

Does Variability Training Facilitate Learning in Two Genetically Divergent Mouse Strains?

by

Megan A. Arnold

A thesis to be submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 7th, 2016

Keywords: Variability, Strain, ASD, Mouse, IRA, Learning

Copyright 2016 by Megan A. Arnold

Approved by

M. Christopher Newland, Chair, Alumni Professor of Psychology
Jeff Katz, Alumni Professor of Psychology
John Rapp, Professor of Psychology

Abstract

Variability is an operant dimension of behavior so the degree of variability observed depends upon the reinforcement contingency. One approach to reinforcing variations is the threshold procedure, in which a particular response or response pattern, out of several possible ones, produces reinforcement if it occurs infrequently relative to other responses. Behavioral variability is also associated with clinically relevant populations. Individuals with Autism Spectrum Disorders (ASD) tend to behave in a highly repetitive manner compared to controls, whereas individuals with Attention Deficit Hyperactivity Disorder (ADHD) tend to behave in a highly variable manner. The purpose of the first experiment was to assess whether the BALB/c mouse strain, a proposed murine model of behavioral deficits associated with ASD, also models invariant responding characteristic of ASD compared to control, C57Bl/6 mice. Behavioral variability is imperative to the acquisition of novel responses, i.e., learning. Concurrent reinforcement of operant variations facilitates the acquisition of a difficult response. The purpose of the second experiment was to determine whether a history of variable responding during a critical period of development, adolescence, affects learning in adulthood using an Incremental Repeated Acquisition (IRA) task with a Performance and Learning condition. On the IRA task, performance sessions approximate behavioral repetitions because the target response does not change from session to session whereas learning sessions require organisms to acquire a new target response each session.

Acknowledgements

The author would like to thank Dr. Christopher Newland for his patience and guidance throughout this project without which, the author would likely still be programming the procedure. She would also like to thank her committee members Drs. Jeffery Katz and John Rapp for their comments on this project. In addition, support for this research was provided by: an Auburn University Cellular and Molecular Biosciences Peaks of Research Fellowship, the Society for the Advancement of Behavior Analysis, and the Auburn University Graduate Student Research Grant for thesis.

Table of Contents

Abstract	ii
Acknowledgments	iii
List of Tables	v
List of Figures	vi
List of Abbreviations	vii
Chapter 1: Introduction	1
Evidence that Variability is an Operant Dimension of Behavior	1
Other sources of Variability	10
Extinction-induced Variability	10
Differences in Variability Between Groups	12
Variability and Learning	14
Animal Model of Autism Spectrum Disorder	17
Incremental Repeated Acquisition: A Measure of Learning	20
References	24
Chapter 2: Does Experience During Adolescence Affect Adulthood? Variability Training in Adolescent Mice and Learning in Two Strains.	30
Abstract	31
Introduction	32
Method	34
Data Analysis Strategy	40
Results.	41
Discussion	46
References	53

Tables	58
Table Captions	59
Figures	60
Figure Captions	65

List of Tables

Table 1.....	58
--------------	----

List of Figures

Figure 160
Figure 2	61
Figure 3	62
Figure 463
Figure 564

List of Abbreviations

5-HT	Serotonin
ADHD	Attention Deficit Hyperactivity Disorder
ASD	Autism Spectrum Disorders
CRF	Continuous Reinforcement
DA	Dopamine
DRA	Differential Reinforcement of Alternative behavior
EIBI	Early Intensive Behavioral Intervention
EXT	Extinction
FI	Fixed Interval
FR	Fixed Ratio
IRA	Incremental Repeated Acquisition
IRI	Interresponse interval
IRT	Interresponse time
ITI	Intertrial Interval
MCL	Maximum Chain Length
MD	Major Depression
MPR	Mathematical Principles of Reinforcement
OVVA	Operant Variability and Voluntary Action theory
PND	Postnatal Day
PQ	Progress Quotient
RA	Repeated Acquisition
RF	Relative Frequency

SHR Spontaneously Hypertensive Rat

TMT Trimethyltin

VI Variable Interval

WKY Wistar-Kyoto

WRF Weighted Relative Frequency

Chapter 1: Introduction

Several disorders are associated with diminished response variability, including Autism Spectrum Disorders (ASD). Behavioral characteristics of ASD include impaired social interactions, anxiety, intellectual disability (often), and stereotypic behavioral patterns. The BALB/c strain has been a putative animal model of the social deficits and neurobiological abnormalities observed in ASD. To date there is no research aimed at determining whether the natural variability observed in this strain is consistent with ASD or if this strain responds to reinforced variations similar to ASD (Brodkin, 2007). Characterizing the behavioral phenotype of this strain would prove informative for future ASD research, particularly if responding in this strain is invariant.

Evidence that Variability is an Operant Dimension of Behavior

There is considerable evidence that variability is an operant dimension of behavior. The degree of observed variability is sensitive to contingency, schedule of reinforcement, and extinction (Neuringer, 2009). Often, the term “variable” connotes high variability, but variability of responding exists along a continuum that ranges from repetitive to highly variable. If a response class shows low or no variability, then one can predict the next response to occur in the set. In contrast, if a response class is highly variable, then one cannot readily predict which response within the class will occur next. It is important to note that variability is not a characteristic of a single response; instead it is a continuous dimension of a response class.

Given that the goal of behavioral science is to observe, predict, and control behavior, variable responding is traditionally considered an undesirable outcome indicative of insufficient experimental control (Kazdin, 2011). In this context, it makes sense that researchers have usually sought to avoid variability of dependent measures, as it is seemingly counterproductive to demonstrating causal relationships between environment and behavior. Furthermore, highly

variable responding also seems to be incompatible with operant conditioning. An operant is a response class in which the probability of members of the class occurring is altered by contingent consequences (Skinner, 1953). Reinforcement describes the process in which a consequence is contingent upon a behavior that results in an increase in the probability of that behavior occurring in the future in a similar context. Although reinforcement can, and typically does, reduce response variability over the past 28 years, Neuringer and colleagues have demonstrated that this need not be the case (e.g., Grunow & Neuringer, 2002; Page & Neuringer, 1985). Highly variable responding *can* be observed, predicted, and controlled. Variability is a dimension of behavior that is sensitive to selection by consequences, similar to topography, frequency, and force.

In an early study of operant variability, Blough (1966) reinforced least frequent interresponse times (IRT) of three pigeons. The reinforcement criterion was based on the 150 most recent IRTs, excluding those less than 0.8s. The variability of IRT durations increased as a function of contingency and was indistinguishable from a random number generator. Thus, the IRT distribution is sensitive to reinforcement and, when required, the patterning of IRTs is random-like. Reinforcement may increase the variability of response patterning in time, but reinforcement can also increase the variability in response patterning across manipulanda. Bryant and Church (1974) demonstrated that organisms could be trained to alternate responses on left (L) and right (R) levers in an unpredictable pattern consistent with a stochastic model. Rats were assigned to one of three groups that differed in the probability of reinforced alternations and stays (100:0, 75:25, and 50:50). An alternation entailed responding on the lever opposite that pressed on the previous trial, whereas a stay entailed pressing the same lever pressed on the previous trial. The probability of alternating lever presses from the previous trial

varied as a function of the reinforcement contingency. Alternations were most probable for the 100:0 group, intermediately probable for the 75:25 group, and least probable for the 50:50 group. The predictability of alternations and stays for the 75:25 group did not statistically differ from random responding. Based on reinforcement criterion the number of alternations and stays was predictable, but whether an alternation or a stay would be performed on any given trial was unpredictable.

Blough (1966) and Bryant and Church (1974) demonstrated that reinforcement can increase the variability of how an already established operant is distributed in time and of switching between two members of the same response class. Pryor, Haag, and O'Reilly (1969) demonstrated that the effect of reinforcement is not limited to increasing the variability of already established responses, but can also expand the size of the response class by reinforcing novel behaviors in two rough-toothed porpoises. Reinforcement of a novel behavior resulted in complex previously unobserved responses in this species. Although previously performed behaviors (e.g., porpoising or tail-slapping) were no longer reinforced, they continued to occur as the porpoises' repertoire expanded to include novel responses.

Another study demonstrating that the degree of variability observed is controlled by the degree of variability required is Grunow and Neuringer (2002, Experiment 1). In this study, rats performed 3-response sequences across two levers and one key. Reinforcement was delivered according to a threshold criterion. In a threshold procedure, only sequences performed relatively infrequently, relative to other sequences resulted in reinforcement. To determine if a sequence met reinforcement criterion, the relative frequency (RF) of all 27 sequences were calculated at the end of each trial. Each time a sequence occurred, the frequency (count) for that sequence incremented by one. The RF was calculated by dividing the count of a sequence by the total

number of sequences performed. Following each reinforcer delivery the weighted relative frequency (WRF) was calculated by multiplying each RF by a weighting coefficient, in this case 0.98. Applying a weighting coefficient weighted recently performed sequences more heavily than sequences not recently performed, decreasing the probability that the same sequence would have produced reinforcement on temporally proximal trials (Denney & Neuringer, 1998). If a sequence did not meet criterion, the WRF was equal to or greater than the threshold, the weighting coefficient was not applied.

Rats were assigned to one of four groups that differed only in threshold criterion. The most stringent threshold criterion was 0.037, intermediate threshold criteria were 0.055 and 0.074, and the most permissive threshold criterion was 0.37. Initially, reinforcers were delivered according to a fixed ratio (FR) 1 schedule: each time the performed sequence's WRF was below threshold, reinforcement occurred. Following FR 1, reinforcement frequency decreased to a variable interval (VI) 1min followed by a VI 5min schedule to determine whether contingencies or reinforcement frequency better accounted for response variability, or U-values. U-values are a measure of entropy or uncertainty. In Equation 1, p is the probability (relative frequency) of a sequence, i and n , or 3^N , is the number of possible sequences.

$$U = \frac{-\sum_{i=1}^{27} (p_i \times \log_2 p_i)}{\log_2 n} \quad (1)$$

U-values can range from 0.00, in which only one of the possible sequences occurred, to 1.00, in which all possible sequences occur with equal frequency. Under the FR 1 condition, the U-values were consistent with the reinforcement criterion: the highest variability was seen for the 0.037 group, second highest variability was observed in the 0.055 group, followed by the 0.074 group, and then the 0.37 group. Interestingly, when the reinforcement schedule was thinned to a VI 1min and VI 5min, the effect of reinforcement frequency depended upon baseline U-values.

For the relatively invariant 0.37 group, variability increased with decreased reinforcer frequency. In contrast, the most stringent (0.037) and intermediate (0.055 and 0.074) groups showed a decrease in variability as reinforcement frequency decreased. Thus, reinforcement increased variability for all groups, but the effect of reducing reinforcement rate on variability depended upon the level of variability observed under FR 1 conditions, reducing reinforcer rate did not increase variability for all groups. These findings suggest that the reinforcement contingency increased U-values more consistently than reinforcement frequency.

Page and Neuringer (1985) were perhaps the first to clearly demonstrate that variability is an operant dimension of behavior by controlling for other potential determinants of sequence variability, such as intermittent reinforcement. In their fifth experiment, reinforcers were delivered according to a Lag n schedule, that is, a sequence resulted in reinforcement if it differed from the previous n sequences. For example, in a Lag 2 schedule of reinforcement, a sequence results in reinforcement if it differed from the previous two sequences. In the first phase of this experiment, reinforcement was delivered according to a Lag 50 schedule: only sequences that differed from the previous 50 sequences resulted in reinforcement. Following the reinforcement contingent condition was the yoked-variable ratio condition during which grain presentation was yoked to the reinforced trials during the last six sessions of the Lag 50 condition. Thus, if the first and fourth trial in the Lag 50 condition resulted in reinforcement, then the first and fourth trial in the yoked condition resulted in reinforcement independent of which sequence the pigeon performed. When reinforcement was contingent upon variation, performed sequences were variable. In the yoked condition, response variability was low even though the number and order of trials that resulted in reinforcement was the same as the Lag 50 condition.

In their third experiment, they demonstrated that the degree of observed variability increased as a function of the lag requirement. As the reinforcement contingency (i.e., the Lag) changed and required a higher degree of response variability to meet criterion, observed reinforcement variability increased. In a Lag n contingency, a sequence that satisfies the reinforcement contingency must differ from the previous n sequences. Pigeons were trained to perform an 8-response sequence across two keys. When the reinforcement criterion for sequence variability was relatively low (Lag 5), the variability of performed sequences tended to vary with the lag requirement and reinforcement increased accuracy to 85 per cent or more during the last five sessions. Increasing the reinforcement criterion to a Lag 10 resulted in a greater increase in response variability, but pigeons were able to maintain a high degree of accuracy. Although, the degree of variability observed for the Lag 5 schedule was sufficient to satisfy the Lag 5 criterion it was insufficient to satisfy the Lag 10 criterion. The degree of variability only increased and met the Lag 10 criterion when the reinforcement contingency changed to Lag 10. Once the Lag increased, pigeons again earned reinforcers on 85% or more of trials during the last five sessions of Lag 10. When the Lag became too large (Lag 50) U-values and accuracy decreased.

In experiment three increasing the Lag resulted in an increase in response variability, experiment four was designed to determine the mechanism that best accounts for this increase in this variability, specifically if responding was random or under the control of previously executed sequences, as might be suggested if a memorial mechanism were in place. If responding is under the control of previously executed sequences, then increasing sequence length would reduce the percentage of reinforced trials, however, if pigeons responded in a quasirandom fashion, then increasing sequence length would result in an increase in the percentage of reinforced trials. In their fourth experiment, the number of responses required per

trial (4, 6, and 8 responses per sequences) varied and the reinforcement criterion was set to a Lag 5 requirement. As the response requirement per sequence increased, the percentage of reinforced trials significantly increased. This finding supported the quasirandom hypothesis, further supporting the conclusion that variability is a reinforcing dimension of behavior and that pigeons were not performing sequences based on remembering previous sequences.

In order to extend the findings from Page and Neuringer (1985, Experiment 4) that responding is random rather than under the control of previously executed sequences, Neuringer (1991) manipulated interresponse intervals (IRI) within a sequence by varying the minimum time between responses. If variable responding was under the control of previously performed sequences, then as the IRI increased, accuracy, or number of trials that resulted in reinforcement, should decrease. Long-Evans rats performed 4-response sequences across two levers and sequences were reinforced according to a Lag 5 schedule. The first three responses in each sequence produced an IRI and responses made during the IRI were ineffective: responding during the IRI did not count toward the sequence and did not extend the IRI. Increasing IRI duration resulted in an overall increase in sequence variability. Within a sequence, response repetitions decreased as length of the IRI increased, but repetitions increased during responses relatively proximal to reinforcement, i.e. the third or fourth response in a sequence. In Neuringer (1991)'s second experiment, he compared the effect of varying IRI lengths for a VAR group to a REP group. In the VAR group, reinforcement occurred according to a Lag 5 schedule, whereas for the REP group, a single target sequence (LLRR) was reinforced according to an FR 1 schedule. For the VAR group, increasing the IRI length resulted in increased accuracy, i.e., rats were more likely to meet the reinforcement criterion. For the REP group, as IRI increased, variability increased (although not to the same extent as the VAR group) and errors were more

likely to occur. These findings are inconsistent with the memory hypothesis: memory of previous responding appears to facilitate repetitive responding while impairing variable responding.

Following reinforcement of invariant responding, extinction increases variability (described below in more detail), in contrast, extinction following reinforced variations can decrease response variability. In one such study, Da Silva Souza, Abreu-Rodrigues, and Baumann (2010) studied the effects of extinction and non-contingent reinforcement on sequence variability following variability training and repetition training. Participants in the variability groups earned points if they performed sequences that satisfied a Lag 2 and threshold criterion. If the weighted relative frequency of a sequence was less than 0.02 and the sequence differed from the previous two sequences then the participant earned a point. For the repetition group, participants earned points for performing a target sequence. Following training for the respective contingencies, both groups experienced an extinction condition or a non-contingent reinforcement condition followed by a return to baseline contingency. For the variability groups, extinction and non-contingent reinforcement decreased U-values and indicated that sequences performed during extinction were less variable than contingent variability training. For the repetition groups, U-values increased during extinction and non-contingent reinforcement and indicated that sequences were more variable than during repetition training.

Neuringer, Kornell, and Olufs (2001) investigated how extinction affected responding following reinforcement of variable sequences. Rats performed 3-response sequences across two levers and one response key. During the 15 baseline sessions, sequences produced reinforcement if the weighted relative frequency was less than the threshold criterion (0.05). Four extinction sessions followed baseline during which reinforcers were never delivered. Extinction reduced

response rate while increasing sequence variability. Interestingly, sequences that occurred with a low probability during baseline increased during extinction to a greater extent than higher probability sequences during baseline. In Experiment 3, rats were assigned to a VAR group or a REP group. Variable responding was reinforced in the VAR group whereas a specific target sequence was reinforced in the REP group. Following acquisition of the respective contingencies, both groups experienced extinction. During baseline, U-values were higher for the VAR group than for the REP group and extinction increased U-values for both groups to similar levels (0.869 and 0.739, respectively) although the authors did not report if extinction U-values significantly differed. Interestingly, the effect of extinction on reinforced variations differed between Da Silva Souza, Abreu-Rodrigues and Baumann (2010) and Neuringer, Kornell, and Olufs (2001, Experiment 3). These differences could be due to reinforcement contingency, species, or reinforcer type. Da Silva Souza, Abreu-Rodrigues and Baumann (2010) used secondary reinforcers that were delivered according to a Lag procedure with humans whereas Neuringer, Kornell, and Olufs (2001, Experiment 3) used primary reinforcers that were delivered according to a threshold procedure with rats. Future studies should systematically investigate how different procedures and reinforcer type affect extinction of reinforced variability across species to elucidate these inconsistencies.

Response variability is also sensitive to discriminative stimuli (Pesek-Cotton, Johnson, & Newland 2011). Using a multiple VARY:FR 4 schedule, rats responded more variably in the presence of the VARY SD and less variably in the presence of the FR 4 SD. During the VARY condition, U-values approached 0.90 whereas U-values were approximately 0.10 in the FR 4 condition. When the VARY and FR 4 component alternated according to a mixed schedule, differences between U-values decreased. Demonstrating that variability is susceptible to

stimulus control provides further evidence that variability is an operant dimension of behavior. An operant must be consistent with the three-term contingency in which a discriminative stimulus occasions a response that produces a reinforcer.

Together, these studies demonstrate that reinforcement need not result in a decreased variability, but making reinforcement contingent upon variable responding results in an overall increase in variability. The degree of variability can be precisely controlled by reinforcement contingency and discriminative stimuli, and affected by the removal of the contingency (i.e., extinction).

Other Sources of Variability

Operant conditioning is not the only means by which responding becomes variable. When baseline responding is relatively invariant removing the reinforcement contingency induces variable responding. Distinct populations can vary naturally (i.e., without explicit training) in baseline levels of response variability: degree of variability observed in responding when the reinforcement criterion does not require variations can differ between populations. Examples of these distinct populations are genetically divergent strains of rodent such as the spontaneously hypertensive rat (SHR) and Wistar-Kyoto rat (WKY), individuals diagnosed with Major Depressive Disorder (MDD) and non-clinical individuals, and individuals diagnosed with ASD and non-clinical individuals.

Extinction-induced variability. Variability can be the product of removing the reinforcement contingency, i.e., extinction, or genetics. During extinction, the most common response performed is the one that previously resulted in reinforcement and occasionally a response dissimilar to the previously reinforced response is performed resulting in an overall increase in variability (Neuringer, Kornell, & Olufs, 2001). Antonitis (1951) trained rats to nose-

poke along a slot in an operant chamber. Early in training, the variability of nose-poke location was high but as training continued the variability of nose-poke location decreased. During extinction sessions, the variability of nose-poke position increased. Reinforcement resulted in a decrease in response variability and response variability increased during extinction. This effect of extinction on response location variability was replicated with pigeons (Eckerman & Lanson, 1969). In their first experiment, any key peck to one of the 20 response keys produced access to grain and median response location was invariant during CRF. Following CRF, the reinforcement schedule changed to extinction (EXT) and variability in response key location increased. Finally, when CRF was reinstated, median response location variability again decreased to original CRF levels. In these instances, it is likely that reinforcement resulted in a decrease in response location variability because the contingency did not require variations.

Extinction-induced variability is necessary for the acquisition of complex operants by shaping or reinforcing successive approximations. Initially a successive approximation of the target operant will produce reinforcement. Once the organism reliably performs this approximation, it is placed on extinction. Under the new contingency only the new approximation produces reinforcement. Placing the first approximation on extinction induces response variations increasing the probability that the organism performs the new criterion response and this behavior contacts reinforcement.

Extinction-induced variability can be used to produce socially preferable alternatives, such as mands, to problem behaviors, such as whining in participants with ASD, (Grow, Kelley, Roane, & Shillingsburg, 2008). Following a functional analysis of the problem behavior and baseline, a differential reinforcement of alternative behavior (DRA) condition began. During DRA, the problem behavior was placed on EXT and reinforcement was contingent upon the

socially acceptable alternative behavior. After the problem behavior was placed on extinction rate and variability increased. For example, whining increased, but so did the socially acceptable alternative of signing or speaking the word “toy”. After several extinction and DRA sessions, frequency of the socially acceptable alternative increased and problem behavior decreased.

Differences in reinforced variability between groups. Differences in variability are observed not only between groups with different reinforcement histories, but also for genetically divergent groups or clinical versus non-clinical populations. Johansen, Killeen, and Sagvolden (2007) compared response variability of nose-poking target locations in spontaneously hypertensive rats (SHR) and Wistar-Kyoto rats (WKY) on FI schedules of reinforcement. The SHR strain is a putative animal model of ADHD, a disorder characterized by increased response variability in humans (Saldana & Neuringer, 1998). For the shortest delay of reinforcement, an FI 1', WKY rats responded most frequently on the target nose-poke and the next most frequent response was on the nose-poke just below the target with few responses to other nose-pokes. In contrast, SHR rats responded most frequently to the target nose-poke but non-target nose-pokes occurred across a larger number of nose-pokes than non-target sequences for WKY rats. So, non-target nose-poke locations were more variable for SHRs than WKYs.

ADHD is associated with increased variability but some disorders, including Major Depression (MD) and Autism Spectrum Disorders (ASD), are associated with diminished variability. Hopkinson & Neuringer (2003) compared the effect of reinforcing variations in participants who met criterion for depression and non-depressed participants. Using a computer-generated task, participants produced 5-link sequences across two keys. During baseline, participants earned reinforcers independent of response variations. Depressed participants responded less variably (i.e., had significantly lower U-values) than non-depressed participants.

Following baseline, reinforcement was contingent upon sequence variability using a threshold procedure. When reinforcement was contingent upon sequence variation, variability (i.e., U-values) significantly increased for both groups and between-groups differences were no longer significant. Thus, when reinforcement is not contingent upon varying, participants with depression differ from non-depressed participants but these between-groups differences disappear with reinforcement.

Autism spectrum disorders are characterized by deficits in social behaviors and perseverative motor movements (e.g., hand flapping) and/or repetitive speech, such as echolalia. Although the severity of these symptoms can vary from individual to individual, the symptoms must interfere with social and occupational functioning (American Psychiatric Association, 2013, 5th edition). Early intensive behavioral intervention (EIBI), developed by Lovaas (1987), has proved successful in improving severe behavioral deficits for children diagnosed with ASD. The Lovaas model consists of 35-40 hours per week of discrete trial training and other behaviorally-based interventions. The rationale for EIBI is that by providing intensive intervention in early childhood, symptomology may be improved such that the child requires and can benefit from less intensive and costly support (McEachin, Smith, & Lovaas, 1993). Low levels of behavioral variability (i.e., perseverative motor movements and verbal behavior) may interfere with adaptation to environmental changes or treatment efficacy of behavioral interventions.

Using a percentile schedule of reinforcement, Miller and Neuringer (2000) reinforced variable responding in adolescents diagnosed with autism, college-aged controls, and child controls. Participants generated sequences consisting of four responses across two keys (L and R). In the first phase of the experiment, reinforcers were delivered according to a probabilistic

(PROB) contingency. Independent of sequence variability, reinforcers were delivered with a 0.5 probability every trial. Following the PROB condition, reinforcers were delivered contingent upon sequence variability. For each of the 16 possible sequences the relative frequencies were calculated by dividing the total number of times a sequence occurs by the total number of sequences that occurred within a session. When a reinforcer is delivered, the relative frequency of each sequence was multiplied by a weighting coefficient (0.97), a tactic that decreases the contribution of the sequence's relative frequency exponentially as it becomes distal from current response. Whether or not a sequence met reinforcement criterion was determined based on a percentile schedule. The relative frequency for the just-performed sequence was compared with the 20 previously performed and ranked sequences. If the just-performed sequence ranked 11th or higher (i.e., 12th – 20th rank) then a reinforcer was delivered. In the first phase of the experiment, reinforcement was not contingent upon sequence variability, 50 per cent of randomly selected trials produced reinforcement.

During the PROB condition, when reinforcement was not contingent upon variation, U-values for participants with ASD were significantly lower than controls. When reinforcement was contingent upon sequence variability, the relative frequency of sequences decreased and U-values increased for all groups. Thus, when reinforcement is not contingent upon variations, these two groups (ASD and control) differ in the degree of response variability.

Variability and Learning

Considerable evidence exists that highly variable responding facilitates the acquisition of complex responses. Whether the source of variability is extinction, genetic, or operant, variability is the sine non qua of operant learning. Without some degree of variability, learning

new behavior, or learning to perform previously acquired behaviors in a novel context, is not possible.

Given that early intensive behavioral interventions can cost upwards of \$40,000 annually, interventions that improve treatment efficiency could substantially reduce the cost and increase accessibility to EIBI (Amendah, Grosse, Peacock, & Mandell, 2011). Although extinction induced-variability can facilitate the acquisition of mands in children and adolescents with ASD, extinction can also induce aggressive responses and increase the frequency and magnitude of self-injurious behavior (Azrin, Hutchinson, & Hake, 1966; Iwata, Dorsey, Slifer, Bauman, & Richman, 1994). As such, operant variability may be a less restrictive alternative to the facilitative effect of extinction on learning, in applied settings.

Variability can be considered the behavioral substrate upon which consequences select, this is particularly apparent in shaping complex operants such as verbal behavior. In order to establish an operant, an initial level of variable responding is required for consequences to select and thus narrow the degree of variability (Neuringer, 2004). This decreased variability of reinforcement can be detrimental to acquiring new operants (Schwartz, 1982b). According to the operant variability and voluntary action (OVVA) theory, this need not be the case as reinforcement does not necessarily result in a reduction in variability but the degree of variability is a function of the reinforcement contingency (Neuringer & Jensen, 2010). If variable responding is required for the acquisition of a new operant and reinforcing variations maintains highly variable responding, then reinforcing variable responding should facilitate the acquisition of difficult, yet similar operants.

Neuringer (1993) investigated how readily a target sequence is selected following variability training. During baseline, rats performed 4-response sequences across two levers and

reinforcers were delivered according to a Lag criterion. Following baseline, two easy, intermediate, and difficult sequences were selected for an *always* and *never* contingency. To compare whether variability training facilitated acquisition of an always-reinforced sequence one group of rats experienced variability training while the other did not. For the variability group, the reinforcement criterion was set to a Lag 3 and the reinforcers were delivered according to a VI 5s schedule. The control group did not receive variability training. For the variability and control rats, frequency of the *always* sequence increased as the experiment progressed, however the control rats did not perform the *always* sequence to the same extent as the variability rats. One common finding is that withholding reinforcement increases response variability when baseline responding is relatively invariant and this increased variability facilitates the acquisition of a new operant (i.e., shaping). Based on Neuringer's (1993) findings reinforced variability can also facilitate the acquisition of a target response.

Grunow and Neuringer (2002, Experiment 2) assessed whether variability training affected how readily rats acquired an easy (LRL) and a difficult (LKK) target sequence. Reinforcement was contingent upon performing sequences with a WRF below a given threshold (0.037, 0.055, 0.074, or 0.37) and target sequences were reinforced according to a VI 1' schedule of reinforcement. Once the interval timed out, the next sequence that met the variability criterion resulted in a single pellet. Sequences that met the variability criterion, but occurred before the interval expired, resulted in a time-out. Concurrently, rats performed a target sequence on a CRF schedule of reinforcement. Earning a reinforcer contingent upon performing the target sequence resulted in three pellet deliveries and reset the VI 1' timer.

When the target sequence was easy (LRL), rats in the high variability group (0.037) learned the target sequence the fastest but all groups (0.055, 0.074, and 0.37) eventually learned

the target sequence. High levels of variability increased the speed of acquisition for easy target sequences, but did not affect the eventual accuracy of performance. However, when the target sequence was difficult (LKK), group differences did not disappear with time. The groups with the most stringent variability reinforcement criterion (0.037 and 0.055) acquired the difficult sequence the fastest and performed the target sequence most frequently. The intermediate group (0.074) did not demonstrate the same level of target sequence acquisition, as did the high variability groups but this group performed an approximation of the target sequence (KKK) most frequently. Finally, the low variability group (0.37) did not acquire the difficult target response and tended to perform the previously reinforced easy target (LRL). Thus, concurrently reinforcing variable sequences facilitates the acquisition of a target sequence and this effect is most pronounced when the target sequence is difficult.

To date, research demonstrating the effect of variability on learning is limited to concurrently reinforcing variable sequences and a target sequence. Interestingly, Miller and Neuringer's (2000) study demonstrated maintenance of reinforced variability during a return to baseline, even though reinforcement was no longer contingent upon variation, which demonstrates promise for the maintenance of operant variability. To date, it remains relatively unknown if establishing a reinforcement history of operant variability continues to serve a facilitative effect on learning with other tasks. Demonstrating such a facilitative effect could improve the efficiency and accessibility of behavioral interventions for disorders such as autism spectrum disorder (ASD).

Animal Model of Autism Spectrum Disorder

One proposed animal model of social deficits characteristic of ASD is the BALB/c mouse strain. BALB/c mice, relative other mouse strains tend to show reduced sociability, barbering,

increased anxiety, and increased aggression (Brodkin, 2007; Kalueff, Minasyan, Keisala, Shah, & Tuohimaa, 2006). In tests of sociability, the BALB/c strain showed significantly more social avoidance than the C57Bl/6 strain (Brodkin, Hagemann, Nemetski, & Silver, 2004).

Furthermore, when group housed adult male BALB/c mice tend to behave more aggressively than other group-housed strains (Guillot & Chapouthier, 1996; Kalueff, Minasyan, Keisala, Shah, & Tuohimas, 2006). Although there are limited studies on stereotypies in BALB/c mice, one study reported more repetitive activity, including twirling and vertical jumping in the homecage, relative to C57Bl/6 mice (Tang, Orchard, & Sanford, 2002) and we have observed high rates of stereotypy in our lab.

In addition to behavioral similarities between BALB/c mice and individuals diagnosed with ASD, there is evidence of neurobiological similarities. For example, like children diagnosed with autism, BALB/c mice also show abnormally low levels of serotonin (Chandana, Behen, Juhasz, Muzik et al. 2005; Chugani, Muzik, Behen, Rothermel et al. 1999). This low level of overall serotonin may also be related to increased aggression and self-injurious behavior (Ferrari, Palanza, Parmigiani, de Almeida, & Miczek, 2005; McCracken, McGough, Shah, Cronin, Hong, Aman, et al. 2002; Popova, 2006; Posey & McDougle, 2002). In addition to abnormally regulated serotonergic system, BALB/c mice also tend to have larger brain to body ratio than other strains (Wahlsten, Hudspeth, & Bernhardt, 1975). Similarly, children diagnosed with autism also show enlarged brain weight, relative to body weight particularly during early childhood (Courhesne, Redcay, Morgan, & Kennedy, 2005; Hazlett, Poe, Gerig, Smith, Provenzale, Ross, et al. 2005; Nicolson & Szatmari, 2003). Taken together, these findings suggest that the BALB/c mouse strain is a viable candidate for an animal model of autism, at least in regard to sociability and neurobiological abnormalities of the serotonergic system and

brain volume and that the C57Bl/6 mouse is likely an appropriate comparison strain. The C57Bl/6 strain shows increased social approach, social sniffing, and direct social contact (e.g., social barbering) in addition to reduced aggression relative to the BALB/c strain (Brodkin, 2007).

Although there is ample evidence that the BALB/c strain is an animal model of social deficits associated with ASD, it remains unclear what other aspects of autism the BALB/c mouse strain does and does not model, specifically perseverative behavior often observed in individuals with ASD. No animal model is a perfect model of a disease or disorder and as such elucidating which characteristics of a disorder the animal models is informative. There exists tentative evidence from our lab that the BALB/c strain may perseverate more than the C57Bl/6 strain. In one study, BALB/c and C57Bl/6 mice responded under an FR schedule of reinforcement and the FR requirement for milk reinforcement increased across sessions. Saturation rate was higher for the BALB/c strain than the C57Bl/6 strain. Saturation refers to the rate at which a reinforcer's effect on responding diminishes as responses become temporally distal to the reinforcer. In an operant variability task, this should result in an increased probability of repeating recently performed sequences in BALB/c mice than C57Bl/c mice. As previously described in operant variability tasks, repeating a recently performed sequence precludes reinforcement (e.g., in a Lag procedure) or decreases the probability of reinforcement (e.g., in a threshold procedure due to the application of a weighting coefficient). It should be noted that differences in the response unit used in Hutsell and Newland (2013) and response units used in operant variability tasks (one response on one operandum versus several responses across multiple operandum) may limit the predictions based on BALB/c and C57Bl/6 saturation rates from one task to another. That being

said, one such study from our lab suggests that BALB/c mice perseverate more than C57Bl/6 mice when the response unit consists of several responses across three operanda.

In an incremental repeated acquisition (IRA) procedure (discussed below), BALB/c mice responded significantly more than C57Bl/6 mice during the acquisition of performance and learning chains with sucrose pellet reinforcement. Briefly, in order to increment the length of performance and learning chains, mice must repeat the chain to satisfy the reinforcement criterion and mastery-based criterion. Although accuracy for performance and learning chains increased across sessions for both BALB/c and C57Bl/6 mice, this increase was greater for the BALB/c strain (Johnson, Bailey, Johnson & Newland, 2010). As previously stated, repeating recently performed sequences is advantageous in an IRA task and if the BALB/c strain is more likely to repeat recently performed responses than the C57Bl/6 strain then may account for the differences in accuracy for these two strains.

If BALB/c mice perseverate more than C57Bl/6 mice, then this would be evidence that the BALB/c could also model perseveration observed in ASD and the C57Bl/6 strain could be an appropriate comparison strain of social deficits and perseveration observed in ASD. Further investigation of strain differences in perseveration is warranted and necessary to further characterize the behavioral phenotypes of these two strains. The proposed study will, hopefully, begin to answer this question by comparing BALB/c mice to C57Bl/6 mice on a task of operant variability.

Incremental Repeated Acquisition: A Measure of Learning

Although there are many methods available to study learning in non-human animals such as concurrent transitions, repeated acquisition, or performance on maze tasks, in our laboratory we commonly use an incremental repeated acquisition (IRA) task. Boren and Devine (1968)

developed a task to measure steady state relearning, within subject, of behavior chains called repeated acquisition (RA). In their study, three rhesus monkeys were placed in a chamber with 12 levers mounted in a line on one wall of the chamber with the levers grouped into four groups of three. During initial training pressing any lever produced a pellet. Once monkeys learned to press levers, the monkeys were then required to press any of the three levers from a group of levers say the left-most group of levers. The next step of training would require the monkeys to perform a two-link chain: for example press a lever from the right-most group, then from the left-most group to earn a food pellet. This training continued up to a four-link chain. Following training, the monkeys were required to press a specific lever from a group to earn a reinforcer, although the initial link was always from the left-most group and moved, in order, to the right. The measurement of interest in RA is the number of errors a monkey makes before acquiring or performing the target sequence and this target sequence changes from session to session. By learning a new target sequence each session, RA allows for the assessment of factors, such as schedule of reinforcement or drugs that influence transitional behavior can be assessed (Paule & McMillan, 1984; Thompson, 1974).

Similar to the RA task, an incremental repeated acquisition (IRA) task requires an animal to learn a target sequence within-session, but in IRA procedures the length of the sequence increments from a 1-link chain (i.e., single response) to an n -link chain, based on performance during the experimental session. Paule and McMillan (1984) developed an incremental repeated acquisition (IRA) task for rats to test the effects of a single injection of trimethyltin (TMT) on learning. In their IRA task, the terminal length of a response sequences was five links. At the beginning of each session 1-link chains produced a reinforcer, for example pressing the R lever produced a reinforcer. After 40 correct responses and a 1-min intercomponent T.O., the chain

length incremented to a 2-link chain such that only R-L chains produced reinforcers. This procedure for incrementing chain length, within session, continued until rats a 5-link chain, 60min elapsed, or 5min elapsed without a response. Although a 5-link chain was the longest possible chain in this experiment, during baseline rats reliably performed 4-link and 5-link chains but performance on these longer chains was more variable than shorter chains and TMT decreased performance. Unlike the traditional RA task, IRA allows for the measurement of two different types of errors. The first type of error is within sequence: an error occurs after the first correct response in a sequence. The second type of error is between sequence: an error occurs prior to the first correct response in a sequence. Furthermore, the pattern of responding that comprises an incorrect sequence provides further information about how an intervention disrupts learning (Cohn & Paule, 1995). For example, was the error due to perseverative lever presses or due to pressing the incorrect lever after a link is added to the chain?

In our laboratory, two components are used in IRA: the performance component and the learning component with each component presented during different sessions. The performance chain does not change from session to session and as such this component is useful for studying nonspecific drug effects on response rate or the effect of a neurotoxicant on psychomotor function (Bailey, Hutsell, & Newland, 2013; Bailey, Johnson, & Newland, 2010). Contrarily, learning chains change from session to session except the target chain remains constant within a session. Also, learning chains are not repeated until all possible learning chains are presented. Thus, each learning session is an opportunity to assess the effect an intervention (e.g., drug or neurotoxicant) has on acquiring a novel response (Bailey, Johnson, & Newland, 2010; Paule & McMillan, 1984). One measure of mastery-based IRA performance developed in our laboratory is progress quotient or PQ (Bailey, Johnson, & Newland, 2010). PQ is an overall measure of

learning that preferentially weights reinforcers earned in later links than reinforcers earned in earlier links, to account for differences in chain length and difficulty. In Equation 2, n denotes the length of a given link and Rf_n is the number of reinforcers earned in a given link and the summation of weighted reinforcers will be divided by the number of unweighted reinforcers earned within the session.

$$PQ = \frac{\sum(n * Rf_n)}{\sum Rf_n} \quad (2)$$

IRA is useful for characterizing the effect of training conditions on the behavior of genetically divergent mouse strains, such as the BALB/c and C57Bl/6 strains (Johnson, Bailey, Johnson, & Newland, 2010). Although IRA is useful for assessing the effect of chemical interventions and genetic differences on learning and steady-state performance, to date no studies have used IRA to assess the effect a previously implemented behavioral intervention has on learning and steady-state performance. Such studies addressing the effect of behavioral interventions have, later, on IRA could prove informative for applied research (e.g., early intensive behavioral intervention for ASD) as IRA is correlated with IQ in humans (Baldwin, Chelonis, Prunty, & Paule, 2012).

References

- Amendah, D., Grosse, S. D., Peacock, G., & Mandell, D. S. (2011). The economic cost of autism: A review. *Autism Spectrum Disorders. Oxford University Press Inc.*
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed). Washington, DC: Author: American Psychiatric Association.
- Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42*(4), 273-281.
- Azrin, N. H., Hutchinson, R. R., & Hake, D. F. (1966). Extinction-induced aggression. *Journal of the Experimental Analysis of Behavior, 9*(3), 191-204.
- Bailey, J. M., Hutsell, B. A., & Newland, M. C. (2013). Dietary nimodipine delays the onset of methylmercury neurotoxicity in mice. *Neurotoxicology, 37*, 108-117.
- Bailey, J. M., Johnson, J. E., & Newland, M. C. (2010). Mechanisms and performance measures in mastery-based incremental repeated acquisition: behavioral and pharmacological analyses. *Psychopharmacology, 209*(4), 331-341.
- Baldwin, R. L., Chelonis, J. J., Prunty, P. K., & Paule, M. G. (2012). The use of incremental repeated acquisition task to assess learning in children. *Behavioral Processes, 91*(1), 103-114.
- Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior, 9*(5), 581-591.
- Boren, J. J., & Devine, D. D. (1968). The repeated acquisition of behavioral chains. *Journal of the Experimental Analysis of Behavior, 11*(6), 651-660.

- Brodkin, E. S., Hagemann, A., Nametski, S. M., & Silver, L. M. (2004). Social approach-avoidance behavior of inbred mouse strains towards DBA/2 mice. *Brain Research, 1002*(1), 151-157.
- Brodkin, E. S. (2007). BALB/c mice: low sociability and other phenotypes that may be relevant to autism. *Behavioural brain research, 176*(1), 53-65.
- Bryant, D., & Church, R. M. (1974). The determinants of random choice. *Animal Learning & Behavior, 2*(4), 245-248.
- Chandana, S. R., Behen, M. E., Juhasz, C., Muzik, O., Rothermel, R. D., Mangner, T. J., ... & Chugani, D. C. (2005). Significance of abnormalities in developmental trajectory and asymmetry of cortical serotonin synthesis in autism. *International Journal of Developmental Neuroscience, 23*(2), 171-182.
- Chugani, D. C., Muzik, O., Behen, M., Rothermel, R., Janisse, J. J., Lee, J., et al. (1999). Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children. *Annals of Neurology, 45*(3), 287-295.
- Cohn, J., & Paule, M. G. (1995). Repeated acquisition of response sequences: the analysis of behavior in transition. *Neuroscience & Biobehavioral Reviews, 19*(3), 397-406.
- Courchesne, E., Redcay, E., Morgan, J. T., & Kennedy, D. P. (2005). Autism at the beginning: microstructural and growth abnormalities underlying the cognitive and behavioral phenotype of autism. *Developmental Psychopathology, 17*(3), 577-597.
- Denney, J. & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior, 26*(2), 154-162.
- Da Silva Souza, A., Abreu-Rodrigues, J., & Baumann, A. A. (2010). History effects on induced and operant variability. *Learning & Behavior, 38*(4), 426-437.

- Eckerman, D. A., & Lanson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior*, *12*, 73-80.
- Ferrari, P. F., Palanza, P., Parmigiani, S., de Almeida, R. M. M., & Miczek, K. A. (2005). Serotonin and aggressive behavior in rodents and nonhuman primates: predispositions and plasticity. *European Journal of Pharmacology*, *526*(1), 259-273.
- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, *9*(2), 250-258.
- Grow, L. L., Kelley, M. E., Roane, H. S., & Shillingsburg, M. A. (2008). Utility of extinction-induced response variability for the selection of mands. *Journal of Applied Behavior Analysis*, *41*, 15-24.
- Guillott, P-V. & Chapouthier, G. (1996). Intermale aggression and dark/light preference in ten inbred mouse strains. *Behavioural Brain Research*, *77*(1), 211-213.
- Hazlett, H. C., Poe, M., Gerig, G., Smith, R. G., Provenzale, J., Ross A., et al. (2005). Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. *Archives of General Psychiatry*, *62*(12), 1366-1376.
- Hopkinson, J. & Neuringer, A. (2003). Modifying behavioral variability in moderately depressed students. *Behavior Modification*, *27*(2), 251-264.
- Hutsell, B. A. & Newland, M. C. (2013). A quantitative analysis of the effects of qualitatively different reinforcers on fixed ratio responding in inbred strains of mice. *Neurobiology of Learning and Memory*, *101*, 85-93.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, *27*(2), 197-209.

- Johansen, E. B., Killeen, P. R., & Sagvolden, T. (2007). Behavioral variability, elimination of responses, and delay-of-reinforcement gradients in SHR and WKY rats. *Behavioral and Brain Functions, 60*(3), 1-11.
- Johnson, J. M., Bailey, J. M., Johnson, J. E., & Newland, M. C. (2010). Performance of BALB/c and C57Bl/6 mice under an incremental repeated acquisition of behavioral chains procedure. *Behavioral Processes, 84*(3), 705-714.
- Kalueff, A. V., Minasyan, A. Keisala, T., Shah, Z. H., & Tuohimaa, P. (2006). Hair barbering in mice: Implications for neurobehavioral research. *Behavioral Processes, 71*, 8-15.
- Kazdin, A. E. (2011). Single-case research designs: Methods for clinical and applied settings (2nd ed.). New York, NY: Oxford University Press
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3-9.
- McCracken, J. T., McGough, J., Shah, B., Cronin, P., Hong, D., Aman, M. G., ... & McMahon, D. (2002). Risperidone in children with autism and serious behavioral problems. *New England Journal of Medicine, 347*(5), 314-321.
- McEachin, J. J., Smith, T., & Lovaas, I. O. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation, 97*(4), 359-372.
- Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis, 33*(2), 151-165.
- Neuringer, A., & Jensen, G. (2010). Operant variability and voluntary action. *Psychological review, 117*(3), 972-933.

- Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, 27(1), 79-94.
- Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, 17(1), 3-12.
- Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning & Behavior*, 21(2), 83-91.
- Neuringer, A. (2004). Reinforced variability in animals and people: Implications for adaptive action. *American Psychologist*, 59(9), 891-906.
- Neuringer, A. (2009). Operant variability and the power of reinforcement. *The Behavior Analyst Today*, 10(2), 319-343.
- Nicolson, R. & Szatmari, P. (2003). Genetic and neurodevelopmental influences in autistic disorder. *Canadian Journal of Psychiatry*, 48, 526-537.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11(3), 429-452.
- Paule, M. G., & McMillan, D. E. (1984). Incremental repeated acquisition in the rat: acute effects of drugs. *Pharmacology Biochemistry and Behavior*, 21(3), 431-439.
- Pesek-Cotton, E. F., Johnson, J. E., & Newland, M. C. (2011). Reinforcing behavioral variability: An analysis of dopamine-receptor subtypes and intermittent reinforcement. *Pharmacology, Biochemistry, and Behavior*, 97, 551-559.
- Popova, N. K. (2006). From genes to aggressive behavior: The role of serotonergic system. *Bioessays*, 28(5), 495-503.
- Posey, D. J. & McDougle, C. J. (2002). Risperidone: A potential treatment for autism. *Current opinion in investigational drugs (London, England: 2000)*, 3(8), 1212-1216.

- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: training for novel behavior. *Journal of the Experimental Analysis of Behavior*, 12(4), 653-661.
- Saldana, L. & Neuringer, A. (1998). Is instrumental variability abnormally high in children exhibiting ADHD and aggressive behavior? *Behavioral Brain Research*, 94(1), 51-59.
- Skinner, B. F. (1953). *Science and human behavior*. SimonandSchuster.com.
- Schwartz, B. (1982b). Reinforcement-induced stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General*, 111, 23-59.
- Tang, X., Orchard, S. M., & Sanford, L. D. (2002). Home cage activity and behavioral performance in inbred and hybrid mice. *Behavioral Brain Research*, 136, 555-569.
- Thompson, D. M. (1974). Repeated acquisition of behavioral chains under chronic drug conditions. *Journal of Pharmacology and Experimental Therapeutics*, 188(3), 700-713.
- Wahlsten D., Hudspeth, W. J. & Bernhardt, K. (1975). Implications of genetic variation in mouse brain structure for electrode placement by stereotaxic surgery. *Journal of Comparative Neurology*, 162(4), 519-531.

CHAPTER 2

Variability Training in Adolescent Mice and Learning in Two strains: Does Experience During Adolescence Affect Adulthood?

By

Megan A. Arnold

Abstract

Variability is an operant dimension of behavior and is imperative to the acquisition of novel responses. Spontaneous and reinforced variability were examined in BALB/c, a putative animal model of autism spectrum disorders, and C57Bl/6 adolescent mice. In adulthood, mice acquired a complex response chain, on an incremental repeated acquisition (IRA) task, in order to determine the effects of adolescent training on adulthood learning. Response variations were compared for 20 BALB/c (10 littermate pairs) and 20 C57Bl/6 (including 9 littermate pairs) adolescent males. Half of the mice from each strain were required to produce a three-response sequence across three levers in a highly variable manner (VAR group). This was established using a threshold procedure in which an infrequently produced sequence was eligible for reinforcement. Littermates of the VAR mice were also required to produce a three-response sequence, but any distribution across the levers was acceptable (ANY group). BALB/c VAR mice responded with more variability than C57Bl/6 VAR mice. Learning in adulthood was assessed using an Incremental Repeated Acquisition (IRA) task with a performance and learning condition. The BALB/c acquired longer response chains and had higher progress quotient (PQ) than the C57Bl/6 strain, as previously reported. Variability training during adolescence did not differentially influence learning in adulthood for either strain. These findings identify limitations on the degree to which the BALB/c strain models invariant responding in ASD and the conditions under which variability training facilitates the acquisition of novel responses.

Variable responding facilitates the acquisition of behavior; without variation, there is nothing upon which reinforcing consequences can select and shape novel behavior (Shahan & Chase, 2002). Reinforcement contingencies usually narrow topographic variability of the reinforced response but, this narrowing of response variability need not be the case. When reinforcement is contingent upon highly variable responding, reinforcing consequences increase variability (Page & Neuringer, 1985). Interestingly, Grunow and Neuringer (2002) demonstrated that reinforcing response variations, using a threshold procedure, also facilitated the acquisition of a concurrently reinforced difficult-to-acquire response sequence.

Although reinforcing variations increases response variability, the degree to which response variability changes as a function of environmental contingencies depends upon the genetics of the individual organism such as strain of rodent or clinical diagnosis. This raises the possibility if procedures that promote variable responding could exert long-term influences over the acquisition of new behavior. Johansen, Killeen, and Sagvolden (2007) showed that spontaneously hypertensive rats (SHR), a putative animal model of ADHD, nose-poke more variable locations than Wistar-Kyoto (WKY) rats when one target nose-poke location was reinforced according to a Fixed Interval (FI) 1' schedule. Although both strains responded on the target location more frequently than non-target locations, the SHR rats responded on non-target locations to a greater extent than WKYs, resulting in greater variability. Some clinical disorders, such as autism spectrum disorders (ASD), are defined, in part, by deficits in response variability (American Psychiatric Association, 2013). Although individuals with ASD show reduced response variability relative to controls, reinforcement of response variations has been shown to increase phoneme variability in children with ASD (Esch, Esch, & Love, 2009) and response

patterns on a keyboard in adolescents with ASD (Miller & Neuringer, 2000) however the degree to which variability increased did not approach the level of variability observed in controls.

The BALB/c mouse strain is a putative animal model of several characteristics of ASD (Brodkin, 2007). Compared to the C57Bl/6 strain, BALB/c mice show reduced social approach to novel mice, less direct contact once approach occurs, aggression, and anxiety (Brodkin, 2007; Guillott & Chapouthier, 1996; Kalueff, Minasyan, Keisala, Shah, & Tuohimaa, 2006; Sankoorikal, Kaercher, Boon, Lee, & Brodkin, 2006). BALB/c mice also share some neurobiological characteristics with individuals with ASD, including smaller corpus callosum and reduced global serotonin (5-HT) synthesis relative to other inbred strains (Ferrari et al. 2005). Although 5-HT synthesis is higher during childhood than adulthood, brain levels of 5-HT are lower in children with ASD than age-matched controls (Chadana et al. 2005; Chugani et al., 1999; McDougle et al., 1996). Reduced 5-HT activity is associated with anxiety, aggression, and repetitive behavior (Hollander Phillips, Chaplin, Zagursky, Novotny, Wasserman, & Iyengar, 2005; Popova, 2006) and administration of SSRIs can alleviate these symptoms (Brodkin et al., 1997; West et al., 2009).

In addition to abnormal levels of 5-HT, autism is also associated with dopamine (DA) overactivity (Lam et al., 2006). One study compared differences in striatal, frontal cortex, hippocampal, and cerebellar levels of DA and 5-HT throughout adolescence in BALB/c and C57Bl/6J mice (Yochum, Medvecky, Cheh, Bhattacharya, & Wagner, 2010). They found that DA and 5-HT levels, in the striatum and frontal cortex, increased with age for both strains, but striatal dopamine increased to a lesser extent for BALB/c mice. In addition, 5-HT levels in the hippocampus increased with age for both of these strains, but 5-HT turnover was greater for the BALB/c strain. Finally, cerebellar DA and 5-HT levels decreased throughout adolescence, but 5-

HT levels remained higher in the BALB/c strain. These findings regarding two neurotransmitter systems (DA and 5-HT) relevant to ASD symptomology, provide further support for the use of the BALB/c as a behavioral and neurobiological model of ASD and for the use of the C57Bl/6 strain as an appropriate comparison strain, particularly in adolescence.

Adolescence is a critical period of development marked by significant neurobiological changes, in particular dopamine pathways (Spear, 2000a). Interestingly, changes in dopaminergic activity seem to moderate reinforcement of variable and invariant behavior (Berridge & Robinson, 1998; Fibiger, LePiane, Jakubovic, & Phillips, 1987; Pesek-Cotton, Johnson, & Newland, 2011). Neurobiological changes that occur during adolescence include reduction in gray matter in the prefrontal cortex and striatum in both humans (Giedd et al. 1999; Sowell et al. 1999) and rodents (Spear, 200b). In rodents, the definition of adolescence depends upon the specific experimental question, but can be limited to post natal day (PND) 21 – 60 (Macri et al. 2002). Interestingly, the number of postsynaptic DA receptors peaks around PND 28 (Spear, 2000b) and the degree to which the number of postsynaptic neurons decreases into adulthood depends upon use (Brenhouse et al., 2008; Laviola et al. 2003).

Given that dopaminergic pathways are relatively plastic during adolescence, compared to adulthood, exposure to environmental manipulations relevant to dopaminergic activity during this critical period of neurobiological development could reasonably be expected to enhance or diminish behavioral outcomes in adulthood. Here, we asked whether BALB/c responding was invariant, relative to the C57Bl/6 strain, on an operant variability task in which response sequences were reinforced according to a threshold procedure. If so, then the BALB/c strain may also be a useful model of perseverative responding in disorders such as ASD. In addition, we asked whether the response history established during adolescence facilitated or impaired

learning in adulthood. Adulthood learning was assessed using an incremental repeated acquisition (IRA) task with a learning component and a performance component.

Methods

Experimental Subjects

The subjects were 20 C57Bl/6 and 20 BALB/c experimentally naïve, male mice purchased as 21-day-old littermate pairs from Harlan Sprague Dawley (Indianapolis). Subjects were group-housed five per cage in amber polysulfone cages on an OptiMICE® housing rack from Animal Care Systems in an AAALAC-accredited laboratory. Mice were separated and individually housed if they became aggressive, to ensure the health of mice in the study. The colony room was maintained on a 12-hr dark-light cycle (lights on at 6:00am). Animal weight was allowed to increase until they reach approximately 24.5g (\pm .5g). This weight was maintained with post-session feeding of 2016 Teklad Global 16% Protein Rodent Diet.

Ten C57Bl/6 and 10 BALB/c mice were assigned to a contingent variability group (VAR) and their littermates were assigned to the control, or noncontingent, variability group (ANY). Mice were group housed with other mice of the same strain and variability criterion unless they became aggressive. Two C57Bl/c mice were excluded because they failed to acquire lever-pressing by PND 40. This did not affect statistical analyses when litter was the repeated measure, an approach that is justified because both mice were from the same litter. One BALB/c VAR and one BALB/c ANY mouse were excluded from IRA analyses because they died before IRA was completed, because they were from different litters, littermate analyses were not performed for IRA. Thus there were 18 C57Bl/6 and 18 BALB/c mice included in IRA analyses.

Apparatus

Experiments were conducted in 10 operant chambers manufactured by Med Associates (Med Associates Inc. St. Albans, VT, model #MED ENV-007, 12.0" L x 9.5" W x 11.5" H) enclosed in sound-attenuating cabinets. Chambers were outfitted with three levers, two retractable levers located on the front panel, each below a yellow stimulus light, to the left (L) and right (R) of a dipper dispenser that dispensed 0.1cc liquid reinforcer (3:1 solution water and sweetened-condensed milk). One nonretractable lever (B) was centered on the back panel directly across from the dipper dispenser. A white-noise generator and a 2.8-W house light were located near the ceiling of the chamber, on the front panel above the two stimulus lights.

Procedure

Autoshaping of lever-pressing. All mice began autoshaping of lever-pressing in 2-hr sessions on PND 24. The first lever autoshaped was either the Left lever (L) or the Right lever (R) and the Back lever was always autoshaped third. The first autoshaped lever (L or R), was counterbalanced within strain and training group, for simplicity of explanation the autoshaping sequence of levers was L first, then R, and finally B. Sessions began with a 70" ITI during which the house light and stimulus lights were off and all levers were retracted. After the ITI ended, the house light and left stimulus light illuminated, and the left lever extended for 7 sec and the dipper arm was raised for 5 sec at which point the left lever retracted, the left stimulus light turned off, and the next ITI began. If the mouse pressed the lever during the 7 sec that it was available, then the milk dipper arm immediately rose. Once the mouse performed 10 responses on the left-lever, autoshaping concluded and a Fixed Ratio 1 (FR1, one response/reinforcer) schedule of reinforcement for left presses began, during which the left stimulus light remained illuminated and the left-lever remained extended for the duration of the session. The session terminated after 60' or once the mouse performed 40 lever presses in the FR1 component. During left-lever

sessions, the right lever remained retracted and the back lever was inactive and blocked by plexiglass. Right-lever and back-lever autoshaping sessions were conducted in a manner similar to left-lever autoshaping except sessions began in the FR1 component. Once autoshaping was complete, meaning mice satisfied the FR1 contingency for the back-lever, the response requirement increased to an FR2 contingency in which two responses distributed across any of the three levers resulted in milk delivery. Upon earning 30 reinforcers under the FR2 contingency, the schedule increased to a FR3 in which any three lever presses across any of the three levers resulted in milk delivery. Once mice earned 30 reinforcers under the FR3 contingency, within 1-hr, the experiment proper began, for all mice, on PND 40.

Variability Training. For ANY mice, the contingency remained the same as the FR3 contingency: each 3-response sequence resulted in access to milk. For the VAR mice, the contingency changed such that milk presentation was contingent upon performing infrequently occurring sequences. The definition of infrequent was updated at the completion of each sequence or trial. After a mouse performed a sequence (e.g., BBB, BLB, LRL, . . . for 27 possible sequences) the count (c) for that sequence incremented by one, then the relative frequency (RF) was calculated for each sequence, by dividing the count by the sum of the counts for each sequence.

$$RF_i = \frac{c_i}{\sum_{i=1}^{27} c_i} \quad (1)$$

In order to weight recently performed sequences more heavily than temporally distal sequences, making recently performed sequences less likely to produce reinforcement on subsequent trials, a weighting coefficient of 0.98 was applied to each sequence after reinforcer delivery (Denney & Neuringer, 1998; Grunow & Neuringer, 2002). At the end of each trial, the weighted count (wc_i)

was divided by the sum of all 27 weighted counts in order to calculate the weighted relative frequency (WRF).

$$WRF_i = \frac{wc_i}{\sum_{i=1}^{27} wc_i} \quad (2)$$

If the weighted relative frequency was below the assigned threshold (0.037 for VAR or 1.000 for ANY) then that trial produced reinforcement. If a sequence did not meet reinforcement criterion, then the weighting coefficient was not applied, but the count (c_i) for that sequence was incremented by one.

Each session began with the house light and two stimulus lights illuminating, the R and L levers extending, and the B lever becoming active (there is no stimulus light associated with the back lever). Pressing a lever produced a 0.5s low tone (2900 Hz). The sequence was recorded (e.g., LLL, LBR, etc) after three presses. If the weighted relative frequency of that sequence (LLL) was below the threshold then the front levers retracted, stimulus lights turned off, the dipper arm raised for 3", and a high tone (4500 Hz) sounded for 3" after which a 3" ITI commenced during which the dipper arm lowered, and the house light turned off. If the relative frequency of the sequence was greater than or equal to the threshold then a 3" ITI commenced. Responses on the back lever, during reinforcement and the ITI, were recorded, but did not count toward responding on the subsequent trial. When the ITI timed out, the R and L levers extended, the B lever became active, and the house light and two stimulus lights illuminated. Sessions terminated after a mouse earned 60 milk deliveries or 1hr elapsed, whichever occurred first.

Sessions during adolescent training were treated as "continuous sessions" for the purpose of calculating sequence frequencies and weighting. This means that data at the end of each session were saved and written into the next session the following day prior to the beginning of experimental procedures and ensures a stringent variability criterion for VAR mice. For example

with continuous sessions, if a mouse performed one sequence frequently in the previous session, then on the first trial of the next session, completing that sequence would not produce reinforcement. This was necessary because pilot testing using noncontinuous sessions revealed that for C57Bl/6 mice the U-values for VAR mice approximated ANY U-values after 5-7 sessions.

Although continuous sessions were necessary to maintain a stringent variability criterion, it also resulted in a decrease in response rates, to near zero levels, for some VAR mice, over the first nine sessions. For these mice, responding occurred at the beginning of the session, but responses did not meet reinforcement criterion, which resulted in response rate decline or cessation of responding. To increase response rates, five ANY trials were introduced at the beginning of each session for VAR mice on session 10. Sequences performed during the first five trials were not included in calculations of variability (U-values), response rate (number of three-response sequences performed), or statistical analyses. This intervention increased response rates and was implemented for the remainder of adolescent training, which concluded on PND 60.

Incremental Repeated Acquisition (IRA). After the variability component was complete, mice were returned to their home cages until they reached adulthood at PND 100, upon which IRA training began. IRA consisted of a performance and a learning condition, conducted in separate sessions as in Bailey, Johnson, and Newland (2010). In the Performance condition, the same chain was acquired each session. In the Learning condition, a different chain was presented each day. The performance and learning sessions began with the illumination of the stimulus lights above each lever, extension of the R and L levers, and activation of the B

lever. For learning sessions, the house light was illuminated and functioned as a discriminative stimulus for learning versus performance session.

IRA sessions always began with the first link of the performance or learning chain. The length of the chain was incremented according to a mastery-based criterion using a backward chaining approach. Six consecutively reinforced responses (e.g., L) in the first link of the chain caused the chain to increment to the second link (e.g., R-L). Three consecutively reinforced chains were required to increment the chain length in links two through five. The maximum possible chain length was a six-link chain. Earning 50 reinforcers in the sixth link ended the session. Correctly completed chains produced 5” access to milk and a 0.5” high tone that sounded at the onset of milk presentation. Incorrect responses resulted in 5” timeout. During the timeout, all stimuli turned off, with the exception that the house light remained illuminated throughout learning sessions. Responses made during the timeout did not count toward the next link requirement, but extended the duration of the timeout by 1s to prevent adventitiously reinforcing responding during the timeout. Following timeout, the chain reset to the first link. There was a unique discriminative stimulus for each link in performance and learning chains. A pure low tone, slowly pulsing low tone, slowly pulsing high tone, pure high tone, rapidly pulsing high tone, and rapidly pulsing low tone were associated with links 1 to 6, respectively. Performance sessions occurred on the first 17 sessions, after which learning sessions were introduced in alternation with performance session and this continued through PND 134 (Table 1).

Data Analysis Strategy

Adolescent training comparison. The primary measure of interest for the last 10 sessions of baseline strain comparison was the U-value, a measure of variability (e.g., Page &

Neuringer, 1985). In Equation 3, the relative frequency (RF) is the probability of a response relative to all other possible responses and n is the number of possible responses. The number of possible sequences is determined by raising the number of levers in the chamber (3) to the length of the sequence (3), thus there are 3^3 or 27 sequences possible.

$$U = \frac{-\sum_{i=1}^{27} RF_i \times \log_2 RF_i}{\log_2 n} \quad (3)$$

U-values can range from 0.00, in which only one of the possible sequences occurred, to 1.00, in which all possible sequences occur with equal frequency. Using a 2x2 ANOVA, U-values for each of the two groups (VAR and ANY), for each strain (C57Bl/6 and BALB/c), were compared to determine differences between- and within-strain. The same statistical analyses were performed to assess differences in response rate. All post-hoc analyses were conducted using Tukey's HSD.

IRA. The primary measure of interest for IRA was progress quotient (PQ), an overall measure of learning that preferentially weights reinforcers earned in later links over those earned in earlier links. In Equation 4, n denotes the position of a link within the chain and Rf_n is the number of reinforcers earned in the n^{th} link. The summation of weighted reinforcers is then divided by the number of unweighted reinforcers earned within the session.

$$PQ = \frac{\sum(n \cdot Rf_n)}{\sum Rf_n} \quad (4)$$

This preferential weighting accounts for differences in chain length and difficulty. Furthermore, PQ captures the mouse's ability to acquire the change better and with a wider dynamic range than other dependent measures (Bailey et al., 2010). Other dependent measures include maximum chain length (MCL), accuracy, number of reinforcers earned, and number of errors. Dependent measures from performance sessions that occurred prior to introducing learning chains, performance sessions that occurred after the introduction of learning chains, and learning

sessions were analyzed in three separate 2(strain) x 2(contingency) RMANOVA. All post-hoc analyses were conducted with Tukey's HSD.

Results

Adolescent Training Comparison

For the last 10 sessions of training during adolescence, there was a significant three-way strain x contingency x session interaction ($F(10, 357) = 2.851, p = 0.002$). For BALB/c VAR mice, U-values significantly increased, while U-values for BALB/c ANY mice significantly decreased across the last 10 sessions. There was a significant strain x contingency interaction on U-values ($F(1, 357) = 13.4, p < 0.001$). BALB/c VAR mice had the highest U-values, C57Bl/6 VAR mice had intermediate U-values, and both BALB/c ANY and C57Bl/6 ANY had U values (averaged across all sessions) that were indistinguishable from each other (Figure 1, top panel). There were no significant changes in U-values for either C57Bl/6 group (Figure 1, top panel).

There were also effects of strain and contingency as well as their interactions on the total number of sequences, a measure of response rate, performed within a session, for the last 10 sessions of adolescent strain comparison. Figure 1 (bottom panel) shows the total number of sequences performed within a session for each group. Response rates significantly increased across sessions ($F(10, 357) = 13.659, p < 0.001$). BALB/c VAR mice performed significantly more sequences within a session than the other three groups, which did not differ from one another (BALB/c ANY vs C57BL/6 ANY $t(357) = 1.165, p = 0.649$; BALB/c ANY vs C57BL/6 VAR $t(357) = 0.911, p = 0.799$; C57BL/6 ANY vs C57BL/6 VAR $t(357) = -0.337, p = 0.987$; Figure 1, Bottom Panel). This high rate of responding for BALB/c VAR mice appeared to drive a main effect of strain ($t(357) = 3.465, p = 0.001$) and a main effect of contingency ($t(357) = -3.900, p < 0.001$).

For both strains, mice in the ANY group showed greater preference for certain sequences, although the preferred sequence varied between mice and drifted over the course of the study, the preferred sequence was one that had one or no changeovers between levers, replicating Pesek-Cotton et al., (2011). In contrast, mice in the VAR groups, for both strains, did not show an extreme degree of preference for any sequence over the others, which indicated the VAR mice varied the sequences they performed within a session. Figure 2 shows the relative frequencies for two littermates, one assigned to the VAR condition and the other to the ANY condition. Relative frequencies for each sequence are taken from session 20 of variability training for representative BALB/c (top panel) and C57Bl/6 (bottom panel) mice.

IRA During Adulthood

Figure 3 shows PQ (top-left panel), MCL (top-right panel), reinforcers earned (middle-left panel), accuracy (middle-right panel), and incorrect responses (bottom-left panel) on performance sessions for each group. The vertical dotted line between sessions 17 and 18 indicates when the first learning session was introduced. Prior to the introduction of learning chains, there were no significant main effects of strain or contingency and no significant interactions on dependent measures. Differences between strains became apparent for dependent measures of the performance chain, once learning chains were introduced.

Due to the large impact of introducing learning chains on performance sessions, Figures 4 and 5 display these measures, beginning at Session 18, with learning and performance chains shown separately. Figure 4 shows data for PQ (top row), MCL (middle row), and total sequences (bottom row) during performance sessions after learning chains were introduced (left) and on learning sessions (right). Figure 5 shows overall accuracy (top row), number of reinforcers

earned (middle row), number of incorrect responses (bottom row) on performance sessions after the introduction of learning chains (left) and on learning chains (right).

Progress Quotient (PQ). On performance sessions, prior to the introduction of learning chains, PQ increased across sessions ($F(16, 518) = 16.532, p < 0.001$), with no effects of or interactions involving strain or adolescence variability condition (Figure 3, top-left panel). After the introduction of learning chains, PQ continued to increase across sessions for all groups ($F(1, 322) = 6.052, p < 0.001$; Figure 3, top-left panel), but this increase was greater for the BALB/c mice than C57Bl/6 mice ($F(1, 322) = 102.862, p < 0.001$; Figure 4, top-left panel). For learning chains there was a significant strain x contingency x sequence interaction ($F(10, 321) = 2.021, p = 0.031$). For some learning chains, as indicated by different sessions, the BALB/c mice had higher PQs than C57Bl/6 mice and for some chains C57Bl/6 VAR and C57Bl/6 ANY mice differed in PQ (Figure 4, top-right panel). Overall, there was a significant main effect of strain on learning chains with BALB/c mice usually having higher PQs than C57Bl/6 mice ($F(1, 321) = 14.901, p < 0.001$). Interestingly, for BALB/c mice, the PQ was indistinguishable between the VAR and ANY groups, whereas C57Bl/6 mice differed from each other on some learning chains ($F(10, 321) = 2.021, p = 0.031$; Figure 4, top-right panel). Finally, for learning sessions, PQ varied with the specific chain being learned chain ($F(10, 321) = 9.553, p < 0.001$; Figure 4, top-right panel).

Maximum Chain Length (MCL). Prior to the introduction of learning chains, MCL increased across sessions for all groups ($F(16, 528) = 14.699, p < 0.001$; Figure 3, top-left panel). Once learning chains were introduced, MCL reached on performance sessions continued to increase for BALB/c mice, but MCL did not continue to increase for C57Bl/6 mice (Figure 3, top-left panel). The introduction of learning sessions revealed a main effect of strain; the

BALB/c mice reached significantly longer chain lengths than C57Bl/6 mice ($F(1, 322) = 174.984, p < 0.001$). For learning sessions, there was a significant strain x session interaction on MCL ($F(10, 321) = 4.790, p < 0.001$). Overall, the BALB/c mice reached longer chain lengths than C57Bl/6 mice ($F(1, 321) = 35.524, p < 0.001$), but this effect depended upon learning session, or the specific chain being learned (Figure 4, bottom-right panel). Finally, there was a main effect of session ($F(10, 321) = 4.790, p < 0.001$), suggesting that some learning chains were more difficult to acquire than others.

Accuracy. Although there was no systematic trend in accuracy before the introduction of learning chains ($F(16, 528) = 1.058, p = 0.393$), there was a significant strain x session interaction ($F(16, 528) = 2.115, p = 0.007$). This significant interaction appears to be driven by some C57Bl/6 mice who responded very accurately on early sessions and by three BALB/c mice who failed to perform a correct response on the first three sessions, but began to respond correctly on the fourth session without further intervention (Figure 3, middle-right panel). On performance sessions after the introduction of learning chains, there was a significant strain x contingency interaction ($F(1, 322) = 8.651, p = 0.004$). This interaction appears to be driven by the C57Bl/6 strain; C57Bl/6 VAR mice responded more accurately than C57Bl/6 ANY mice on some sessions (Figure 5, top-left panel). These session specific differences by contingency for the C57Bl/6s appears to have also driven a significant main effect of contingency ($F(1, 322) = 9.535, p = 0.002$). For learning sessions, there was a significant contingency x session interaction ($F(10, 321) = 2.560, p = 0.005$; Figure 5, top-right panel). Again, this effect appears to be driven by the C57Bl/6 mice – on some sessions C57Bl/6 ANY mice responded more accurately than C57Bl/6 VAR mice and on other sessions C57Bl/6 VAR mice responded more accurately than C57Bl/6 ANY mice (Figure 5, top-right panel). Furthermore, there was a significant main effect

of session ($F(10, 321) = 2.641, p = 0.004$) indicating that specific learning chains were more difficult to acquire than others.

Reinforcers earned. On performance sessions before learning sessions were introduced, there was a significant strain x session interaction ($F(16, 528) = 1.923, p = 0.016$; Figure 3, middle-left panel). On early sessions, C57Bl/6 mice earned more reinforcers than BALB/c mice. This finding is likely due to the fact that three BALB/c mice failed to earn any reinforcers for the first three sessions. In addition, the number of reinforcers earned significantly increased across sessions ($F(16, 528) = 2.892, p < 0.001$; Figure 3, middle-left panel). On performance sessions, after the introduction of learning sessions, there was a significant contingency x session interaction ($F(9, 322) = 2.019, p = 0.037$; Figure 5, middle-left panel). Overall, VAR mice earned more reinforcers than ANY mice ($F(1, 322) = 12.085, p = 0.001$). This effect appears to be driven by the C57Bl/6 strain – C57Bl/6 VAR mice earned many more reinforcers than C57Bl/6 ANY mice for some performance sessions after learning chains were introduced. For learning sessions there was a significant contingency x session interaction ($F(10, 321) = 1.908, p = 0.043$). Again, this interaction appears to be driven by the C57Bl/6 group with differences between C57Bl/6 VAR and C57Bl/6 ANY mice on some learning sessions (Figure 5, middle-right panel). Finally, the number of reinforcers earned depended upon the specific learning chain indicating that some chains were more difficult to acquire than others ($F(10, 321) = 3.131, p = 0.001$).

Incorrect responses. On early performance sessions, prior to the introduction of learning chains, the number of incorrect responses increased across sessions for all groups ($F(16, 528) = 7.358, p < 0.001$; Figure 3, bottom-left panel). This increase in incorrect responses is likely an artifact of mice reaching longer chain lengths, because each novel link initially occasioned

incorrect responding. After learning sessions were introduced, there were no consistent effects of strain or contingency on the number of incorrect responses (Figure 5, bottom-left panel). For learning sessions, there was a significant strain x contingency x session interaction ($F(10, 321) = 1.963, p = 0.037$; Figure 5, bottom-right panel). On the first two sessions, C57Bl/6 ANY mice made fewer incorrect responses than C57Bl/6 VAR mice, but made a similar number of incorrect responses as other groups until the last four sessions. For the last four sessions, C57Bl/6 VAR mice made fewer incorrect responses than C57Bl/6 ANY mice but a similar number of incorrect responses as both BALB/c groups.

Learning Chains. Closer inspection of learning chains that interacted with strain or contingency for dependent measures (sessions 18, 26, 29, 31, and 37) did not reveal a systematic effect of *type* of learning chain that produced these differences (data not shown). For example, this effect was observed with the first link of the chain was on the R (session 18), on the L (session 26), and on the B (sessions 29, 31, and 37).

Discussion

The aims of the current study were (1) to assess reinforced and spontaneous response variability in two mouse strains during adolescence and (2) to assess whether establishing a history of variable or invariant responding in adolescence affected learning and steady-state performance, in adulthood, on an IRA task. The results from this study extend, to adolescent mice, prior research demonstrating that response variations are sensitive to their consequences in adult pigeons, rats, and humans (e.g., Page & Neuringer, 1985; Neuringer, 2004). Interestingly, there was no main effect of strain on variability in adolescence. This was due to the large difference in U-values for the BALB/c VAR and ANY mice. Although U-values for the two C57Bl/6 groups significantly differed, the magnitude of this difference was much less than for

BALB/c strain (Figure 2, top panel). Differences between these two strains, in the VAR condition, could reflect how these strains differ in the coupling of reinforcers to a response. Hutsell and Newland (2013) applied Mathematical Principles of Reinforcement (MPR; Bradshaw and Killeen, 2012) to response-rate functions under various fixed-ratio schedules of reinforcement when the reinforcer was sweetened condensed milk (as in this study) or sucrose pellets, for three mouse strains (BALB/c, C57Bl/6, DBA/2). They found that a reinforcer is coupled with fewer previous responses, as reflected in a parameter termed *saturation rate*, in the BALB/c strain than in the C57Bl/6 strain. Stated differently, the delay-of-reinforcement gradient reaches far back in time for the C57Bl/6 mice so a reinforcer can strengthen more responses whereas in the BALB/c strain a reinforcer had a powerful but temporally concentrated effect. Machado (1997) proposed that operant variability arises when switching between response devices is reinforced. It is possible that the BALB/c's superior performance in the present study may have occurred because of the reinforcement of recent switches or changeovers. The shallower delay-of-reinforcer gradient for the BALB/c mice could have selected runs of responses on a lever. Although the threshold procedure does not directly reinforce the number of changeovers in a sequence, in fact, sequences with no changeovers were reinforced if they occurred infrequently switching from one response device to another is imperative for maintaining highly variable responding in this task (Figure 2). This account is speculative but it is not inconsistent with Machado's (1997) conclusion that reinforced switching is the mechanism that underlies operant variations. Future research should assess reinforced changeovers, in mice, to extend these findings regarding the behavioral mechanism of operant variations.

In addition to differences in variability during adolescence, there were strain differences in response rate, as measured by the number of sequences performed within a session. Overall,

BALB/c mice performed more sequences within a session than C57Bl/6 mice, as in a previous study using simple lever-pressing (Hutsell & Newland, 2013). VAR mice responded at a higher rate than ANY mice. This finding also replicates previous work in our lab with rats (Pesek et al., 2011) but in the present study, this effect was only present after the implementation of training trials at the beginning of a session, which were introduced in an effort to reduce extinction of lever-pressing observed for several mice and are likely responsible for the higher response rates observed for VAR mice, after their introduction. This lack of effect on response rates prior to the implementation of training trials may have been due to procedural differences between Pesek-Cotton et al. (2011) and the current study. In Pesek-Cotton et al. (2011) responding was reinforced according to a lag schedule in which each session began with several trials that always resulted in reinforcement, in order to establish a look-back window for subsequent reinforcement. Once we added five ANY trials to the beginning of each VAR session, differences between VAR and ANY mice emerged.

One goal of the first phase of this study was to assess whether establishing a history of variable responding would facilitate the acquisition of novel response chains in adulthood. The effect of adolescent variability training appears to be strain specific. Establishing a history of variable responding during adolescence in BALB/c mice did not improve their already superior performance on IRA. In contrast, establishing a history of variable responding during adolescence for C57Bl/6 mice resulted in a subtle and inconsistent improvement on some dependent measures of IRA. Variability training during adolescence for C57Bl/6 mice did not improve PQ or MCL prior to the introduction of learning chains, but this history did increase reinforcers earned, and decrease incorrect responses on some performance sessions after learning chains were introduced (Figure 3, sessions 18-37). Also, on some learning sessions there were

differences between C57Bl/6 VAR and C57Bl/6 ANY mice on PQ, accuracy, reinforcers earned, and number of incorrect responses (Figure 4 & Figure 5).

This study replicated strain differences on IRA, with BALB/c mice having higher PQ and reaching longer MCL than C57Bl/6 mice (Johnson et al., 2010). Given that there were no differences between BALB/c VAR and ANY mice on dependent measures of IRA and occasional differences between C57Bl/6 VAR and ANY mice on dependent measures of IRA, it is possible that variability training may be more beneficial for groups that do not perform well on a task, such as the C57Bl/6 mice on IRA. It is unclear why strain differences between BALB/c and C57Bl/6 mice and adolescent training differences for the C57Bl/6 mice did not emerge until after the introduction of learning chains. It is possible that the impact of variability training may be greater when it occurs simultaneous with the acquisition of a complex task, as in Grunow and Neuringer (2002). Based on the results of this study, variability training as an early intervention may not be sufficient to facilitate subsequent learning or the maintenance of variability training's effect may not carry over following an extended delay. As such, introducing learning chains may have brought about sufficient variability for strain and history effects to emerge on IRA. Future studies should address the effect of concurrent variability training on IRA or assess the effects of variability training on subsequent learning following a shorter delay, for these two strains.

Although the findings of this study confirm strain differences in IRA reported by Johnson et al. (2010), with BALB/c mice having higher PQ and reaching longer MCLs than C57Bl/6 mice on performance sessions after learning chains were introduced, here BALB/c performance was deficient compared to previous studies with this strain (Bailey, Hutsell, & Newland, 2013; Shen, Pope, Hutsell, & Newland, 2015). In the present study, C57Bl/6 mice rarely reached MCLs greater than three and BALB/c mice rarely reached MCLs greater than four (Figure 3,

top-right panel Figure 4, bottom-right panel) . Furthermore, for all groups, PQ never exceeded 3.0, indicating that even when mice reached a MCL of 4, they earned few, or no, reinforcers at the longest chain length. Therefore, it appears that establishing a history of performing sequences consisting of three responses in adolescence limited the MCL that mice reached and PQ in adulthood, for all groups. This artifact of training may have restricted the range of dependent measures to the extent that between-group differences were undetectable, particularly between VAR and ANY mice.

For the C57Bl/6 strain, subtle and inconsistent differences between VAR and ANY appeared, but this effect adolescent training may have been more consistent or robust if not for this artifact of training three-response sequences in adolescence. For the BALB/c strain, which normally reaches much longer MCL and higher PQ in adulthood, adolescent training differences may have emerged at longer chain lengths. Shen et al. (2015) reported similar MCL and PQ for their adolescent BALB/c group as was observed in this study. They did not report measures of IRA for this group into adulthood, so it is unclear if the inferior performance of the adolescent trained mice, relative to adult trained mice, would have continued into adulthood. Future studies comparing the effects of establishing a history of variant or invariant responding, for these two strains, should control for the length of sequence in training and adulthood by increasing the response unit in adolescence to six lever-presses. Doing so may bring about more consistent differences for VAR and ANY mice, particularly for the C57Bl/6 strain and help elucidate the effects of training in adolescence on performance in adulthood.

The BALB/c strain has been proposed as an animal model of some characteristics of ASD, such as low sociability, increased anxiety and depression, increased brain size, decreased corpus callosum, and low levels of 5-HT (Brodin, 2007), the findings of this study indicate that

the BALB/c strain may not be a good model of some of the perseverative responding characteristic of ASD. When reinforcement was contingent upon variable operant responding, U-values for the BALB/c strain exceeded what would be expected in an animal model of ASD. For human adolescents with ASD, contingent reinforcement of variations significantly increased U-values, but the increase in U-values did not approach and certainly did not exceed that of control participants (Miller & Neuringer, 2000). This conclusion regarding the utility of the BALB/c as an animal model of perseverative responding characteristic of ASD was further supported by the superior performance of the BALB/c strain, compared to C57Bl/6 strain, on IRA because measures of IRA are correlated with IQ scores in humans (Baldwin, Chelonis, Prunty, & Paule, 2012). One study found that BALB/c mice show increased home cage activity, specifically increased vertical jumping relative to the C57Bl/6 strain (Tang, Orchard, & Sanford, 2002). If this high level of home cage activity is indicative of locomotor perseveration in the BALB/c strain, then these findings suggest a functional distinction between two forms of perseveration. One, which we sought to model here, but did not observe, is perseveration of an arbitrary response maintained by an explicit reinforcer. A second, which is modeled by stereotypies, is visible as an excessive rate of some form of a naturally occurring response class, such as increased home cage activity specifically increased vertical jumping, relative to the C57Bl/6 strain (Tang, Orchard, & Sanford, 2002). In our lab we have observed similar home cage behavior in the BALB/c strain. This functional distinction between forms of perseveration is informative for future studies using the BALB/c strain to model characteristics of ASD.

Finally, establishing a history of variant or invariant responding in adolescence using a threshold procedure does not appear to greatly affect measures of learning in adulthood on an IRA task. It is possible that the facilitative effect of variability training is greatest when it is

concurrent with the acquisition of a difficult response, as in Grunow and Neuringer (2002). Future research should control for the length of delay between variability training and subsequent testing and the length of the response unit during variability training to better control for procedural differences between these two tasks. Doing so will likely elucidate the conditions under which variability training facilitates the acquisition of a response in these two strains of mice.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author
- Bailey, J. M., Hutsell, B. A., & Newland, M. C. (2013). Dietary nimodipine delays the onset of methylmercury neurotoxicity in mice. *NeuroToxicology*, *37*, 108 – 117.
- Bailey, J. M., Johnson, J. E., & Newland, M. C. (2010). Mechanisms and performance measures in mastery-based incremental repeated acquisition: behavioral and pharmacological analyses. *Psychopharmacology*, *209*(4), 331-341.
- Baldwin, R. L., Chelonis, J. J., Prunty, P. K., & Paule, M. G. (2012). The use of incremental repeated acquisition task to assess learning in children. *Behavioral Processes*, *91*(1), 103-114.
- Berridge, K. C. & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*, 309-369.
- Bradshaw, C. M. & Killeen, P. R. (2012). A theory of behavior on progressive ratio schedules with applications in behavioural pharmacology. *Psychopharmacology*, *222*, 549 - 564
- Brenhouse, H. C., Sonntag, K. C., & Anderson, S. L. (2008). Transient D1 dopamine receptor expression on prefrontal cortex projection neurons: relationship to enhanced motivational salience of drug cues in adolescence. *The Journal of Neuroscience*, *28*, 2375-2382.
- Brodkin, E. S., McDougle, C. J., Naylor, S. T., Cohen, D. J., & Price, L. H. (1997). Clomipramine in adults with pervasive developmental disorders: A prospective open-label investigation. *Journal of Child and Adolescent Psychopharmacology*, *7*, 109 - 121
- Brodkin, E. S. (2007). BALB/c mice: low sociability and other phenotypes that may be relevant to autism. *Behavioural brain research*, *176*(1), 53-65.

- Chandana, S. R., Behen, M. E., Juhasz, C., Muzik, O., Rothermel, R. D., Mangner, T. J., ... & Chugani, D. C. (2005). Significance of abnormalities in developmental trajectory and asymmetry of cortical serotonin synthesis in autism. *International Journal of Developmental Neuroscience*, 23(2), 171-182.
- Chugani, D. C., Muzik, O., Behen, M., Rothermel, R., Janisse, J. J., Lee, J., et al. (1999). Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children. *Annals of Neurology*, 45(3), 287-295.
- Denney, J. & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior*, 26(2), 154-162.
- Esch, J. W., Esch, B. E., & Love, J. R. (2009). Increasing vocal variability in children with Autism using a lag schedule of reinforcement. *The Analysis of Verbal Behavior*, 25, 73 – 78.
- Ferrari, P. F., Palanza, P., Parmigiani, S., de Almeida, R. M. M., & Miczek, K. A. (2005). Serotonin and aggressive behavior in rodents and nonhuman primates: predispositions and plasticity. *European Journal of Pharmacology*, 526(1), 259-273.
- Fibiger, H. C., LePiane, F. G., Jakubovic, A., & Phillips, A. G. (1987). The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *The Journal of Neuroscience*, 7, 3888-3896.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., ... & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2, 861-863.
- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, 9(2), 250-258.

- Guillott, P-V. & Chapouthier, G. (1996). Intermale aggression and dark/light preference in ten inbred mouse strains. *Behavioural Brain Research*, 77(1), 211-213.
- Hollander, E., Phillips, A., Chaplin, W., Zagursky, K., Novotny, S., Wasserman, S., & Iyengar, R. (2005). A placebo controlled crossover trial of liquid fluoxetine on repetitive behaviors in childhood and adolescent Autism. *Neuropsychopharmacology*, 30, 582-589.
- Hutsell, B. A. & Newland, M. C. (2013). A quantitative analysis of the effects of qualitatively different reinforcers on fixed ratio responding in inbred strains of mice. *Neurobiology of Learning and Memory*, 101, 85-93.
- Johansen, E. B., Killeen, P. R., & Sagvolden, T. (2007). Behavioral variability, elimination of responses, and delay-of-reinforcement gradients in SHR and WKY rats. *Behavioral and Brain Functions*, 3, 1 – 11.
- Johnson, J. M., Bailey, J. M., Johnson, J. E., & Newland, M. C. (2010). Performance of BALB/c and C57Bl/6 mice under an incremental repeated acquisition of behavioral chains procedure. *Behavioral Processes*, 84(3), 705-714.
- Kalueff, A. V., Minasyan, A. Keisala, T., Shah, Z. H., & Tuohimaa, P. (2006). Hair barbering in mice: Implications for neurobehavioral research. *Behavioral Processes*, 71, 8-15.
- Lam, K. S., Aman, M. G., Arnold, L. E. (2006). Neurochemical correlates of autistic disorder: A review of the literature. *Research in Developmental Disabilities*, 27, 254 - 289
- Laviola, G., Macri, S., Morley-Fletcher, S., & Adriani, W. (2003). Risk-taking behavior in adolescent mice: psychobiological determinants and early epigenetic influence. *Neuroscience & Biobehavioral Reviews*, 27, 19-31.

- Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior*, *68*, 1 – 25.
- Macri, S., Adriani, W., Chiarotti, F., Laviola, G., 2002. Risk taking during exploration of a plus-maze is greater in adolescent than in juvenile or adult mice. *Animal Behaviour*, *64*, 541-546.
- Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis*, *33*(2), 151-165.
- Neuringer, A. (2004). Reinforced variability in animals and people: Implications for adaptive action. *American Psychologist*, *59*(9), 891-906.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(3), 429-452.
- Pesek-Cotton, E. F., Johnson, J. E., & Newland, M. C. (2011). Reinforcing behavioral variability: An analysis of dopamine-receptor subtypes and intermittent reinforcement. *Pharmacology, Biochemistry, and Behavior*, *97*, 551-559.
- Popova, N. K. (2006). From genes to aggressive behavior: The role of serotonergic system. *Bioessays*, *28*(5), 495-503.
- Sankoorikal, G. M. V., Kaercher, K. A., Boon, C. J., Lee, J. K., & Brodtkin, E. S. (2006). A mouse model system for genetic analysis of sociability: C57Bl/6J versus BALB/cJ inbred mouse strains. *Biological Psychiatry*, *59*, 415-423.
- Shahan, T. A. & Chase, P. N. (2002). Novelty, stimulus control, and operant variability. *The Behavior Analyst*, *25*, 175 – 190.

- Shen, A. N., Pope, D. A., Hutsell, B. A., & Newland, M. C. (2015). Spatial discrimination reversal and incremental repeated acquisition in adolescent and adult BALB/c mice. *Behavioural Processes, 118*, 59 – 70.
- Sowell, E. R., Thompson, P. M., Holmes, C J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience, 2*, 859-861.
- Spear, L. P. (2000a). Neurobehavioral changes in adolescence. *Current Directions in Psychological Science, 9*, 111-114.
- Spear, L. P. (2000b). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews, 24*, 417-463.
- Tang, X., Orchard, S. M., & Sanford, L. D. (2002). Home cage activity and behavioral performance in inbred and hybrid mice. *Behavioral Brain Research, 136*, 555 – 569.
- West, L., Brunssen, S. H., Waldrop, J. (2009). Review of the evidence for treatment of children with autism with selective serotonin reuptake inhibitors. *Journal for Specialists in Pediatric Nursing, 14*, 183 - 191

Table 1

Session	Chain	Chain Type
1 - 17	RBLBRL	Performance
18	BLBRLR	Learning
19	RBLBRL	Performance
20	BLRLBR	Learning
21	RBLBRL	Performance
22	BRBLRL	Learning
23	RBLBRL	Performance
24	BRLRBL	Learning
25	RBLBRL	Performance
26	LBLRBL	Learning
27	RBLBRL	Performance
28	LBRBLB	Learning
28	RBLBRL	Performance
29	LRBRLB	Learning
30	RBLBRL	Performance
31	LRLBRB	Learning
32	RBLBRL	Performance
33	RLBLRB	Learning
34	RBLBRL	Performance
35	RLBRLB	Learning
36	RBLBRL	Performance
37	RLRBRB	Learning

Table Captions

Table 1. A table showing the order of performance and learning chains for each session. The left column shows the session number in IRA, the middle column shows the target chain for that session with the right-most link representing the link closest to reinforcement, and the right column shows the condition in effect for a given session (performance or learning).

Figure 1.

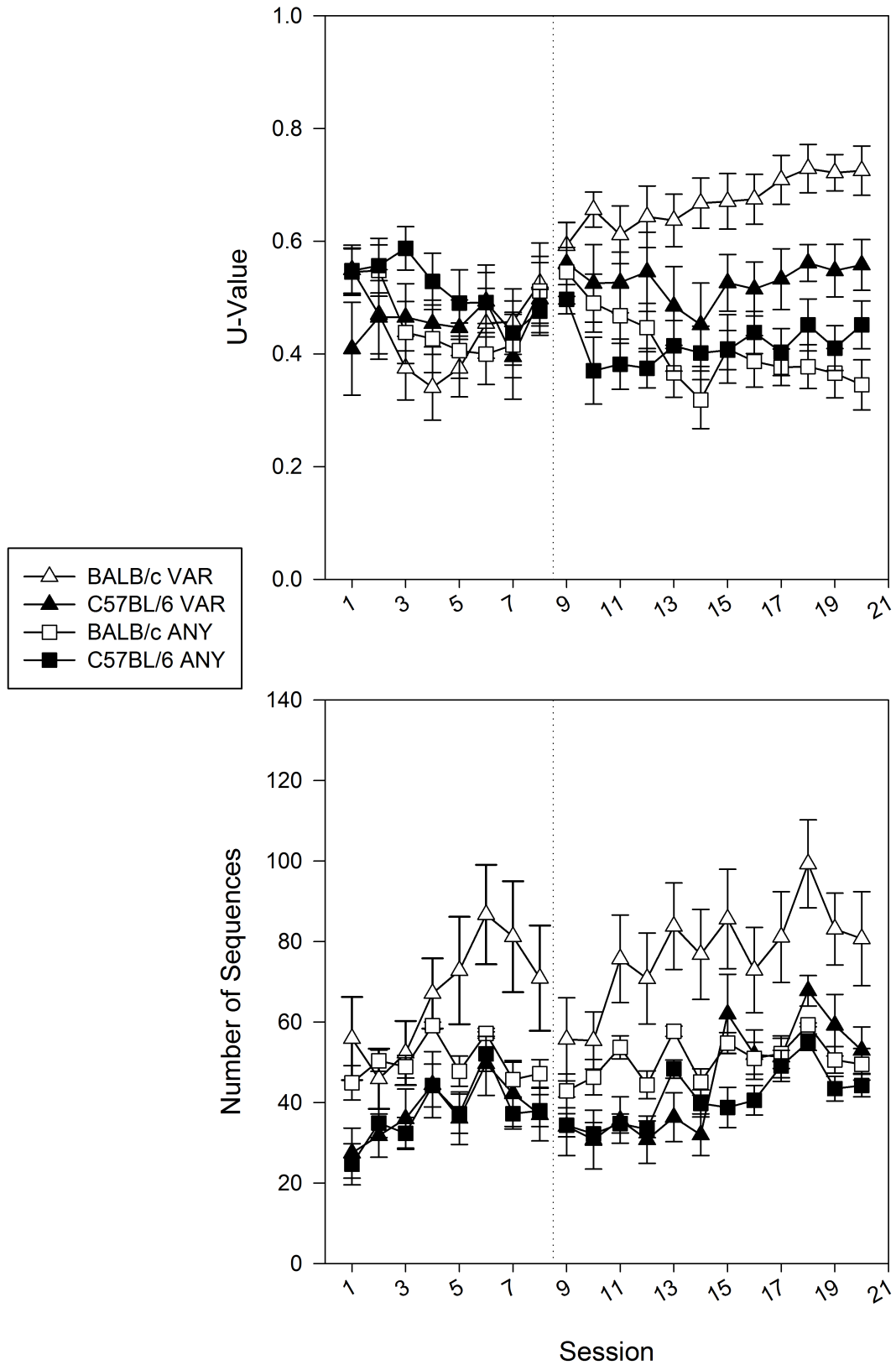


Figure 2.

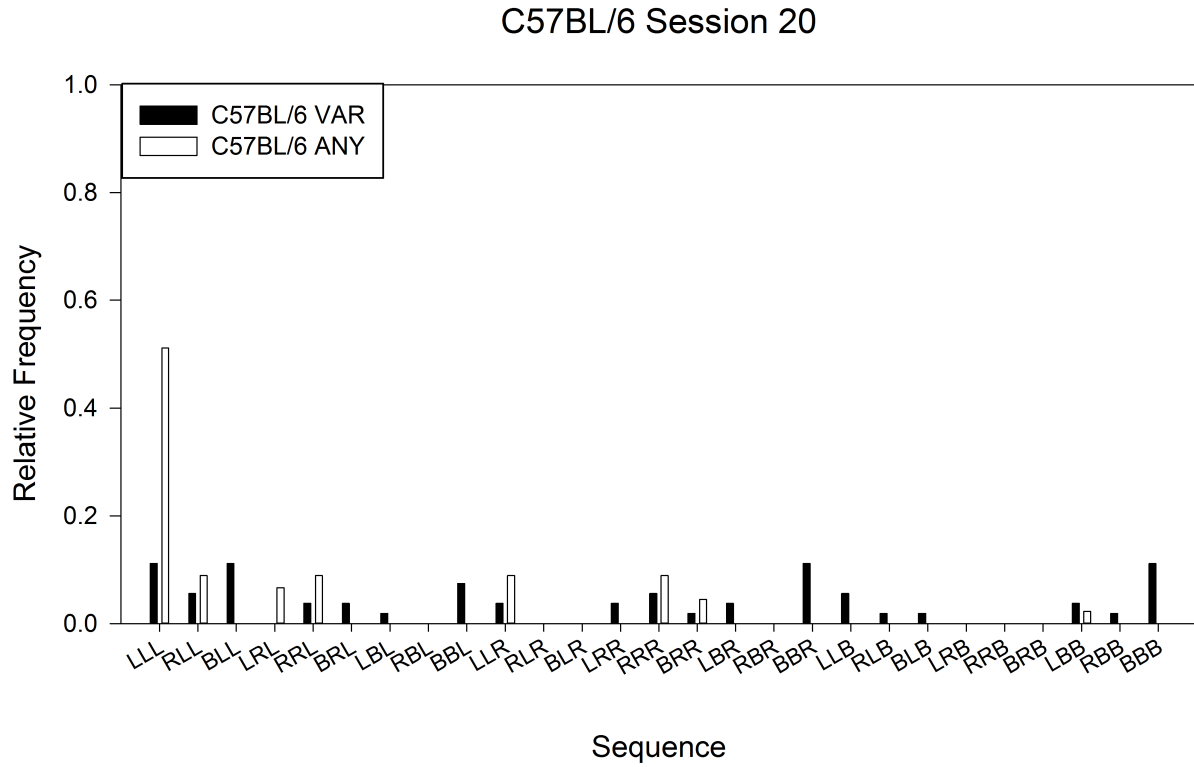
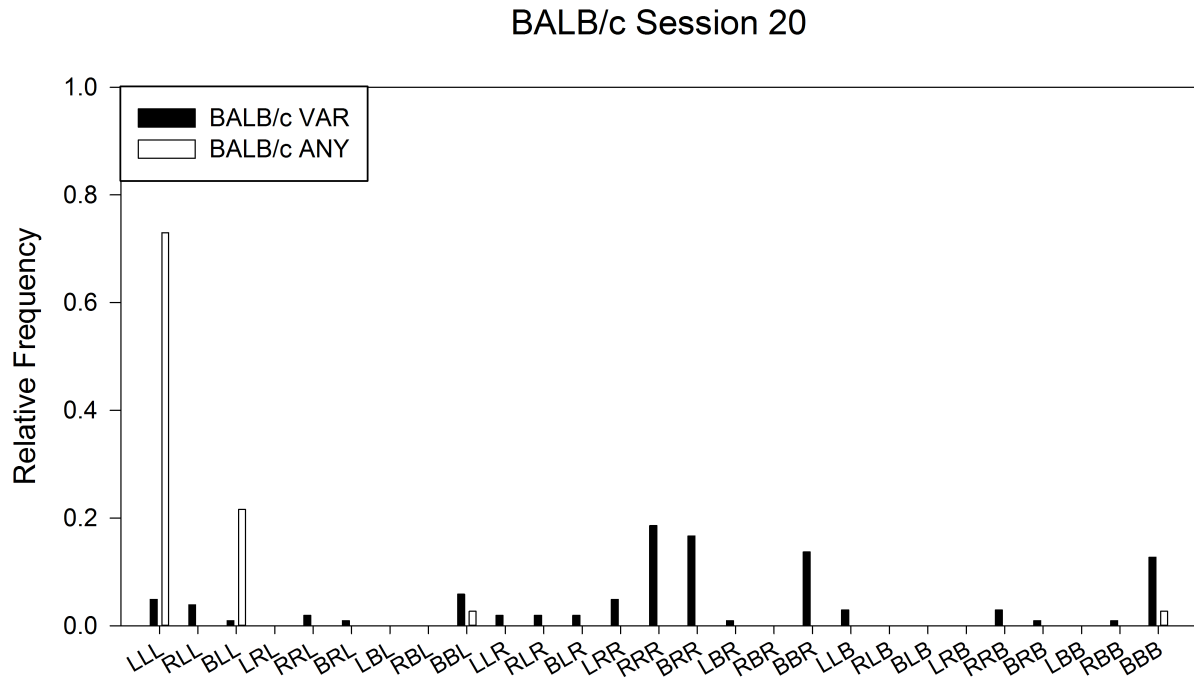


Figure 3.

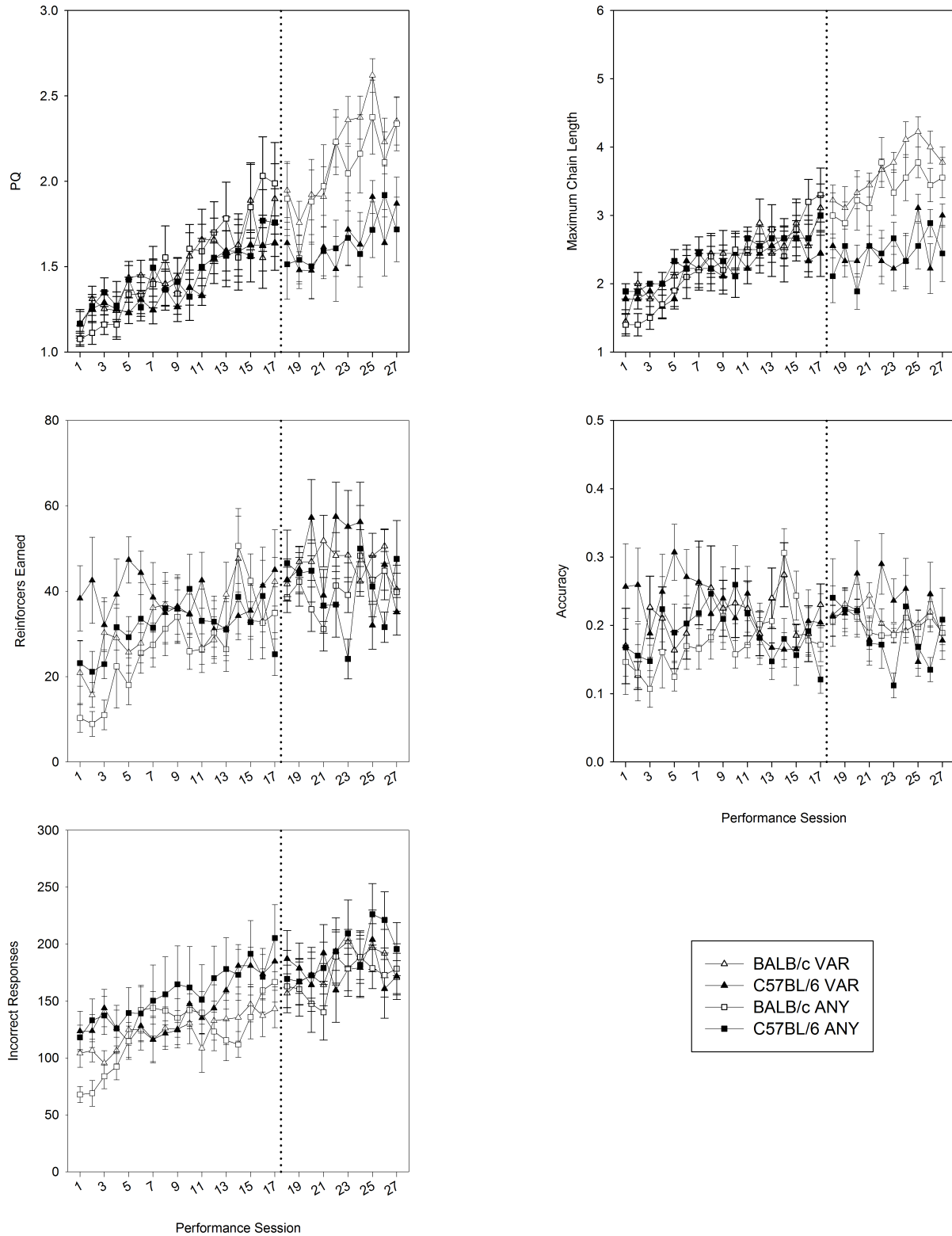


Figure 4.

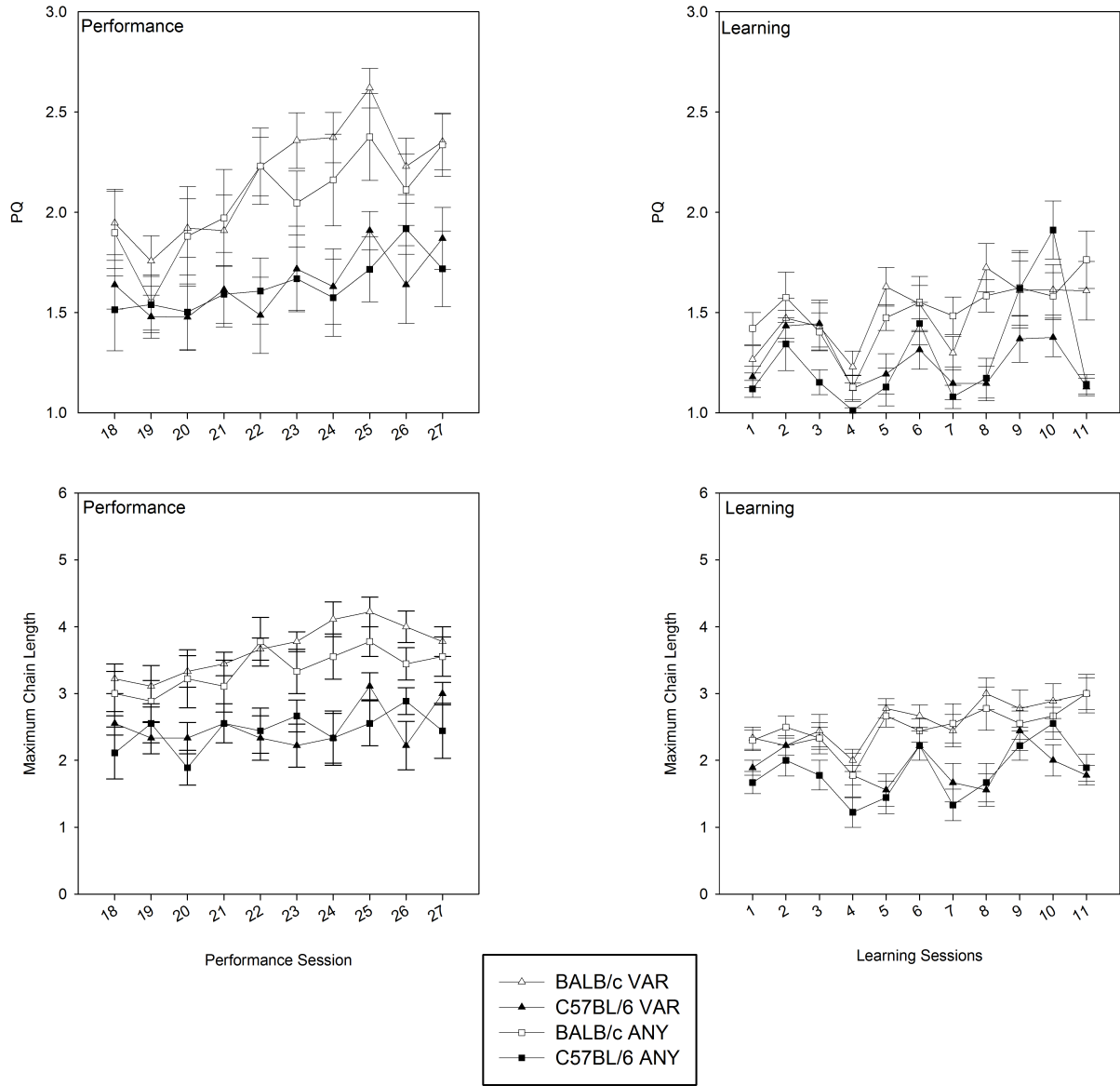


Figure 5.

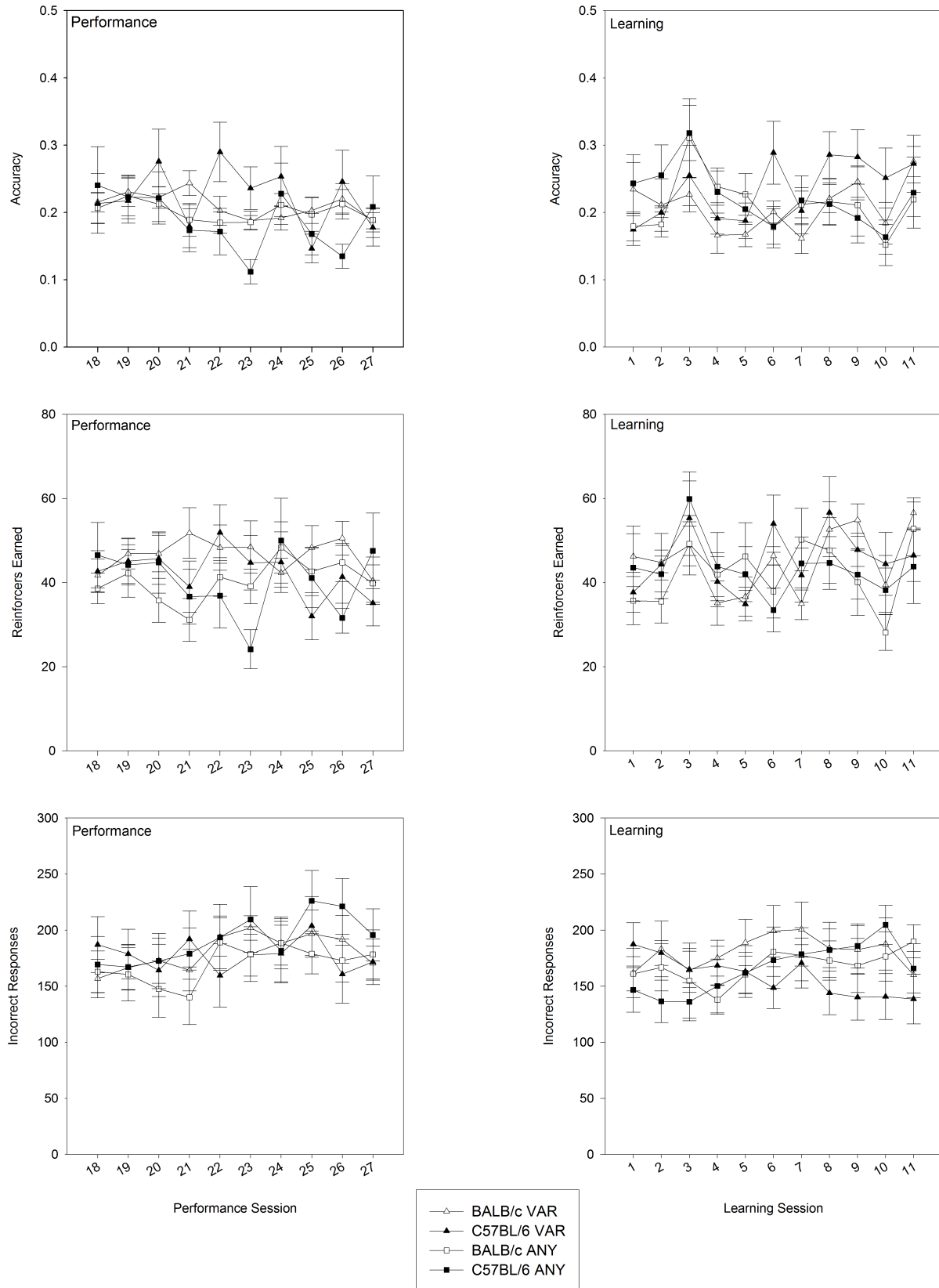


Figure Captions

Figure 1. Mean (\pm SEM) U-values (top-panel) and number of sequences performed (bottom-panel) across sessions during adolescent training. The break in data paths represents the introduction of five training trials at the beginning of the session in order to reduce the effects of extinction observed for some VAR mice.

Figure 2. Individual data from representative BALB/c (top-panel) and C57Bl/6 (bottom-panel) littermates on the final session of VAR (filled bars) or ANY (open bars) training in adolescence. The relative frequency for each of the 27 sequences is calculated by dividing the number of times a given sequence occurred within a session by the total number of sequences performed within the session.

Figure 3. Mean (\pm SEM) PQ (top-left panel), MCL (top-right panel), reinforcers earned (middle-left panel), accuracy (middle-right panel), and incorrect responses (bottom-left panel) for performance sessions in IRA. The break in data paths represents the introduction of learning chains.

Figure 4. Mean (\pm SEM) PQ (top row) and maximum chain length (bottom row) on performance sessions after the introduction of learning chains (left column) and learning sessions (right column).

Figure 5. Mean (\pm SEM) Accuracy (top row), reinforcers earned (middle row), and incorrect responses (bottom row) on performance sessions after the introduction of learning chains (left column) and learning sessions (right column).