3-23-2016

Reinforcement of Variability and Implications for Creativity

Harvey Ray Bayliss

Follow this and additional works at: http://scholarcommons.usf.edu/etd

Part of the Social and Behavioral Sciences Commons

Scholar Commons Citation
Bayliss, Harvey Ray, "Reinforcement of Variability and Implications for Creativity" (2016). Graduate Theses and Dissertations. http://scholarcommons.usf.edu/etd/6065

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Reinforcement of Variability and Implications for Creativity

by

Harvey R. Bayliss

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Applied Behavior Analysis Department of Child and Family Studies College of Behavioral and Community Sciences University of South Florida

Major Professor: Andrew Samaha, Ph.D. Kwang-Sun Blair, Ph.D. Sarah Bloom, Ph.D.

Date of Approval:
March 21, 2016

Keywords: percentile schedule, drawing, automated assessment, Autism Spectrum Disorder

Copyright © 2016, Harvey R. Bayliss
TABLE OF CONTENTS

List of Figures ................................................................................................................................. ii

Abstract ........................................................................................................................................ iii

Chapter One: Introduction .............................................................................................................. 1

Chapter Two: Methods .................................................................................................................... 7
  Participants and Setting ................................................................................................................... 7
  Repetitive Behavior and Rigid Routine Questionnaire Pre-Assessment .................................... 8
  Materials and Equipment .............................................................................................................. 8
  Measures and Data Collection ..................................................................................................... 9
  Dependent Variable .................................................................................................................... 9
  Independent Variable ................................................................................................................ 10
  Pre-Assessment Procedures ........................................................................................................ 11
    Preference Assessment ............................................................................................................... 11
    Token Training ........................................................................................................................ 12
    Reinforcer Assessment ............................................................................................................ 12
  Experimental Design and Procedures ......................................................................................... 13
    General Session Protocol ......................................................................................................... 13
    Random Ratio (RR) .................................................................................................................. 14
    Reinforcement of Same Drawings (Same-Draw) ..................................................................... 14
    Reinforcement of Varied Drawings (Diff-Draw) ...................................................................... 14
    Reinforcement of Varied Lines (Diff-Line) ............................................................................. 15
  Social Validity and Treatment Fidelity Measures ....................................................................... 15

Chapter Three: Results .................................................................................................................... 17
  Preference Assessment ................................................................................................................ 17
  Reinforcer Assessment .............................................................................................................. 17
  Drawing Sessions ...................................................................................................................... 18
  Repetitive Behavior and Rigid Routine Questionnaire Post-Assessment .................................. 21

Chapter Four: Discussion ............................................................................................................... 31

References ....................................................................................................................................... 36

Appendices ..................................................................................................................................... 40
  Appendix A: Repetitive Behavior and Rigid Routine Questionnaire ........................................ 41
  Appendix B: Repetitive Behavior and Rigid Routine Questionnaire Scoring ........................... 42
  Appendix C: Treatment Fidelity Checklist .................................................................................. 43
  Appendix D: USF IRB Approval Letter ....................................................................................... 44
LIST OF FIGURES

Figure 1: Results of the pre-intervention Repetitive Behavior and Rigid Routine Questionnaire for each factor and participant ................................................................. 16

Figure 2: Results of the MSWO preference assessments for each participant .......................... 22

Figure 3: Results of MSWO preference assessments across assessments for Reggie ............... 23

Figure 4: Results of the reinforcer assessments for each participant ........................................ 24

Figure 5: Results of the variability of responses by Paul .......................................................... 25

Figure 6: Distribution of ED rankings for Paul from sessions 25-46 plotted in the form of a cumulative distribution on the top panel (note the double log-axes) and chronologically on the bottom panel ........................................................................ 26

Figure 7: Results of variability of responses by Reggie .......................................................... 27

Figure 8: Results of variability of responses by Doris ............................................................ 28

Figure 9: The figures in the left column display the AED of drawings for the final two sessions of each condition and the average of number of tokens delivered within each condition for each participant ......................................................... 29

Figure 10: Results of the pre-intervention and post-intervention Repetitive Behavior and Rigid Routine Questionnaire for each factor and participant ........................................ 30

Figure 11: Average number of lines per drawing between Diff-Draw and Diff-Line for each participant .............................................................................................................. 35
ABSTRACT

One of the defining characteristics of Autism Spectrum Disorder (ASD) is repetitive, rigid, or stereotyped patterns of behavior. A proposed approach to treating such patterns is to provide reinforcement for response variability. Though research demonstrates that the variability of responses can be influenced by contingencies of reinforcement, no studies have examined the effects of placing contingencies on different units of behavior. The purpose of this study was to examine effects of two modified percentile schedules on variety of completed drawings and individual lines drawn by students with ASD who had been referred for engaging in rigid patterns of behavior. For all three participants that completed drawing sessions, results indicated that drawing variability increased the most when reinforcement was contingent on the variability of the completed drawing, as opposed to a random ratio schedule of reinforcement or reinforcement being contingent on individual lines being varied.
CHAPTER ONE:
INTRODUCTION

Autism Spectrum Disorder (ASD) is considered one of the fastest growing developmental disabilities in the United States (Lord & Bishop, 2009). The Centers for Disease Control and Prevention estimated approximately 1 in 68 children are diagnosed with ASD (CDC, 2015). One of the defining characteristics of ASD is repetitive or stereotypic behavior (Hertzig & Shapiro, 1990; Honey, Leekam, Turner, & McConachie, 2007). Stereotypic behavior includes repetitions of body movements (e.g., hand flapping), and fixated interests (e.g., preoccupation with particular objects), but can also include rigid routines and insistence on sameness (DSM–5; American Psychiatric Association, 2013; Baron-Cohen, 1989; Bodfish, Symons, Parker, & Lewis, 2000).

Positive reinforcement-based interventions are among the most commonly researched approaches to modify behavior in school settings (Northup, Vollmer, & Serrett, 1993). Positive reinforcement has been demonstrated to increase academic behavior in students (Broughton & Lahey, 1978), and is effective at increasing academic behavior specifically in special education classrooms (Iwata & Bailey, 1974). There is also evidence technology-based interventions can improve students’ with ASD outcomes (Goldsmith & LeBlanc, 2004). However, reinforcement can quickly lead to response repetition and a narrowing of response topography (Iversen, 2002; Ross & Neuringer, 2002).
One proposed treatment approach to repetitive behavior is to reinforce response variability. A number of studies have demonstrated that variability itself can be taught and strengthened. Generally, the approach involved reinforcing behavior that is topographically different from previous responses (Goetz & Baer, 1973; Page & Neuringer, 1985; Pryor, Haag, O’Reilly, 1969). Miller and Neuringer (2000) found participants diagnosed with ASD produced low response variability in comparison to typically developing individuals. Participants were asked to play a computer game where pressing a left and right button occasionally resulted in reinforcement. Sixteen possible sequences of 4-response trials were possible. In the intervention phase, reinforcement was contingent on variation of the sequence of responses in comparison to response sequences previously emitted. When variability was reinforced, the participants diagnosed with ASD responded with greater variability than during baseline. These findings suggest that behavioral variability in individuals diagnosed with ASD is sensitive to contingencies of reinforcement. This is particularly noteworthy because the absence of variability is a core characteristic of ASD.

Miller and Neuringer (2000) arranged a percentile schedule of reinforcement. During a percentile schedule of reinforcement, reinforcer delivery is determined by ordering the previous $x$ responses from highest to lowest according to the dimension being shaped, and determines if a response meets a selected level of performance in that ordering (e.g., the level of response is greater than 80% of the previous 20 responses). Thus, the criterion for reinforcement is recalculated after each response and readjusts according to the current performance of the participant (Galbicka, 1994; Miller & Neuringer, 2000). Though Miller and Neuringer (2000) used a percentile schedule of reinforcement, the authors suggested future research explore implementing a lag schedule of reinforcement for increasing
response variation. Lag schedules are effective at increasing response variability (Camilleri & Hanley, 2005; Lee, McComas, & Jawor, 2002; Napolitano, Smith, Zarcone, Goodkin, & McAdam, 2010). Lag-n schedules arrange for the delivery of reinforcement given an occurrence of behavior that is different from the previous n responses (Lee, McComas, & Jawor, 2002; Page & Neuringer, 1985).

In addition to the reinforcement of varied responses being a potential intervention to reduce response repetition, it may also have implications for teaching creativity. Runco and Jaeger (2012) discussed a standard definition of creativity that has two parts: first, creativity must be original, and second it must be useful. Runco and Jaeger described originality in terms of difference or uncommonness. Originality is not found in stereotypic responses, but in varied responses. As differentness can be measured in terms of the variability of responses (i.e., how different a response is from previous responses), it allows for empirical testing of a dimension of creativity as responses topographies can be measured and compared. Creativity plays an important role in education, especially in terms of learning to problem-find and problem-solve (Fasko, 2000; Jacobs & Dominowski, 1981; Martinsen, 1995). Though there is limited research, generating novel or varied responses might be considered aspects of problem-solving techniques such as “brain-storming” (Neuringer, 2004).

Whereas variability is crucial in creativity for a response to be considered useful, it must be germane to the context. Runco and Jaeger (2012) described usefulness in the terms of the response’s fit and appropriateness. For example, if a student is asked to draw a house but the student draws a dog, the response is certainly uncommon, but does not belong to the response class expected by the instruction, and therefore would not be useful.
Variability has been measured using two different approaches, each with its advantages and disadvantages: human judges and automated assessments. When assessing the variability of responses, human judges have been used to decide if the current response is different from those that came before and decide if it is germane to the context (Diener, Wright, Smith & Wright, 2014; Goetz & Baer, 1973; Holman, Goetz, & Baer, 1977; Pryor, Haag, & O’Reilly, 1969). One limitation of using human judges is that the criteria for deciding if a response is different can be very complicated and difficult to operationalize. This limitation has been addressed by limiting possible variance definitions. Goetz and Baer (1973) pre-determined 20 possible block-formation definitions to avoid ambiguity in determining what should be considered a new response. Praise statements were delivered for block-formation that met one of the 20 definitions and were not observed previously in the session. Pryor, Haag, and O’Reilly (1969) determined novel responses emitted by porpoises by having three observers determine if the response was observed in the past. Two out of the three judges had to determine that the response was novel for it to be scored. Pryor, Haag, and O’Reilly wrote that they ended the study in part because the observational system became too unwieldy.

Rather than relying on a human judge to determine variation, an alternative approach was the use of an automated assessment (e.g., Miller & Neuringer, 2000; Neuringer, 1986; Ross & Neuringer, 2002). Ross and Neuringer (2002) delivered reinforcement contingent on participants varying three dimensions (area of rectangle, location, and shape) of drawing a rectangle on a computer screen. Reinforcement was delivered according to whether the area, shape (ratio of height to width), and position of each drawn rectangle occurred at a relative frequency of less than 5%. Relative frequency was determined by the count of a
response that occurred within a pre-defined category, divided by the sum of all occurrences across all categories (a total of 4096 possibilities) and trials. When a response met reinforcement criteria, the computer delivered auditory feedback indicating that points were earned. This allowed for the use of precise and reliable measurement of responses and immediate feedback that might be difficult for human judges to deliver; however, Ross and Neuringer’s approach was incapable of characterizing the degree of difference between two responses, only that they were different. An approach that is capable of characterizing the degree of differences between two responses might have more face validity because it can be compared directly to judgments made by humans.

In addition to methodological issues regarding the characterization of degree of difference, there is also an open question regarding at what level contingencies can be placed. In art, reinforcement may be provided in two distinct phases of creation: during the creative process and after the creation process in response to the product. For example where praise functions as a reinforcer, a parent may praise a child immediately after the child mixes blue and white paint together and depicts a sky; however, reinforcement might also be provided at the end of the creative process in response to the product of all of the child’s responses. Goetz and Baer (1973) examined the effects of immediate praise statements being delivered contingent on the participant producing block formations that were not observed in previously in that session. The trainer delivered praise to responses immediately if the formation of a few blocks was different from other formations created during the session. Photographs were taken of the final structures (which consisted of a series of formations). Data were also collected on the number of different block-formations in the final structures. Results indicated that in the praise of different forms condition the number of different forms in each final structure was greater than baseline or when
praise was provided for repeated formations; however, the authors did not provide any consequence contingent on properties of the final structures. This raises a question about on what scale the effect of reinforcement for varied responses might be best applied. Goetz and Baer provided immediate reinforcement for varied formations (the smaller units that make up structures). But, if the ‘natural lines of fracture’ for varied responses occurs at the level of finished products then a more direct contingency might be to provide consequences on finished products themselves. On the other hand, if smaller units of responses (e.g., Goetz & Baer’s ‘formations’) are more sensitive to contingencies then providing consequences for final products would at best serve through delayed (and thus weaker) reinforcement. No studies have explored the question of at what level contingencies should be placed to lead to varied responding.

The purpose of this study was to examine effects of a modified percentile schedule on drawing variety using a computer program with automated assessment, comparing the effectiveness of reinforcing individual lines versus varied completed drawings—specifically in students with ASD that demonstrate limited variability in creative responses.
CHAPTER TWO:

METHODS

Participants and Setting

Four students diagnosed with ASD, ages 12-14 and attending a local public middle school were selected to participate in the study. Paul was a 14-year-old male, Reggie was a 12-year-old male, Doris was a 13-year-old female, and Mark was a 13-year-old male. All participants had the receptive ability to follow one-step instructions (i.e., “draw a picture.”), could follow instructions to use a computer, and could sit for 3 min. All of the participants were non-verbal and in lower-functioning Exceptional Student Education (ESE) classrooms. Paul, Reggie, and Doris were in one classroom; Mark was in different classroom. All four participants readily pointed to objects they wanted, and Reggie used signs for more and please. The participants’ teachers completed the Repetitive Behavior and Rigid Routine Questionnaire (see below; appendix 1) for each participant. Criterion for exclusion in the study was based on students’ scores from Factor 1 (repetitive motor movements) being less than 9. A student was excluded from the study after completing token training due to severe behavior problems that interfered with the study. Sessions occurred in the classroom. The teacher determined session times. Sessions typically occurred during structured breaks from academic activities, where the students engaged with educational applications on iPads®.
Repetitive Behavior and Rigid Routine Questionnaire Pre-Assessment

Teachers were asked to complete questions about repetitive motor movements, rigidity/adherence to routine and preoccupation with restricted patterns of interest via a scored questionnaire (for social validity purposes). The questionnaire was based on the Repetitive Behaviour Questionnaire-2 (Leekam et al., 2007), but did not include questions about unusual sensory interest. Factor 1 consisted of questions regarding repetitive motor movements. Factor 2 consisted of questions regarding rigidity/adherence to routines. Factor 3 consisted of questions regarding preoccupations with restricted patterns of interest. An example question addressing rigidity/adherence to routine is “Does your student play the same music, game or video, or read the same book repeatedly?” An example question addressing preoccupation with restricted patterns of interest is “Does your student arrange toys or other items in rows or patterns?”

Criterion for exclusion in the study was based on students’ scores from Factor 1 being less than 9. This exclusion criterion was to discriminate repetitive sensory motor behavior (e.g., hand flapping) from rigidity of routines and insistence of sameness behavior. The questionnaire was repeated after the intervention to see if it produced any teacher-observable changes in behavior.

Figure 1 (See page 16) displays the results of the pre-intervention Repetitive Behavior and Rigid Routine Questionnaire. All participants met the inclusion criteria of having Factor 1 scores lower than 9. One teacher completed questionnaires for Paul, Reggie, and Doris, while the questionnaire for Mark was completed by a different teacher.

Materials and Equipment

An iPad® with Internet access was used for the automated assessment program. The researcher delivered physical tokens when indicated by a briefly flashing background on the
computer program. The program consisted of two components: a user interface and database of previously created reference drawings consisting of individual lines or complete pictures, depending on the condition. The user interface presented a blank screen and allowed participants to draw images touching and dragging his or her finger on the drawing area. Depending on the condition and user performance (either a line or an entire picture), the program signaled to the participant and therapist when the current drawing has meet criteria for reinforcement with a green flashing background. A rectangle marked done could be clicked on at any time before the trial ended by the participant to indicate when the drawing was finished. When a drawing was finished, the contents of the drawing area cleared in preparation for the next trial.

Measures and Data Collection

Dependent Variables

Dependent variables included the average number of lines drawn, the average Euclidean Distance (AED) of drawings, the AED of lines, and the number of tokens delivered. A response was defined as an individual line created by the participant. A line was initiated by the participant’s finger touching the drawing area and dragging his or her finger through the area. When the finger no longer touched the drawing area, the line terminated. Euclidean Distance (ED) was calculated between two images by taking the difference between pixels at each position in the current image and reference image, and averaging those differences. An AED score was obtained for each line by calculating ED scores between the current line and the previous 20 lines and averaged. In addition, AED scores for each of the previous 20 lines were calculated by comparing each line against every other line (resulting in a set of 20 ED scores).
percentile ranking of the current line was calculated by comparing the AED of the most recently completed line to that set of scores.

Difference scores and rankings for drawings were obtained as described above, only taking into account completed drawings as a whole instead of individual lines. A drawing consisted of the combinations of all lines produced by the participant during a single trial. A trial was terminated after 30 s from a response initiation or when the participant clicked on the done button. Drawing and line variability was calculated and reported for each session based on the obtained AEDs of the lines drawn during that session only relative to each other, not the previous 20.

**Independent Variable**

The criterion for token delivery in different draw and different line phases described below were responses that received a difference ranking greater than 80% of the previous 20 responses. In separate conditions, variability was placed on individual lines or completed drawings. Percentile rankings often have a uniform distribution. Conceivably, this would allow us to control the average rate of token delivery. For example, setting the minimum percentile of difference scores for new drawings to .80 should result in token delivery 20% of the time. However, in the course of running the first participant Paul, we discovered that percentile rankings of new drawings were in fact not uniformly distributed. This led to a lower than expected rate of token delivery. To correct this, we based criterion levels of token delivery on the distribution of percentile rankings of new drawings and lines obtained during the 16 sessions including and preceding the last session during RR .2.
Pre-Assessment Procedures

Prior to drawing sessions, each participant completed a preference assessment and token training. A reinforcer assessment was then conducted using the drawing program to determine that tokens provided in session functioned as a reinforcer.

Preference Assessment

Eight items were selected based on preferences identified through interviews with teachers and direct observations of the participants. Items included: a puzzle, a toy that made various sounds, Play-Doh®, a ball, balloons, a rain stick, a stuffed toy animal, magazines, books, various colored ribbons, and a plastic toy caterpillar. Following the identification of stimuli, a multiple-stimulus without replacement preference assessment was conducted (DeLeon & Iwata, 1996). On the first trial, all stimuli were presented in a row in front of the participant and he or she was asked to “pick one.” An approach response was defined as touching or pointing to an item within 10 s. The participant received the item immediately for interaction. Each selected stimuli was not presented again in subsequent trials. The position of each stimulus changed randomly across trials. The session ended when all items have either been approached or presented by themselves. Five sessions were conducted. The purpose of the preference assessment was to identify highly preferred items that could be earned with tokens during token training, reinforcer assessments, and drawing sessions. Only the top two scoring items were used in subsequent phases.
**Token Training**

Following the preference assessment, participants were taught that the tokens could be traded in for access to a choice of two items. Participants were given 10 tokens and prompted to exchange them for a choice of two preferred items identified by the preference assessment. Prompts were faded using a most-to-least approach. The phase continued until participants exchanged 10 tokens in a row without prompting. Data were collected on the number of trials completed and tokens collected by the participant to determine if the tokens would independently be traded in for access to items. All participants met criterion within 18 trials.

**Reinforcer Assessment**

Following token training, participants were asked to draw using the drawing program for 10-min sessions. Data were collected on the number of drawings completed in each session. A reversal design was used. During baseline, no tokens were delivered after each drawing was completed. The session was terminated after 10 min if the participant did not make a drawing response for 2 min or if the participant walked away from his or her seat. During the token condition, a token was delivered on a fixed ratio-1 schedule after the completion of each drawing. Sessions were terminated after 10 min after the participant did not make a drawing response for 2 min or if the participant walked away from his or her seat. In both conditions a less-preferred item was available for interaction, and each drawing trial lasted 15 s. At the end of the session, the participant could exchange his or her tokens for access to a selection of high-preferred items. The purpose of the reinforcer assessment was to determine if tokens are functioning as reinforcers before tokens are delivered during drawing sessions.
Experimental Design and Procedures

Participants were exposed to the following phases: a Random Ratio (RR) condition, Different Drawing (Diff-Draw), and Different Line (Diff-Line) condition. Effects of arranging the response requirements of the percentile schedule on drawings versus lines (as assessed through changes in line variability across those phases) were demonstrated using a reversal design. The RR condition acted as a control condition for each participant. If AED scores were high or on an increasing trend during RR (a result consistent with some findings that variable schedules could produce variability), then we conducted a condition reinforcing sameness (Same-Draw) prior to Diff-Draw as a control condition.

General Session Protocol

Sessions consisted of 5 trials or until the session was terminated by the participant. An iPad® with the drawing program was presented to the participant. A session lasted approximately 90s. The number of sessions conducted per day ranged from 1 to 17. For each trial, the trainer provided verbal instructions to draw a specific object (e.g., “Draw some flowers”). Each trial lasted 15 s, or until the participant clicked the done button. The session timer paused briefly (less than 1 s) while the drawing was saved. If no drawing occurred, the trainer provided the verbal prompt, “Your turn to draw.” If the participant engaged in behavior that was incompatible with the instructions provided, the student was redirected with the verbal prompt, “your turn to draw.” At the end of the trial, contents of the drawing area cleared. At the end of the session, the participant had the opportunity to exchange his or her tokens for a choice of two items. Access to the items was provided for 3 min.
Random Ratio (RR)

The purpose of this phase was to provide participants with experience earning tokens at a rate similar to that which they would experience during the subsequent conditions, while not having the delivery of tokens contingent on the variability of the drawing produced. During RR sessions, completion of each drawing resulted in the delivery of a token according to a RR .2 schedule. Prior to implementing RR .2, participants were first exposed to a RR .5 in order to thin the schedule more gradually following exposure to FR-1 during the reinforcer assessment.

Reinforcement of Same Drawings (Same-Draw)

For Doris, a same-draw condition was implemented because her ED scores were high and on an increasing trend during RR. Tokens were delivered following drawings that had ED scores 20% lower than the previous 20. If drawings did not meet varied criteria, the trainer verbally prompt the participant with the next set of trial instructions. The purpose of this condition was to demonstrate experimental control of variability if the other conditions did not.

Reinforcement of Varied Drawings (Diff-Draw)

During Diff-Draw sessions, at the end of each trial, the drawing program signaled if the product scored as a varied response. The trainer provided the participant with a token for each varied drawing indicated by the computer program. If drawings did not meet varied criteria, the trainer verbally prompt the participant with the next set of trial instructions. The criterion for determining token delivery was that the completed drawing’s ED ranked above a determined percentile of the previous 20 drawings. The criterion for determining the percentile needed for token delivery was adjusted so that responses resulted in token delivery approximately 20% of
occurrences. The purpose of this condition was to determine if reinforcement of varied drawings would increase the ED of future drawings.

**Reinforcement of varied lines (Diff-Line)**

During each trial, the drawing program signaled if individual lines scored as a varied response within a session. If lines did not meet varied criteria, drawing continued until the trial ended. If reinforcement was signaled, a token was provided, and the trial continued for the remainder of the trial duration. The criterion for determining token delivery was if a line’s ED ranked above a determined percentile of the previous 20 lines drawn. The criterion for determining the percentile needed for token delivery was adjusted so that responses resulted in token delivery approximately 20% of responses made divided by the average number of lines produced by each participant. The purpose of this condition was to determine if there was a difference between providing reinforcement for varied completed drawings versus providing reinforcement within a trial for varied lines.

**Social Validity and Treatment Fidelity Measures**

Social validity was addressed through pre and post-intervention Repetitive Behavior and Rigid Routine Questionnaires (see above). Treatment fidelity was measured via a treatment fidelity checklist that was completed for each session (see Appendix 3). The checklist consisted of nine questions featuring Yes-or-No questions to determine if the session was properly conducted, and if the drawing program was working properly. The treatment fidelity score was 100% across all conditions and participants.
Figure 1. Results of the pre-intervention Repetitive Behavior and Rigid Routine Questionnaire for each factor and participant. Factor 1 consisted of questions regarding repetitive motor movements. Factor 2 consisted of questions regarding rigidity/adherence to routines. Factor 3 consisted of questions regarding preoccupations with restricted patterns of interest.
CHAPTER THREE:

RESULTS

Preference Assessment

Figure 2 (See page 22) displays the results of the MSWO preference assessment for the four participants. The two highest scoring stimuli were selected for use during token training, reinforcer assessments, and drawing sessions. For Paul, a toy that made different sounds and a plastic toy caterpillar were selected. Further visual analyses of preference assessments for Reggie were needed to discriminate the highest preferred items. Figure 3 (See page 23) displays the results of approaches made by Reggie across assessments. An increase in ranking of the magazine was observed across sessions. A magazine and a puzzle were determined to be the more preferred items. For Doris, a magazine with a crayon and a puzzle were selected for subsequent sessions. For Mark, ribbons and the caterpillar toy were selected.

Reinforcer Assessment

Figure 4 (See page 24) displays the results of the reinforcer assessments for Paul, Doris and Reggie. Mark was excluded from the study after token training, due to problem behavior in the classroom. Each participant showed increases in the number of drawings when drawing resulted in a token delivery that was exchangeable for access to a high-preferred item after the session. Each participant varied the selection of the two items when exchanging tokens after the session. Each participant produced lower quantities of drawings in the baseline condition in
comparison to the token condition. Paul completed twelve drawings in the first baseline session, but the amount of drawings decreased in the second session and no more than three drawings were completed in any other baseline session. The drawings produced during the reinforcer assessment built the data bank of images used to determine the EDs of subsequent lines and drawings.

**Drawing Sessions**

Figure 5 (See page 25) displays the AED of drawings, the AED of lines, and the number of tokens earned across sessions for Paul in the top, middle, and bottom panels, respectively. After a downward trend in the AED of drawings was observed in RR .2 condition, a Diff-Draw 80% (Diff-Draw 80) condition was implemented. In Diff-Draw 80, for a token to be delivered, a drawing’s ED had to rank in the top 80% of the previous 20 drawings (meaning the ED had to be higher than 15 of the previous 20 drawings). During this condition an initial increase in AED of drawings was observed from sessions 8 to 23; however, the AED of drawings began to decrease (sessions 24 to 27) back to levels in range of the RR .2 condition. Throughout the Diff-Draw 80 condition, the density of token delivery was leaner than expected. No tokens were earned from sessions 18-31. After verifying that the software was calculating AED scores correctly, we began to suspect that the distribution of percentiles for new drawings was not uniform as expected. One reason the distribution might not be uniform is if the EDs of drawings were on a downward trend.

We first attempted to address this problem by lowering the criterion for token delivery by an arbitrary amount, and a change was made to Diff-Draw 60. However, inspection of the raw EDs for individual lines and drawing revealed that our assumption about a downward
trend was not true. Another possible explanation may have to do with how EDs were calculated. Because EDs of the last 20 drawing were recalculated with the addition of each new line and drawing, the comparison pool of EDs fluctuated in irregular ways. To study the problem in more detail, we plotted the cumulative distribution of rankings and the rankings across each successive drawing during sessions 25-46 (the top and bottom panels of Figure 6, see page 26). These data clearly show that the distribution of rankings was not uniform. For example, if they followed a uniform distribution, the average should be around .5. However, as can be seen in the bottom panel, the distribution of rankings was very right-tailed and the average ranking was closer to .1.

Thus for Paul, and all subsequent participants, we selected target percentiles that closely approximated rankings at or above 80% of the previous 20 drawings following an inspection of the distribution of rankings for each participant’s new drawings and lines.

For Paul, Diff-Draw 25 was an effective percentile of .76, meaning that about 24% of Paul’s new drawings were ranked at or above the 25% percentile of the previous 20 drawings. When Diff-Draw 25 was implemented, increases in both AED of drawing and line were observed.

Figure 7 (See page 27) displays the AED of drawings and the AED of lines across sessions for Reggie. Figure 7 also displays the number of tokens earned across sessions. Through the RR .2 condition, the AED of drawing and line seemed stable. A Diff-Draw 25 condition was implemented for session 9, based on the percentile schedule that was effective for Paul. The session resulted in each drawing earning a token, leading to the inspection of the distribution of the rankings for each participant to determine the percentile for criterion. It was determined that a Diff-Draw 45 schedule was an effective percentile of .83 for Reggie. During the Diff-Draw 45 condition, an increase was observed in both the AED of drawing and line. For
both Paul and Reggie, in the Diff-Draw condition used for reversal, both AED of drawing and lines increased and remained above RR .2 levels.

As was done for Diff-Draw conditions, the distribution of ED rankings was inspected to determine the appropriate percentile for the Diff-Line condition, using the average number of lines per drawing to attempt to keep the average number of tokens per session the same. Diff-Line 90 for Paul and Diff-Line 95 for Reggie were implemented once a clear change in trend and level was reached in Diff-Draw. For both participants, Diff-Line resulted in a downward trend in AED of drawings for both participants. The levels of AED of lines also dropped but remained variable for both participants. When there was a clear change in trend and level of AED, a reversal back to Diff-Draw was implemented. For both Reggie and Paul, an upward trend of AED of drawing occurred after the condition change, and the level of AED of lines immediately increased while remaining variable. Once there was a clear change in trend and level of AED of drawings, a reversal back to Diff-Line resulted in similar effects to the first Diff-Line condition for both participants.

Figure 8 (See page 28) displays the AED of drawings and the AED of lines across sessions for Doris. Figure 8 also displays the number of tokens earned across sessions. Through 18 sessions in the RR .2 condition, the AED of both drawings and lines remained variable. Research has shown that variable schedules can result in increases in variability (Eckerman & Lanson, 2013), so in order to obtain a stable baseline prior to examining the effects of percentile schedules on drawings and lines, we implemented a Same-Draw 0. Zero was selected because approximately 20% of Doris’s drawings were ranked last in terms of difference from the previous 20. During Same-Draw 0, a gradual decrease in AED of drawings occurred. During Same-Draw 0, the AED of lines increased, and remained variable. It was determined that a Diff-
Draw 30 schedule was an effective percentile of .80 for Doris, meaning that 20% of Doris’ new drawings were ranked above the bottom 30% of her previous 20. During the Diff-Draw 30 condition, an increase was observed in both the AED of drawing and line. When there was a clear change in trend and level, a Diff-Line 85 condition was implemented. Similar to Reggie and Paul, a downward trend occurred in the AED of drawings, and an immediate decrease occurred in level of AED of lines. These effects were replicated in the reversal.

Figure 9 (See page 29) displays the AED of drawings in the final two sessions in each condition for the participants. ED scores were highest in the adjusted Diff-Draw condition for each participant. For Reggie, only one session of Diff-Draw 25 was implemented and represented in the figure. The Diff-Draw and Diff-Line schedules used in the reversal for each participant resulted in similar averages of token delivery per session.

**Repetitive Behavior and Rigid Routine Questionnaire Post-Assessment**

Figure 10 (See page 30) displays the results of the questionnaire both before and after the intervention. For Reggie, a decrease of repetitive motor movements (Factor 1) was reported while rigidity/adherence to routine (Factor 2) and preoccupation with restricted patterns of interest (Factor 3) increased. For Paul, decreases in Factors 2 and 3 were reported. For Doris, decreases in Factors 1 and 2 were reported, but the score for Factor 3 remained the same.
Figure 2. Results of the MSWO preference assessments for each participant.
Figure 3. Results of MSWO preference assessments across assessments for Reggie. The graph depicts the three highest preferred stimuli from Figure 2. Magazines and a puzzle were selected for the reinforcer assessment.
Figure 4. Results of the reinforcer assessments for each participant.
Figure 5. Results of the variability of responses by Paul. The top panel displays the AED of drawings across sessions. The middle panel displays the AED of lines across sessions, and the bottom panel displays the number of tokens earned during each session. A Random ratio .2 (RR .2) condition was followed by Different Draw 80% (Diff-Draw 80), Different Draw 60% (DD 60), then by a Different Draw 25 (DD 25) and Different Line 90% (DL 90) reversal.
Figure 6. Distribution of ED rankings for Paul from sessions 25-46 plotted in the form of a cumulative distribution on the top panel (note the double log-axes) and chronologically on the bottom panel.
Figure 7. Results of variability of responses by Reggie. The top panel presents the AED of drawings across sessions. The middle panel presents the AED of lines across sessions, and the bottom panel presents the number of tokens earned each session. A Random ratio .2 (RR .2) condition was followed by Different Draw 25% (DD 25), then by a Different Draw 45% (DD 45) and Different Line 95% (DL 95) reversal.
**Figure 8.** Results of variability of responses by Doris. The top panel presents the AED of drawings across sessions. The middle panel presents the AED of lines across sessions, and the bottom panel presents the number of tokens earned each session. A Random ratio .2 (RR .2) condition was followed by Same Draw 0 (SD 0), then by a Different Draw 30% (DD 30) and Different Line 85% (DL 85) reversal.
Figure 9. The figures in the left column display the AED of drawings for the final two sessions of each condition and the average of number of tokens delivered within each condition for each participant. For Reggie, only one session of DD 25 was conducted. The figures in the right column display the AED of drawings for fifth and sixth session and the average number of tokens (from sessions 1-6) within each condition. For figures in the right column, RR .5 conditions were omitted for all participants. For Paul, the first DL 90 condition consisted of four sessions and was omitted. For Reggie the DD25 session was omitted. Numbers within parentheses represent first or second implementation of the condition (for reversal).
Figure 10. Results of the pre-intervention and post-intervention Repetitive Behavior and Rigid Routine Questionnaire for each factor and participant.
CHAPTER FOUR:

DISCUSSION

The present study examined effects of two modified percentile schedules of reinforcement on the variability of drawing topography. The largest increases in variability occurred when contingencies favored novel drawings, rather than the components of drawings. Surprisingly, reinforcing variability in lines resulted in decreases in variability for both lines and drawings. Similar to previous literature (Goetz, & Baer, 1977; Pryor, Haag, & O’Reilly, 1969; Ross & Neuringer, 2002) the results of this study support that variability of the topography of responses can be reinforced. Previous studies using automated assessments to measure variability (e.g., Ross & Neuringer, 2002) have determined if a response drawn on a computer was different to previous responses or not previously exhibited, but did not measure by how much the responses were different (topographically). By measuring the ED of lines and drawing, it was possible to deliver tokens contingent on how different a response looked in comparison to the history of responses as opposed to how long it had been since that response was last emitted. Thus, this study extends the use of percentile schedules in a novel way to reinforce response variability to drawings.

Across all of the participants, Diff-Line conditions resulted in higher AEDs of drawings in comparison to RR conditions for the participants; however, it seems that Diff-Draw was more effective in increasing the AED of both drawings and lines than Diff-Line. We hypothesized that
during the Diff-Line condition; there would be an increase in AED of line. However, for each participant, the AED of lines was typically higher during Diff-Draw in comparison to Diff-Line. One possibility for this effect might be that token delivery contingent on varied lines increased the number of lines produced by the participant. In our study, the number of lines was free to vary for each drawing. Thus, participants might have been able to increase the number of tokens they received in each session simply by creating drawings with more lines. Figure 11 (See page 35) displays the average number of lines per drawing in the Diff-Draw and Diff-Line for each participant. These data show that in fact each participant drew more lines per drawing on average in Diff-Line than in Diff-Draw. The increase in number of lines might have been related to a decrease in the average of the ED for lines – one easy way to draw lots of lines would be to make a few quick and similar strokes on the drawing area. Future researchers might investigate whether similar effects are obtained when limits are placed on the number of lines per drawing.

A characteristic of the ED measurement used in this study is that the number assigned to each drawing is not just a function of that drawing, but of the relationship between that drawing and the previous 20. Said another way, each image’s value changes based on the relationship to those occurring before it. Some might view this as problematic because the same drawing could appear in two separate conditions but have very different EDs because of the comparison drawings differ. One possible way to address this would be to use a standardized pool of comparison drawings (e.g., those drawn during baseline). However, participants might simply learn to draw one thing over and over that just happened to differ from those drawings in the reference set under such circumstances. As the comparison is always changing with the measurement used in this study, the participants learned to vary from previous responses that were always changing based on the participants history of responding. U-values, a dependent
variable found in other studies, has similar properties to AED as used here because both are calculated based on recent performance (e.g., Miller & Neuringer, 2000; Neuringer, 1986; Ross & Neuringer, 2002). Thus, there is precedent for using measures that do not necessarily reflect characteristics of a single performance.

As the proposed standard definition of creativity suggests a creative response is both different and useful (Runco & Jaeger, 2012), future research could further examine the usefulness along with the variability of responses. This study attempted to set a parameter on usefulness via the instructions provided before each session: “Draw some flowers;” however, none of the participants made any responses (across conditions) that resembled flowers to the trainer. Future research could recruit participants of varying artistic experiences and set specific criteria for response to meet to qualify as useful before while placing contingencies on variability.

As a flashing green background signaled token delivery, it is possible that the green flashing background could have acquired secondary conditioned reinforcing properties. Future research could examine if tokens could be faded if the signal does become a secondary conditioned reinforcer.

Two of three participants’ teachers reported seeing decreases in stereotypy, while the third reported no change. Also, two of three participants’ teachers reported less rigidity and adherence to routines, while the third report an increase. Although the generality of these results is unclear, future research should extend these findings to see the degree to which reinforcing variability can collateral effects in core symptoms related to ASD.

By measuring the degree of differences between drawings, this study allowed for empirical testing of a dimension of creativity. While there may never be agreement on what is
creativity, what components make up creativity, or what components of are most important to creativity, it seems that “differentness” plays a role in being creative. The conclusion made by Goetz and Baer (1973) is appropriate to this study: “the definition of ‘creativity’ is no less arbitrary than it has ever been, but one facet of arbitrariness has been subjected to experimental analysis.”
Figure 11. Average number of lines per drawing between Diff-Draw and Diff-Line for each participant.
REFERENCES


**Appendix A: Repetitive Behavior and Rigid Routine Questionnaire**

**Repetitive Behavior and Rigid Routine Questionnaire**

Adapted from: Repetitive Behaviour Questionnaire-2 (Leekam et al., 2007)

<table>
<thead>
<tr>
<th>Does Your Student:</th>
<th>Never or rarely</th>
<th>One or more times daily</th>
<th>15 or more times daily (or at least once an hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Arrange toys or other items in rows or patterns?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2. Repetitively fiddle with toys or other items? (e.g. spin, bang, tap, twist or flick anything repeatedly?)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3. Spin him/herself around and around?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4. Rock backwards and forwards, or side to side, either when sitting or standing?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5. Pace or move around repetitively? (e.g. walk to and fro across a room, or around the same path?)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6. Make repetitive hand and/or finger movements? (e.g. flap, wave, or flick his/her hands or fingers repetitively?)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Does your student:</th>
<th>Never or rarely</th>
<th>Mild or occasional</th>
<th>Marked or notable</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Have fascination with specific objects?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8. Like to look at objects from particular or unusual angles?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9. Have any special objects he/she likes to carry around? (e.g., a teddy, a blanket, or a book)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10. Collect or hoard items of any sort?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11. Insist on things at home remaining the same? (e.g., furniture being arranged in certain ways.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12. Get upset about minor changes to objects?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13. Insist that aspects of daily routine must remain the same?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14. Insist on doing things in a certain way or redoing things until they are “just right?”</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15. Play the same music, game or video, or read the same book repeatedly?</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Appendix B: Repetitive Behavior and Rigid Routine Questionnaire Scoring.

Repetitive Behavior and Rigid Routine Questionnaire Scoring Instructions

Scores can be summarized into three factors. Scores are disturbed as follows:

**Factor 1- Repetitive Motor Movements.** Includes items: 2, 3, 4, 5, 6.
**Factor 2- Rigidity/Adherence to Routine.** Includes items: 11, 12, 13, 14, 15.
**Factor 3- Preoccupation with Restricted Patterns of Interest.** Includes items: 1, 7, 8, 9, 10, 15.

Mean responses for the three subscales are obtained by adding up scores for each factor and dividing by the number of items within that factor.
Appendix C: Treatment Fidelity Checklist

TREATMENT FIDELITY CHECKLIST

<table>
<thead>
<tr>
<th>#</th>
<th>YES</th>
<th>NO</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the user interface active (online connection, mouse responsive)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the computer set up before trainer asks participant to participate?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you ask for the participant if they want to sit and draw?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants Response: ☐ YES ☐ NO  *If response is NO, terminate session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the database registering lines drawn?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you provide vocal instructions for each trial when signaled by the computer?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you provide tokens each time signaled by the computer?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you provide choice of 3 minutes to backup reinforcers or continuing on to next session?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you provide 3 minutes to backup reinforcers when chosen by participant?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you prompt to the next trial, if the participant terminated the trial?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL: | Out of |

Percentage: %
Appendix D: USF IRB APPROVAL LETTER

September 30, 2015

Harvey Bayliss  
ABA-Applied Behavior Analysis  
Tampa, FL  33612

RE: Expedited Approval for Initial Review  
IRB#: Pro00023429  
Title: Reinforcement of Variability and the Implications for Creativity

Study Approval Period: 9/30/2015 to 9/30/2016

Dear Mr. Bayliss:

On 9/30/2015, the Institutional Review Board (IRB) reviewed and APPROVED the above application and all documents contained within, including those outlined below.

Approved Item(s):
Protocol Document(s):
  Variability and Creativity Protocol v.1 9/28/15

*Note, no research activities can begin without submitting the required letter of support and receiving an approval through the Amendment process.

Consent/Assent Document(s)*:
  Parental Consent Form v.1 9/28/15.pdf
  Child Verbal Assent Script V1 9/28/15 (not a stamped document)

*Please use only the official IRB stamped informed consent/assent document(s) found under the "Attachments" tab. Please note, these consent/assent document(s) are only valid during the approval period indicated at the top of the form(s).

It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review
research through the expedited review procedure authorized by 45CFR46.110 and 21 CFR 56.110. The research proposed in this study is categorized under the following expedited review category:

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

This study involving data pertaining to children falls under 45 CFR 46.404 – Research not involving greater than minimal risk.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval via an amendment. Additionally, all unanticipated problems must be reported to the USF IRB within five (5) calendar days.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

Kristen Salomon, Ph.D., Vice Chairperson
USF Institutional Review Board