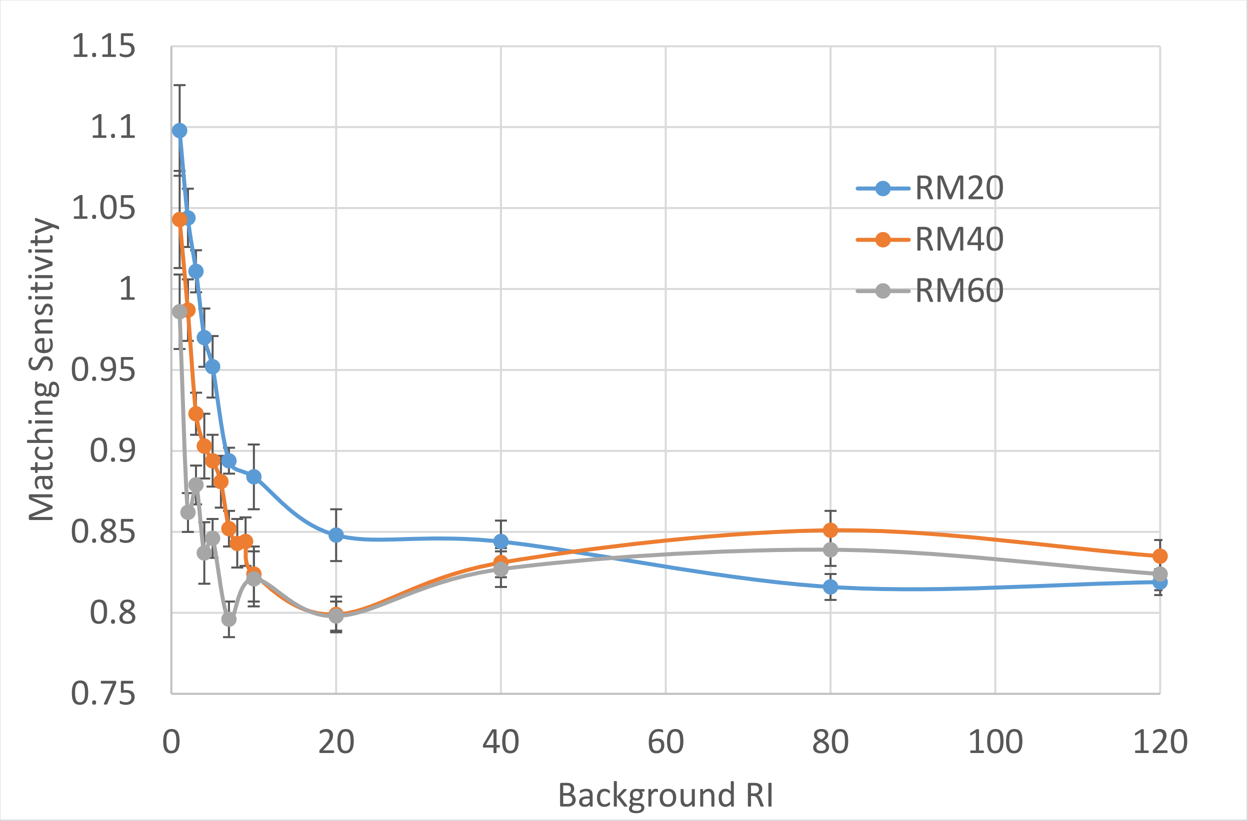


Figure 39

Impact of Behavior Population Quantity on Concurrent Schedule Sensitivity

(Experiment 3 Behavior on Concurrent Schedules Review)

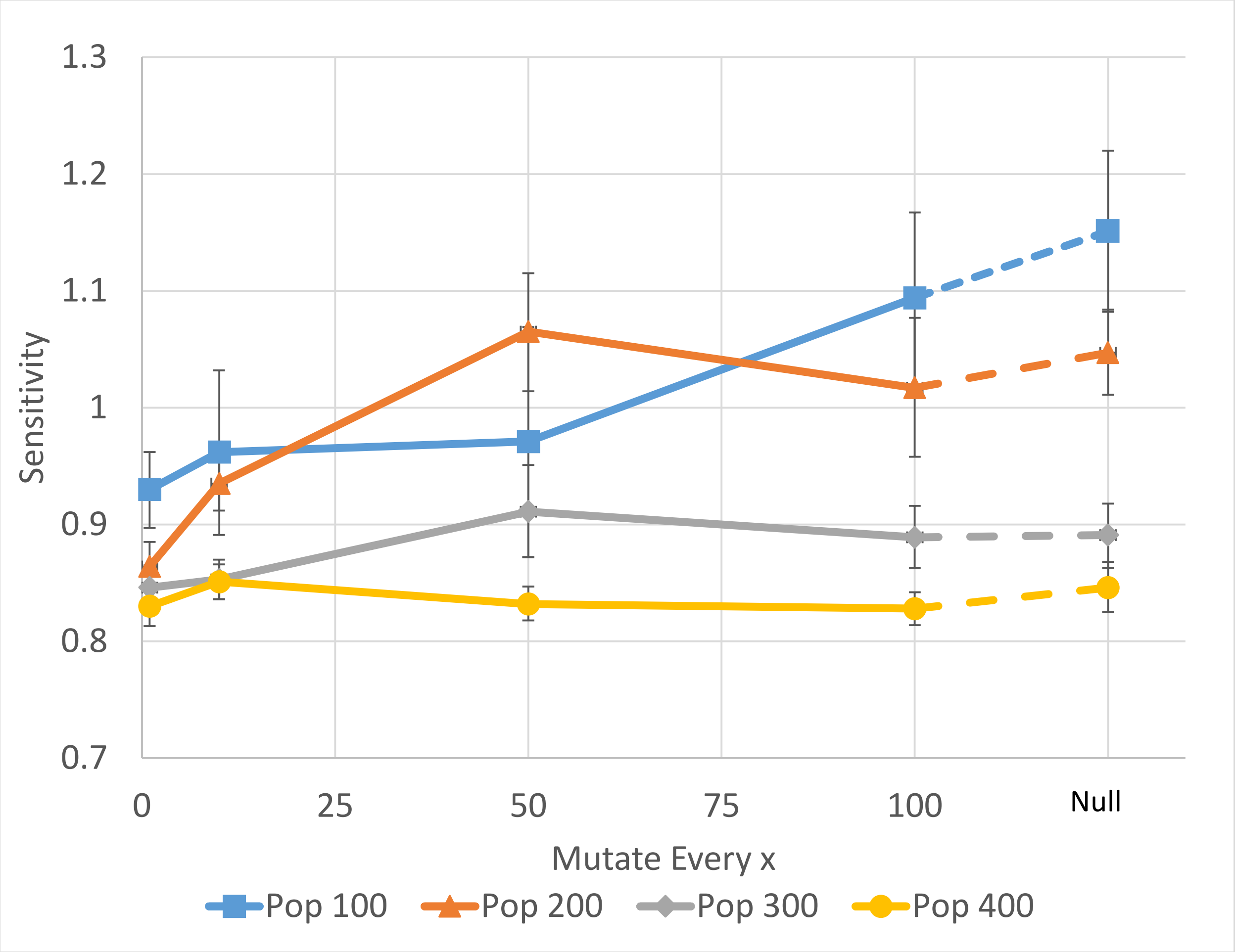


Figure 40

Impact of Selection Modifier Curvature on Concurrent Schedule Sensitivity

(Experiment 3 Behavior on Concurrent Schedules Review)

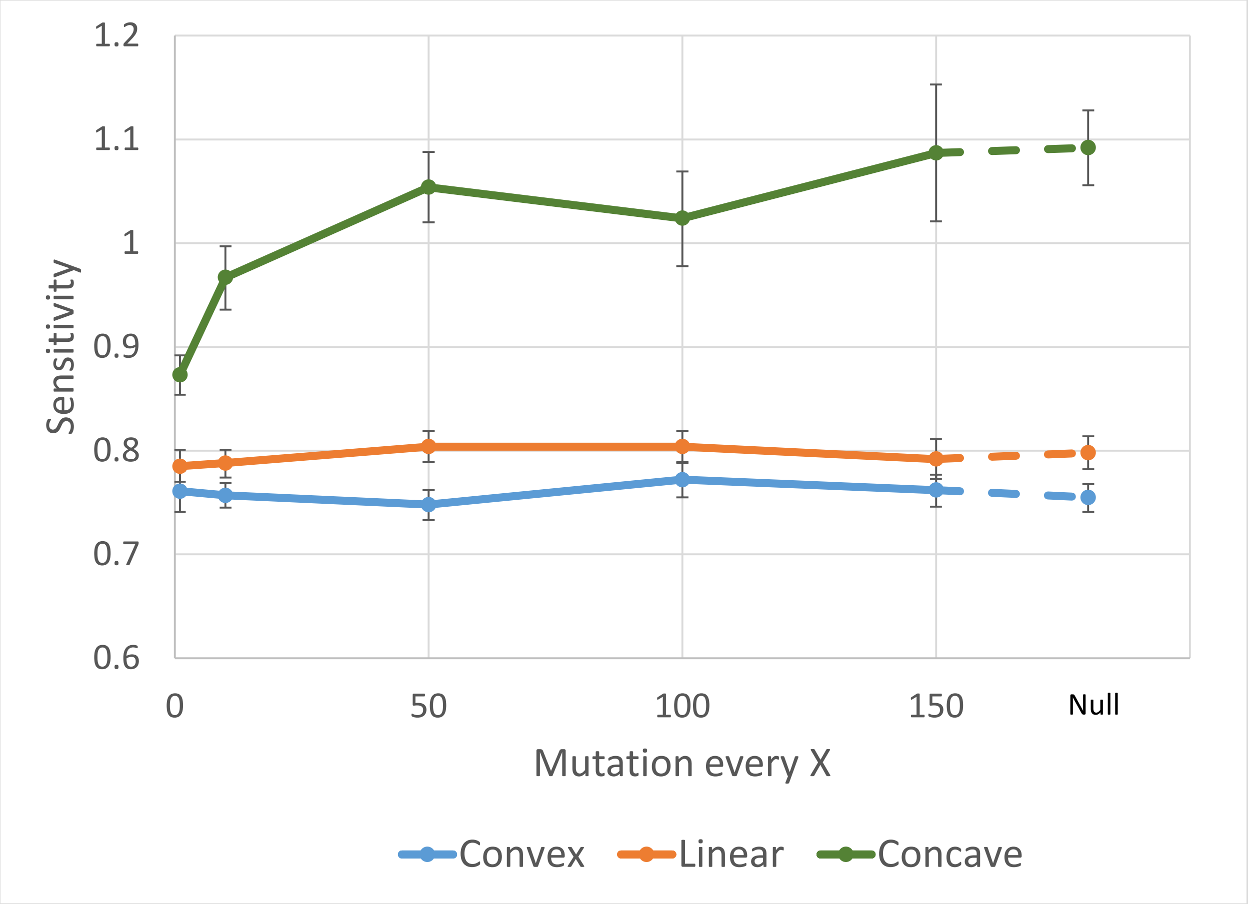


Figure 41

AOs with under 10% behaviors on target during testing

Wall and Background Condition (Experiment Three, Modification Six)

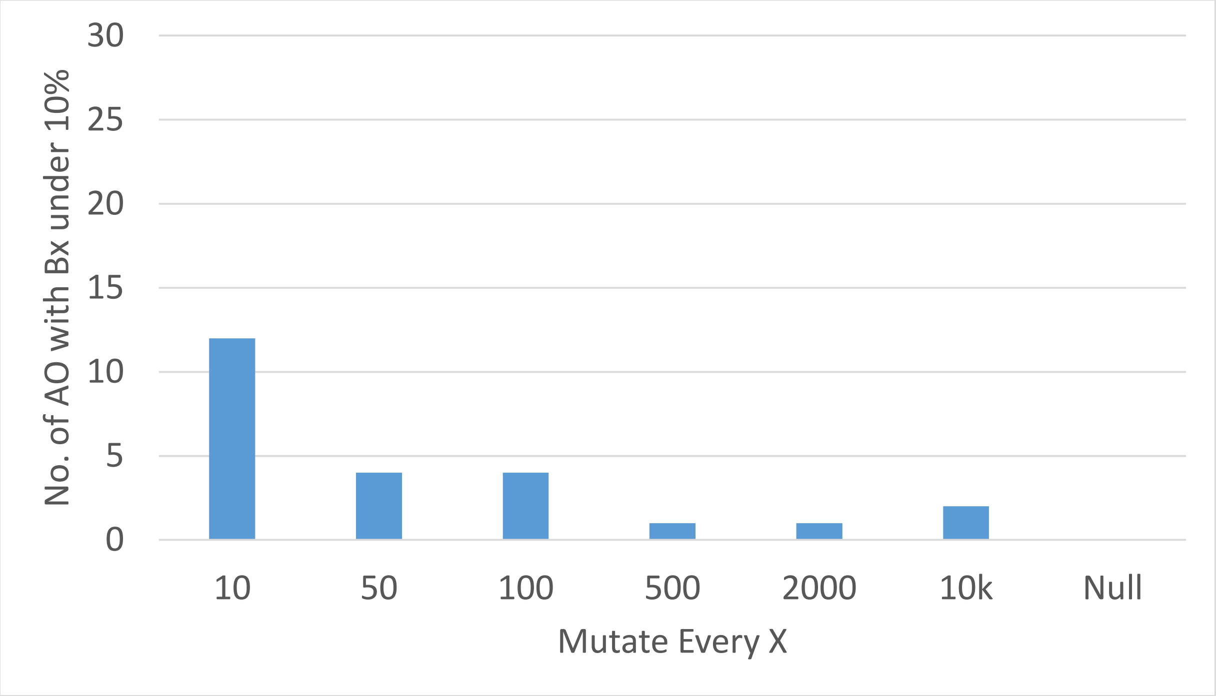


Figure 41 Note. Population mutation rates 1 and 5 were excluded due to not reaching 30 AOs that passed criteria after 90 AOs were run.

Appendix A

Implementation of Stimulus Generalization Environment

The stimulus generalization experiment was, at first, designed to determine if the modified ETBD combined with SST could create animal-like stimulus generalization gradients without any additional modification. Unfortunately this was not found to be the case and the additional modifications to the modified ETBD in order to enable stimulus generalization are recorded in modifications five and six. This section will focus on the development of the stimulus environment.

In rearrangement style (e.g. wavelengths of light, sound frequency) stimulus generalization experiments with live organisms, creating similar stimuli to the conditioned stimulus is nearly self-evident and generally intuitive. For example, in terms of visual wavelengths, wavelengths that are only incrementally different from the conditioned stimulus are easily obtained through the use of a prism. Through rotation of the prism, different wavelengths can be obtained, and the stimuli similarity/difference from the conditioned stimulus is easily calculated by obtaining the difference in wavelength.

In order to convert this physical measure into the abstract stimulus space the AO will inhabit, we must first consider how these signals are converted into neural impulses in live organisms. For example, once light enters the eye, different cone cell subtypes will activate, depending on which wavelengths of light are present. Each of the cone cell subtypes have a range of wavelengths that they respond to, and these wavelengths overlap. This overlap is not just present at the sensory level, but also at the processing level (Ito & Komatsu, 2004). There have been studies that have found neuronal activity that correspond to stimulus similarity for faces (Onat & Büchel, 2015) in the inferotemporal cortex and the width of generalization (Kahnt et al., 2012) in the striatum.

The idea of overlapped sensory and processing neuron activation can be extrapolated to the stimulus element space. Any singular physical stimulus, like a blue light, is assumed to be processed into multiple stimulus elements. The difference between two similar physical stimuli can be symbolically represented by changing the number of overlapping stimulus elements. For the majority of the current study, five stimulus elements were assumed to be linked to one stimulus type, although different numbers of linked stimulus elements were tested in modification four.

Experiment Series Three Methods

Phase One. This experiment was designed to examine the ability of AOs animated by the modified ETBD to exhibit stimulus generalization gradients when exposed to overlapping stimulus combinations. The experimental protocol was based on the procedure in Guttman and Kalish (1956). AOs were run on two training schedules and 11 testing schedules. The training schedules were 20k generations each, while the testing schedules were 200 generations each.

During the first training schedule, the AO was exposed to five wall stimulus elements and five training-wall stimulus elements. There was no reinforcement on the target class, but there was background reinforcement with a RI 10 and a reinforcement magnitude of 40 on 200 background reinforcement targets.

During the second training schedule, AOs were exposed to the five training-wall stimulus elements, and five other stimulus elements representing the discriminative stimulus (Trained Stimuli). For the target class, the reinforcement rate was RI 10 with a reinforcement magnitude of 5. There was no background reinforcement present during this schedule. During testing, the AOs were exposed to 11 different conditions, defined by the specific combination of Trained Stimuli and untrained stimuli. These conditions were presented in a random order. The non-wall stimuli in this experiment can be thought of as discrete steps along a spectrum like wavelengths of light, or a number line. If the stimulus elements were numbers on a number line from 1 to 15, the Trained Stimuli would be in the middle, numbers 6 through 10. To represent slightly different stimulus environments, the AOs were presented with different combinations of Trained Stimuli and untrained stimuli. For example, one of the conditions contains four Trained Stimuli, and one untrained stimuli. Using the number line analogy, this would be stimulus elements numbered 5 through 9, or stimulus elements numbered 7 through 11. With only one stimulus element that is untrained, this is considered the “+1” or “-1” condition. This system continues to the “+5” condition in the positive direction or the “-5” condition in the negative direction, which contains none of the Trained Stimuli and five untrained stimulus elements. All testing conditions have five wall stimulus elements alongside the Trained Stimuli or untrained stimuli, for a total of 10 stimulus element in the local environment. See Figure 26 for a visual representation of the non-wall stimuli per condition. Modifications one, two, and three were implemented in experiment three, using an entropy-based observation 2%, a behavioral population of 200, and a 200 phenotype background target class with RI 10 and a RM 40.

Preliminary testing showed that none of the testing conditions show greater amounts of behavior than the baseline operant-level rates. This finding, while unexpected, does not deviate from our understanding of organisms that are animated by the modified ETBD. AOs have been shown in past experiments, as well as the second experiment in this series, to adapt extremely quickly to changing environmental conditions. The testing conditions do not have reinforcement based on target class behavior. Combined with the presence of background reinforcement, the AOs can be expected to quickly shift their behavior away from the target classes. In order to deal with this challenge, behavioral populations need to be modified to be more perseverative. However, behavioral populations must still be able to maintain the ability to adapt or else the AO would no longer be able to respond to changes in its environment after having learned a particular response.

Modification Five. One simple method of making AOs more perseverative is by reducing the amount of parents generated during the selection step when entropy is low and keeping behaviors from the original population instead. Reducing the number of parents generated also reduces the rate of mutation, since mutation is only completed on new child behaviors. Linking the number of parents to the level of entropy allows the organism to respond quickly when the behavioral population is relatively evenly distributed over the phenotype range (when entropy is high) but slows down the rate of recombination and mutation as the behavioral population is biased by reinforcement. Initial testing of the concept used the following concave power function to convert entropy to selection rate,

$$S = ax^{b}, \tag{8}$$

Where S is the percentage of parents selected (a value between 0 and 1), x is the entropy of the behavioral population, with a and b as constants (a = 0.0625, b = 2). When $ax^{b} > 1$, S was set to 1. This equation and the corresponding constants were chosen due to multiple factors. First, it was important to have an equation that covered the appropriate entropy range. The average starting entropy was found to be 4.555 with a maximum of 4.630 for 100k randomly generated populations with 200 behaviors each. S in equation 8 is equal to one when the entropy is four, giving approximately 12% of the total range where the rate of selection is at its maximum. Second, this equation would allow the AO to retain the ability to adapt quickly when there is little to no prior reinforcement influence on the behavioral population. Finally, the equation and constants were chosen to generate a concave function. The function form (Shown in Figure 27) causes the selection percentage to drop quickly at high entropies and to hover close to zero at low entropies. This theoretically slows down rate of behavior dispersion in behavior populations that are heavily influenced by reinforcement, giving more time for another reinforcer to appear and reinforce the bias.

Another important aspect that needs to be considered is how to allow AOs to be responsive to changes in the reinforcement environment even after the behavioral population has been shaped by a prior reinforcement paradigm. One of the key aspects of the ETBD that allows an AO to adapt quickly to changes in reinforcement patterns in the environment is mutation. However, at low entropies, mutation does not happen due to the low number of child behaviors. In order to compensate for this issue, a minimum population mutation rate was created. When the number of children is below the threshold needed for mutation to occur, a counter will note the number of generations that have passed without mutation occurring. When the number of generations met a chosen threshold, a single behavior from the population is chosen and mutated using the bit-flip-by-individual method. Then the counter restarts from zero. The minimum population mutation method was designed to allow for fine tuning of the rate of mutation. If the minimum rate of population mutation is too high, then the AO might not be able to “remember” the reinforcement on the target class. However, if the minimum population mutation rate is too low, then the AO might not be able to adapt to changes in reinforcement within a reasonable time period.

In order to test the concave, entropy-based selection modifier and the minimum population mutation rate, the procedure described in phase one was run with the two additional modifications. The minimum mutation rate was set to trigger after the following numbers of generations: 1, 5, 10, 15, 20, 30, 40, 50, 75, 100, 500, 1k, 2k, 5k, and 10k. An additional condition (null) tested the concave, entropy-based selection modifier without the minimum population mutation rate.

The next experiment set in phase one was designed to test the impact of additional training on stimulus generalization. AOs were run on the same schedules as the rest of experiment three. The target class was reinforced using the following random interval rates: 20, 40, 60, 80, 100, and 120. The AOs were run with the following reinforcement magnitudes: 10, 20, 30, 40, and 50. The stimulus environment contained a total of 10 stimulus elements in each schedule, with five wall stimulus elements and five target stimulus elements, which vary based on the schedule as shown in Table 3. During training, five ‘training-wall’ stimulus elements were used, which were different from the five ‘wall’ stimulus elements used in testing. The following training lengths (in generations) were tested: 10k, 20k, 40k, and 80k. Modifications one, two, three, and five were implemented using an entropy-based observation 2%, a behavioral population of 200, a 200 phenotype background target class with RI 10 and a RM 40, and a concave selection modifier function with a minimum population mutation rate of 100.

Modification Six. Experiments during modification five showed high levels of variation between individual organisms during the testing schedules. Under the same conditions, some AOs showed high amounts of behavior during testing, others showed some responding during some test conditions and not others, while other AOs showed baseline levels of behavior across all testing conditions. While variation is to be expected even in live organisms (See Figure 2, right plot), it is rare for a live organism to have baseline levels of responding, at least during the first round of testing. In order to better characterize the source of the variation in AOs, three potential causes were examined. First, most live organism stimulus generalizing trainings include a stringent training performance acceptance criterion. In order to create comparable criteria for the AOs, the last 200 generations of the training period were examined. If 80% of all of an AO's behaviors were within the target class, it was considered to have passed the criteria. Second, the presence of wall stimulus elements might also be a source of variability. In order to maintain the same number of stimulus elements while removing wall stimulus elements, all non-wall stimulus elements were doubled. Finally, background reinforcement might have also been interfering with learning stimulus generalization gradients.

In this experiment, AOs were tested with an 80% training cut off and with combinations of the previously mentioned conditions (wall stimulus element / no wall stimulus element, and background reinforcement / no background reinforcement). During training, five unique training-wall stimulus elements were used instead of the five wall stimulus elements used in testing. Modifications one, two, and three were implemented using an entropy-based observation 2%, a behavioral population of 200, and a 200 phenotype background target class with RI 10 and a RM 40 (if applicable). Modification five was also implemented using the concave power function described earlier and the following minimum population mutation rates: 10, 15, 20, 30, 40, 50, 75, and 100. AOs were run in sets of 30, until enough AOs were generated that met the training condition to meet a sample size of 30. If there were more than 30 AOs that qualified, only the first 30 AOs were used in the final analysis. AOs were run with minimum population mutation rates lower than 10, but there were not enough AOs that met the training criteria after 90 AOs so testing was discontinued due to monetary and time limitations. The data from the under-sampled groups were excluded from the group analysis in order to allow for fairer comparison.

Concurrent Schedule Review. Multiple experiments were completed to determine if the entropy-based selection modifier and the minimum population mutation rate impacted the ability of the modified ETBD to demonstrate animal-like behavior on concurrent schedules. First, the selection modifier without the population mutation minimum was tested with different strengths of background reinforcement. A 200 phenotype background was used, with the following ratio intervals: 1, 2, 3, 4, 5, 10, 20, 40, 80, and 120; Three reinforcement magnitudes were used for each RI: 20, 40, and 60. Modifications one and two were implemented using entropy-based observation 2% and a behavioral population of 200. Second, the section modifier was tested with multiple minimum population mutation rates. The following minimum mutation rates were used: none, 1, 10, 50, 100, 150, and 500. Modifications one, two, and three were implemented using an entropy-based observation 2%, a behavioral population of 200, and a 200 phenotype background target class with RI 10 and a RM 40. Third, the section modifier was tested with multiple minimum mutation rates across different populations. The following minimum mutation rates were used: none, 1, 10, 50, 100, 150, and 500. The following behavioral populations were tested: 100, 200, 300, and 400. Modifications one and three were implemented using an entropy-based observation 2% and a 200 phenotype background target class with RI 10 and a RM 40. Finally, the curvature of the selection modifier was varied between concave, linear, and convex by adjusting the constants in Equation 8. For the concave function, a = 0.0625 and b = 2. For the linear function, a = 0.25 and b = 1. For the convex function, a = 0.5 and b = 0.5. The following minimum population mutation rates were used: none, 1, 10, 50, 100, 150, and 500. Modifications one, two, and three were implemented using an entropy-based observation 2%, a behavioral population of 200, and a 200 phenotype background target class with RI 10 and a RM 40.

Data Analysis for Experiment 3

The absolute rates of behavior in all conditions were examined. A repeated measures ANOVA and two contrasts were conducted on the final results. Based on the Tukey Bonferroni Sidak method (Keppel & Wickens, 2004, p. 119), a family wise alpha of 0.05 is appropriate for 10 degrees of freedom with three comparisons. Both Gaussian and exponential functions were fitted to modification six results, for the stimulus generation forms generated with a population mutation rate of 100. Both function types were fitted using the Excel solver add-on. The exponential function, due to its form, were solved for each half of the stimulus generalization gradient separately. The fitted parameters for each half were reflected across the midline of the stimulus generalization gradient to generate two different exponential fits. The r-squared was calculated for each of the two potential exponential function fits and function with the highest r-squared was used as the best exponential fit.

Experiment Series Three Results

Phase One, Modification Five. The results of first test of the selection and mutation modifiers can be seen in Figure 28. Each condition shown is average behavior of 30 AOs across the different stimulus element combinations, with ‘0’ denoting the training set. Each positive or negative step away from zero point removes one more of the training set stimulus elements and includes an additional non-training set stimulus element. As seen in the figure, when all training stimulus element are present the AO has the highest behavioral response during testing and this decreases as the number of training stimulus elements decreases. The total behavior generated during the testing phase shows that high mutation rates tend to have lower amount of total behavior, as seen in Figure 29. However, once the mutation falls below one mutation every 50 generations, it is difficult to determine if the amount of behavior continues to increase, due to the overlapped confidence intervals. The high degree of overlap in the confidence intervals is likely indicative of a high degree of variation between AOs. Analysis of individual AO behavior found that, even when general population mutation rate is extremely low, five to ten AOs in each group have less than 10% of their behavior contained within the target class, as shown in Figure 30. Different amounts of training were examined as well, and as shown in Figure 31. In this figure, the 'area under the surface' was calculated by summing together the peak behavior of each stimulus generalization gradient in the matrix of reinforcement rates and reinforcement magnitudes that had the same number of training generations. It was found that there is a transition between 20k training generations and 40k and above training generations. This difference is likely to be significant since the confidence intervals have only a small overlap. For this experiment, the 90% confidence intervals were also quite large in comparison to the data collected, with the confidence interval being between approximately 75%-85% of the size of the data collected.

Modification Six. This study was completed to examine the variation within the data collected previously. The stimulus generalization gradient figures for the four conditions –Figure 32: Wall and Background, Figure 33: Wall and No Background, Figure 34: No Wall and Background, and Figure 35: No Wall and No Background– all show that AOs have the highest behavioral response during testing and this decreases as the number of training stimulus elements decreases like the previous experiment. The maximum behaviors for all population mutation rates in the 'wall and background' condition are higher than the results from modification five, indicating that 80% training cut off increases rates of behavior. In the two 'no wall' conditions, when there is less population mutation, have a more ‘rounded’ and less triangular appearance. This appearance was less prominent in AOs with more rapid population mutation rates.

The variation was examined by looking at the standardized confidence interval. The standardized CI is the sum of the confidence intervals divided by the number of conditions. In Figure 36, the top plot shows the standardized confidence intervals for PMR 5, 10, 15, and 20. The middle plot shows the standardized confidence intervals for PMR 30, 40, 50, 75, and 100. The bottom plot shows the standardized confidence intervals for PMR 500, 1k, 2k, 5k, 10k and Null. The standardized CIs for the four conditions: 1) both wall and background (blue circle), 2) wall and no background (orange circle), 3) background with no wall (grey diamond), and 4) no wall and no background (yellow diamond) are plotted separately. Results from PMR 1 were excluded from the plots due to having substantially higher CIs in most conditions which would likely bias the data, and not increase its explanatory power. Based on the results, removing background reinforcement had a minimal impact on the standardized confidence intervals in all conditions. The absence of walls led to substantially lower CIs when the population mutation rate was slow, but only for conditions with three or more trained stimulus elements present. As the population mutation rate increases, the CIs in the more trained conditions increases as well. At the highest population mutation rates, there is minimal to no difference between the standardized CIs for the wall groups and the non-wall groups. When only one trained stimulus element were present, the walled conditions seemed to have marginally lower standardized CIs than the non-wall groups. In all other conditions, the non-walled groups had either equal or lower CIs.

A repeated measures ANOVA was performed on all combinations of groups and population mutation rates and a statistically significant difference ($p < 0.05$) was found between the behavior of at least two groups (See Tables 11 and 12) for all combinations. The effect size, calculated as omega squared (ω^2), showed a large effect size for all combinations. A planned contrast between behavior when all five trained stimulus elements were present and the two conditions where no trained stimulus elements were present was also performed and a statistically significant difference was found between these groups (See Tables 13 and 14, $p < 0.05$). The effect size for all groups was below 0.2, indicating a small effect.

In each of the four groups, the Gaussian function was found to fit the data better than the Exponential function. In the group with both wall stimulus elements and background reinforcement, the Gaussian function was found to have an r squared of 0.95 compared to the Exponential function's r squared of 0.80. In the group with wall stimulus elements and no background reinforcement, the Gaussian function was found to have an r squared of 0.97 compared to the Exponential function's r squared of 0.88. In the group with background reinforcement present and no wall stimulus elements, the Gaussian function was found to have an r squared of 0.93 compared to the Exponential function's r squared of 0.80. Finally, in the group with no wall stimulus elements and no background reinforcement, the Gaussian function was found to have an r squared of 0.96 compared to the Exponential function's r squared of 0.84. See Figure 37 for a visual representation of the function fits.

Concurrent Schedule Review. Three sets of experiments were completed to determine if AOs with the modified selection function and the modified population mutation rate still perform as expected on concurrent schedules. The impact of background reinforcement on matching sensitivity was investigated and the results are shown in Figure 38. Similar to the results shown in Figure 15, after the background RI is slower than an RI 20, the matching sensitivity is relatively stable, hovering between 0.8 and 0.85, for the reinforcement magnitudes investigated. Reinforcement rates faster than RI 20, show higher than expected matching sensitivities, and even overmatching at the highest reinforcement rates. The behavior population size was also found to have an impact on matching sensitivity. As shown in Figure 39, the matching sensitivity is relatively clustered at the fastest population mutation rates, with 200, 300, and 400 behavior populations clustering around approximately 0.85 sensitivity, while the 100 behavior population has a matching sensitivity of 0.93. As the population mutation rate slows down, the differences between the populations becomes more pronounced, with the 100 and 200 behavior populations showing high amount of overmatching and the 300 behavior population showing slightly elevated matching sensitivity stabilizing around 0.89 and the 400 behavior population showing relatively flat matching sensitivity, starting at 0.83 for high levels of population mutation and ending at 0.84 for no population mutation. Finally, the curvature of the selection modification function was investigated and the results are shown in Figure 40. The concave function used in this study shows increasing matching sensitivity, starting at 0.87 for high population mutation levels and ending at 1.09 for no population mutation. The linear and convex conditions showed stable mostly flat matching sensitivity, hovering around 0.76 and 0.78 for convex and linear respectively.

Experiment Three Discussion

AOs animated by the modified ETBD with the concave selection function and the population mutation modifier are able to generate monotonically decreasing stimulus generalization curves, with the peak of the behavior at the discriminative stimulus. The slope varied based on the height of the peak. Importantly, the form of the stimulus generalization gradients were found to be Gaussian nature, like those of live organisms (Ghirlanda & Enquist, 2003). However, the AOs, when run on concurrent schedules had consistent overmatching

matching sensitivity rates, unlike live organisms. Taken together, these results point to the modified ETBD to likely to have the capacity to show animal-like stimulus generalization gradients but a different solution for durability of learning needs to be found.

The parameters for durability and flexibility of learning (i.e. the parameters for the selection modifier and the population mutation rate) had specific functional ranges. The selection modifier function form was only able to generate stimulus generalization gradients using a concave form of the three forms tested. Further exploration of the raw data revealed that the concave function form was the only form to collapse the behavioral populations into one behavioral phenotype. In the modified ETBD, this would prevent recombination from having any impact since both parents would have the same genotype. The normal mutation rate would be likely be zero as well since the concave function form also minimizes the number of child behaviors at low levels of entropy. While could potentially be considered “switching off” the ETBD portion of the function, this is not necessarily a bad result. Living organisms, when presented with a laboratory operandi, do not typically interact with it in multiple ways. They often will just interact it in one way, at least until they are put on an extinction schedule causing an extinction burst (Pierce & Cheney, 2013, p. 122). This is also the end result of a SST trial. During SST trials, the stimulus elements are trained to just one behavioral response, unlike the distribution of behaviors found with the ETBD. Once the number of generations reaches the population mutation limit multiple times, the population will slowly diverge from the single phenotype and learning can again occur. While this was able to replicate the stimulus generalization gradients seen in live organisms, the concave function led to overmatching on concurrent schedules. Both the linear and the convex forms of the function performed well on concurrent schedules leading to animal-like behavior. However, they did not cause a convergence of behaviors within the behavioral population and were thus unlikely to be able to have durable stimulus generalization gradients.

A similar phenomena was found when examining the impact of behavior population size. Populations with 100 and 200 behaviors, when used in conjunction with a concave selection modifier function, would lead to overmatching, but populations with 300 or 400 behaviors would not. This is likely due to the fact that larger behavioral populations are less likely to converge entirely on behavior, and thus require less mutation to move back into a more dispersed behavior pattern. Taken together, the results of the experiments on the selection modifier function form, the population size, and the background reinforcement rates all suggest an underlying robustness in the modified ETBD. While only a specific combination of selection modifier function form and small population size can generate the stimulus generalization gradients, all other forms and population sizes still lead to animal-like behavior on concurrent schedules.

The variability between AOs was a potential concern that was explored in modification six. The variability in the data was found to hamper analysis. For example, the training time data (Figure 31) showed a trend towards greater generalization with more training time, but the large confidence intervals prevented a more precise analysis. One possible cause of this variability was AOs having difficulty learning the discriminative stimulus. As seen in Figure 30, the rate of poor learners (i.e. AOs that only responded less than 10% of the time within the target class during testing) was 20% or higher, even at the slowest population mutation rates. After adding a learning criteria, ‘80% of behavior must be within the target class for the last 200 generations of training,’ the rate of poor learners was substantially reduced, as seen in Figure 41. After the criteria was implemented, all groups with population mutation rate of 50 or higher had less than five poor learners each. The peak height increased by approximately 20% after the change. The higher criteria and the lower number of poor learners is more in line with what is typically seen in animal studies of stimulus generalization (Hanson, 1959). However, still having up to five poor learners despite having a training performance criteria was an unexpected result. This suggests that the rate of unlearning was extremely rapid, and potentially occurred in the last 40 generations. This may be a weakness of this instantiation of learning durability.

Despite the general concern over variability, the presence of variability is not necessarily a problem in and of itself. In most studies of generalization, there is generally a high degree of variability even within the first round of generalization testing. The degree of variability is even higher in the later rounds of testing (as seen in Figure 2). Data in stimulus generalization studies is often averaged across participants prior to analysis, or at least organized into subgroups (Zaman et al., 2022). In the present experimental series, the learning criterion lead to stable confidence intervals (as seen in Figure 36) when walls were present, and still maintained the monotonically decreasing Gaussian form expected of stimulus generalization gradients. When walls were not present, the confidence intervals for the conditions near the discriminative stimulus decreased as the speed of population mutation decreased. The stimulus generalization gradients, while still Gaussian in form, seemed to have more rounded peaks. Similarly rounded peaks can be found in conditions where discrimination is poorer (as seen in Figure 3) or potentially when motivating operators, like starvation, are in effect (as seen in Figure 4). This suggests a link between these conditions and either reduced variability on a neurological level, or a predisposition to ignore neutral stimuli like walls under certain conditions.